

UWB Communication Characteristics for Different Distribution of Pedestrian

Chien-Hung Chen¹, Min-Hui Ho², Shu-Han Liao², and Chien-Ching Chiu²

¹Department of Computer and Communication Engineering, Taipei College of Maritime Technology
Shilin, Taipei, Taiwan, R.O.C.

²Department of Electrical Engineering, Tamkang University, Tamsui, Taipei, Taiwan, R.O.C.

Abstract— A comparison of UWB communication characteristics for different distribution of pedestrian is investigated. The impulse responses of these cases are computed by applying shooting and bouncing ray/image (SBR/Image) techniques and inverse Fourier transform. The frequency dependence utilized in the structure on the indoor channel is accounted for in the channel calculation. The bit error rate (BER) performance for UWB indoor communication is calculated. The outage probability for binary antipodal-pulse amplitude modulation (B-PAM) system has been presented. Numerical results have shown that the multi-path effect by pedestrian is an important factor for BER performance. Finally, it is worth noting that in these cases the present work provides not only comparative information but also quantitative information on the performance reduction.

1. INTRODUCTION

UWB technology has received significant interests, particularly after the Federal Communications Commission (FCC)'s Report and Order in 2002 for unlicensed uses of UWB devices within the 3.1–10.6-GHz frequency band [1]. The analysis and design of an UWB communication system require an accurate channel model to determine the maximum achievable data rate, to design efficient modulation schemes, and to study associated signal-processing algorithms [2]. Reference [3] proposes a deterministic propagation model to analyze the channel capacity of a narrowband 2.45-GHz 8×8 MIMO system within a small room for different distribution of pedestrian. However, to the best of our knowledge, there is no paper dealing with the effect of pedestrian on the indoor channel for the UWB communication system. In this paper, a comparison of UWB communication characteristics for different distribution of pedestrian in real environments is investigated. Results of this research provide valuable insights into the BER performance and outage probability in the UWB communication system. In Section 2, a channel modeling and system description is presented. In Section 3, we show the numerical results. Finally, the conclusion is drawn in Section 4.

2. CHANNEL MODELING AND SYSTEM DESCRIPTION

2.1. Channel Modeling

The following two steps are used to calculate the multi-path radio channel.

2.1.1. Frequency Responses for Sinusoidal Waves by SBR/Image Techniques

The SBR/Image method can deal with high frequency radio wave propagation in the complex indoor environments [4, 5]. It conceptually assumes that many triangular ray tubes are shot from the transmitting antenna (TX), and each ray tube, bouncing and penetrating in the environments is traced in the indoor multi-path channel. If the receiving antenna (RX) is within a ray tube, the ray tube will have contributions to the received field at the RX, and the corresponding equivalent source (image) can be determined. By summing all contributions of these images, we can obtain the total received field at the RX. In real environment, external noise in the channel propagation has been considered. The depolarization yielded by multiple reflections, refraction and first order diffraction is also taken into account in our simulations. Note that the different values of dielectric constant and conductivity of materials for different frequency are carefully considered in channel modeling.

2.1.2. Inverse Fast Fourier Transform (IFFT) and Hermitian Processing

The frequency responses are transformed to the time domain by using the inverse Fourier transform with the Hermitian signal processing [6]. By using the Hermitian processing, the pass-band signal is obtained with zero padding from the lowest frequency down to direct current (DC), taking the conjugate of the signal, and reflecting it to the negative frequencies. The result is then transformed

to the time domain using IFFT [7]. Since the signal spectrum is symmetric around DC. The resulting doubled-side spectrum corresponds to a real signal in the time domain. The impulse response of the channel can be written as follows [8]:

$$h_b(t) = \sum_{n=1}^N a_n \delta(t - \tau_n) \quad (1)$$

where N is the number of paths observed at time. $\delta(\cdot)$ is the Dirac delta function. a_n and τ_n are the channel gain and time delay for the n -th path respectively.

2.2. System Block Diagram

The transmitted UWB pulse stream is:

$$x(t) = \sqrt{E_{tx}} \sum_{n=0}^{\infty} p(t - nT_d) d_n \quad (2)$$

where E_{tx} is the average transmitted energy and $p(t)$ is the transmitted waveform. T_d is the duration of the transmitting signal. $d_n \in \{\pm 1\}$ is a B-PAM symbol and is assumed to be independent identically distributed (i.i.d.). To be radiated in an efficient way, however, a basic feature of the pulse is to have a zero dc (direct current) offset. Several pulse waveforms might be considered, provided that this condition is verified. Gaussian derivatives are suitable. Actually, the most currently adopted pulse shape is modeled as the second derivative of a Gaussian function. The second derivative Gaussian waveform $p(t)$ can be described by the following expression:

$$p(t) = \frac{d^2}{dt^2} \left(\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{t^2}{2\sigma^2}} \right) \quad (3)$$

where t and σ are time and standard deviation of the Gaussian wave, respectively.

The average transmit energy symbol E_{tx} can be expressed as

$$E_{tx} = \int_0^{T_d} p^2(t) dt \quad (4)$$

Block diagram of the simulated communication system is shown in Fig. 1. The received signal $r(t)$ can be expressed as follows:

$$r(t) = [x(t) \otimes h_b(t)] + n(t) \quad (5)$$

where $x(t)$ is the transmitted signal and $h_b(t)$ is the impulse response of the channel, $n(t)$ is the white Gaussian noise with zero mean and variance $N_0/2$. The correlation receiver samples the received signal at the symbol rate and correlates them with suitably delayed references given by

$$q(t) = p(t - \tau_1 - (n - 1)T_d) \quad (6)$$

where τ_1 is the delay time of the first wave. The output of the correlator at $t = nT_d$ is [9, 10]

$$Z(n) = \int_{(n-1)T_d}^{nT_d} \left\{ \left[\sqrt{E_{tx}} \sum_{n=0}^{\infty} p(t - nT_d) d_n \right] \otimes h_b(t) \right\} \cdot q(t) dt + \int_{(n-1)T_d}^{nT_d} n(t) q(t) dt = V(t) + \eta(t) \quad (7)$$

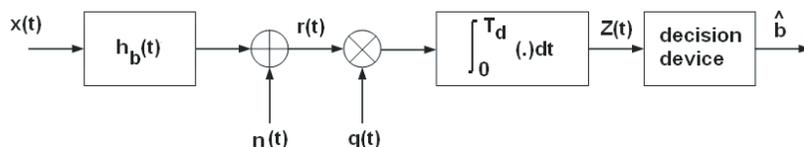


Figure 1: Block diagram of the simulated communication system.

It can be shown that the noise components $\eta(t)$ of Eq. (7) are uncorrelated Gaussian random variable with zero mean. The variance of the output noise η is

$$\sigma^2 = \frac{N_0}{2} E_{tx} \tag{8}$$

The conditional error probability of the N th bit is thus expressed by:

$$P_e[Z(n)|\vec{d}] = \frac{1}{2} \operatorname{erfc} \left[\frac{V(n)}{\sqrt{2}\sigma} \cdot (d_N) \right] \tag{9}$$

where $\operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-y^2} dy$ is complementary error function and $\{\vec{d}\} = \{d_0, d_1, \dots, d_N\}$ is the binary sequence.

Finally, the average BER for B-PAM IR UWB system can be expressed as

$$\text{BER} = \sum_{n=1}^N P(\vec{d}) \cdot \frac{1}{2} \operatorname{erfc} \left[\frac{V(n)}{\sqrt{2}\sigma} \cdot (d_N) \right] \tag{10}$$

where $P(\vec{d})$ is the occurring probability of the binary sequence \vec{d} .

3. NUMERICAL RESULTS

The channel characteristics for different distribution of pedestrian in the indoor environments are investigated Fig. 2 is the top view of indoor environment with dimensions of 10 m (Length) \times 10 m (Width) \times 4.5 m (Height).

There are four different distribution of pedestrian considered in the simulation. Four different numbers of pedestrian with 0, 4, 12 and 36 are simulated. The transmitting and receiving antenna are modeled as a UWB antenna with simple omni-directional radiation pattern and vertically polarized. The transmitting antenna is located at Tx (5, 5, 4) m with the fixed height of 4 m which is located in the center of the indoor environment, as shown in Fig. 2(a). There are 361 receiving points for indoor environment. The locations of receiving antennas are distributed uniformly with a fixed height, 1 m. The distance between two adjacent receiving points is 0.5 m. The maximum number of bounces is set to be seven and the first order diffraction is also considered in the simulation. In the Fig. 2(a), there are four pedestrians in the position marked as A where each A represents one pedestrian. Similarly, there are twelve pedestrians on the position which is as marked B. Finally, there are thirty-six pedestrians on the position A, B and C. Fig. 2(a) shows the static stand pedestrians and Fig. 2(b) shows the pedestrians moving one step randomly. Fig. 2(c) shows the pedestrians moving randomly with many steps, and where applicable, pedestrian random movement. UWB channel characteristics in the indoor environment with pedestrian random movement are investigated.

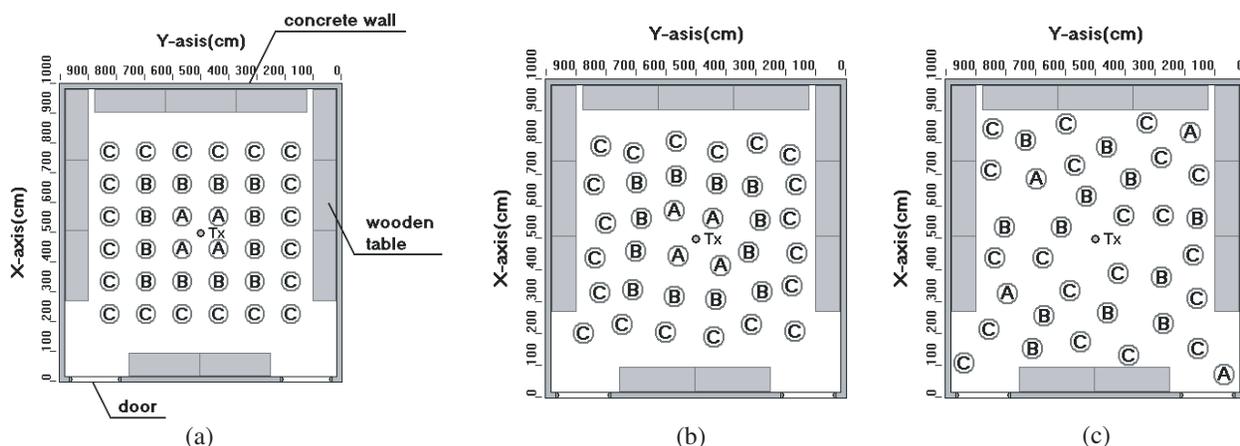


Figure 2: Top view of the indoor environment with dimensions. Marks A, B and C are the positions of the pedestrian. (a) Static stand pedestrians, (b) pedestrians moving one step randomly, (c) pedestrians moving randomly with many steps.

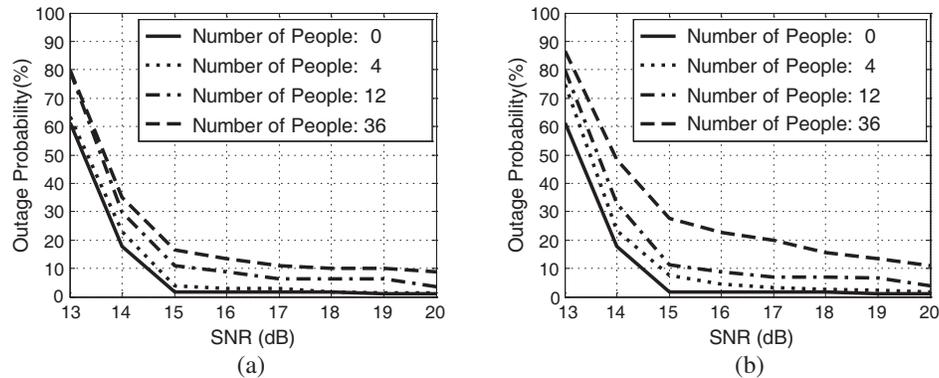


Figure 3: Outage probability versus SNR. (a) Pedestrians moving one step, (b) pedestrians moving randomly with many steps.

By using the impulse responses of these multi-path channels, the bit error rate (BER) performance for binary pulse amplitude modulation (BPAM) impulse radio UWB communication system are calculated. Based on the BER performance, the outage probability for given 361 receiving locations of the transceiver can be computed. Outage probability statistical characteristic taken into account over all R_x location. At 100 Mbps transmission rate and for a $\text{BER} < 10^{-6}$ the outage probability versus SNR are calculated, as shown in Fig. 3. The SNR is defined as the ratio of the average power to the noise power at the front end of the receiver. Fig. 3(a) shows outage probability for different distribution of pedestrians moving one step randomly. It is seen that the outage probabilities at $\text{SNR} = 15$ dB are about 18% and 2% respectively for the 36 moving pedestrian and with on moving pedestrian. It is clear that the BER performance for without pedestrian is better due to the less severe multi-path effect.

In Fig. 3(b) shows outage probability for different distribution of pedestrian random movement. It is seen that the outage probabilities at $\text{SNR} = 15$ dB are about 28% and 2% respectively for the 36 moving pedestrian and with on moving pedestrian. The outage probabilities at $\text{SNR} = 15$ dB pedestrian moving one step is 18% and pedestrians moving randomly with many steps increases about 55% to 28% for the 36 pedestrian. It is seen that the values of the parameters change a lot by comparing the results in Fig. 3(a) and Fig. 3(b).

4. CONCLUSION

A method for analyzing and calculating the channel statistical characteristics of UWB indoor communication systems has been presented. A realistic complex environment is simulated in this paper. A comparison of UWB communication characteristics for different distribution of pedestrian are presented. We analyze the static stand pedestrians and pedestrians moving one step and pedestrians moving randomly with many steps. The outage probabilities for 100 Mbps B-PAM and for a $\text{BER} < 10^{-6}$ versus SNR are calculated. Numerical results show that the values of outage probability increase as the number of pedestrian increase. It is clear that the multi-path effect is severe when the number of pedestrian increases. It is found that value of the parameters change a lot for 36 pedestrian by comparing the results in Fig. 3(a) and Fig. 3(b). The pedestrian's random movement positions, directly increasing the multipath scattering presented in the environment. It is found that the outage probability for the 36 pedestrian is the largest due to the strong pedestrian random movement. The performance of outage probability with pedestrian is worse than that without pedestrian in UWB environment. This is can be attributed to multi-path effect which is severe when pedestrian exist in the room. Finally, it is worth noting that in these cases the present work provides not only comparative information but also quantitative information on the performance reduction.

ACKNOWLEDGMENT

The authors would like to acknowledge the many helpful suggestions of anonymous reviewers.

REFERENCES

1. Nasr, K. M., "Hybrid channel modelling for ultra-wideband portable multimedia applications," *IET Microwaves Antennas and Propagation*, Vol. 2, No. 3, 229–235, Apr. 2008.

2. Irahauten, Z., H. Nikookar, and G. J. M. Janssen, “An overview of ultra wide band indoor channel measurements and modeling,” *IEEE Microwave and Wireless Components Letters; IEEE Microwave and Guided Wave Letters*, Vol. 14, 386–388, Aug. 2004.
3. Ziri-Castro, K. I., W. G. Scanlon, and N. E. Evans, “Prediction of variation in MIMO channel capacity for the populated indoor environment using a radar cross-section-based pedestrian model,” *IEEE Transactions on Wireless Communications*, Vol. 4, 1186–1194, May 2005.
4. Chen, S. H. and S. K. Jeng, “An SBR/image approach for indoor radio propagation in a corridor,” *IEICE Trans. Electron.*, Vol. E78-C, 1058–1062, 1995.
5. Chen, S. H. and S. K. Jeng, “SBR image approach for radio wave propagation in tunnels with and without traffic,” *IEEE Trans. Veh. Technol.*, Vol. 45, 570–578, 1996.
6. Oppermann, I., M. Hamalainen, and J. Iinatti, *UWB Theory and Applications*, John Wiley & Sons, 2004.
7. Kamen, E. W. and B. S. Heck, *Fundamentals of Signals and Systems Using the Web and Matlab*, Prentice-Hall, 2000.
8. Gabriella, M., D. Benedetto, and G. Giancola, *Understanding Ultra Wide Band Radio Fundamentals*, Prentice Hall, 2004.
9. Homier, E. A. and R. A. Scholtz, “Rapid acquisition of ultra-wideband signals in the dense multi-path channel,” *IEEE Conference on Ultra Wideband Systems and Technologies*, 105–109, 2002.
10. Gargin, D. J., “A fast and reliable acquisition scheme for detecting ultra wide-band impulse radio signals in the presence of multi-path and multiple access interference,” *2004 International Workshop on Ultra Wideband Systems*, 106–110, May 2004.

Design of a 1.575 GHz Helical LTCC Chip Antenna for GPS Application

Tao Huang and Yali Qin

Key Lab of Fiber Telecommunication, College of Information Engineering
Zhejiang University of Technology, Hangzhou, China

Abstract— Based on the characteristics of high performance and miniaturization of Low Temperature Co-fired Ceramic (LTCC) technology, this paper presents the design simulation of a miniaturized helical LTCC chip antenna, the helical radiating patch is realized by using via holes to connect upper- and lower-layer radiating elements which are in staggered arrangement. By employing the Ansoft HFSS 3-D EM simulator to design and simulate, the antenna could be got with the central frequency 1.575 GHz, bandwidth 60 MHz and size $10\text{ mm} \times 3\text{ mm} \times 0.77\text{ mm}$. The simulation results show the antenna completely meet the needs of GPS and other wireless communication products with small volume and high integrability.

1. INTRODUCTION

The exploding growth of wireless communications systems leads to an increasing demand for integrated compact low-cost antennas, to satisfy these needs, Low Temperature Cofired Ceramic (LTCC) technology has been extensively studied by many researchers due to its good RF properties, its excellent hermeticity, its flexibility in realizing an arbitrary number of layers [1, 2]. The most direct way of reducing the dimensions of the antenna is the use of very high dielectric constant of the LTCC material. However, in the view point of the antenna, the high dielectric constant causes serious problems to the design of antennas with high gain and good radiation pattern, especially, for endfire antenna. Also, another concept to reduce the size of a microstrip patch antenna at a given resonant frequency is to increase its electrical surface-current length by changing the radiating element's shape, such as utilization of a meandering line structure and helical structure [3, 4].

In this paper, a miniaturized helical LTCC chip antenna is proposed for GPS application, the Ansoft HFSS 3-D EM simulator, based on the finite-element method (FEM), is employed for design simulation [5]. The simulation results for return loss and radiation pattern are presented.

2. PRINCIPLE OF HELICAL LTCC CHIP ANTENNA

Figure 1(a) shows the geometry of the proposed structure, the helical radiating patch is realized by using via holes to connect upper- and lower-layer radiating elements which are in staggered arrangement. Figure 1(b) shows the radiating element of helical LTCC chip antenna, the currents on the horizontal segments (parallel to Y axis) of the radiating element have same direction and, hence, they have main contribution to the desired radiation field. On the other hand, the directions of currents on any two symmetrically vertical segments (parallel to X axis) and the via hole segments (parallel to Z axis) are opposite, which have much lower contribution to the desired radiation field. The lowest resonance of the helical LTCC chip antenna occurs when the total length in the Y -direction is approximately a quarter of waveguide wavelength, but for a fixed total length in the X -direction and Z -direction, if the total length in the Y -direction is less than approximately one eighth of waveguide wavelength, the antenna gain and other antenna performances will all become worse with the decrement of the spacing between the radiating element, so blindly decrease in the physical size of antennas shouldn't be the sole aim for designing antennas, a variety of factors must be integrated to seek for the best balance point between the physical size and the electric performance parameter when we design the LTCC antenna.

3. PROPOSED ANTENNA DESIGN

Figure 2 shows the geometry of the proposed stacked helical LTCC chip antenna mounted FR-4 substrate, the dimension of the FR-4 circuit board is $53\text{ mm} \times 20\text{ mm} \times 0.77\text{ mm}$ and the size of the ground plane is $40\text{ mm} \times 20\text{ mm}$. The antenna occupies a small volume of $10\text{ mm} \times 3\text{ mm} \times 0.77\text{ mm}$ and is fabricated in a ceramic substrate that has a relative dielectric constant of 27, the upper- and lower-layer radiating elements are in staggered arrangement, the dimension of the radiating element is $2.54\text{ mm} \times 0.3\text{ mm}$, the spacing between every Radiating Element is 0.67 mm , the height of the via hole which connect upper- and lower-layer radiating element is 0.7 mm . A $50\ \Omega$ microstrip

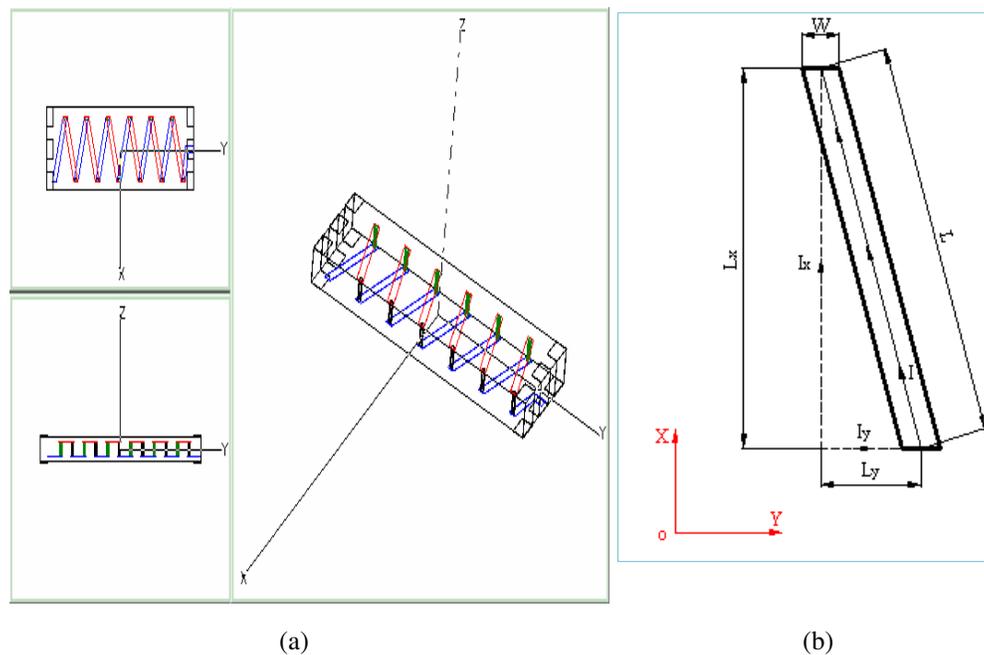


Figure 1: (a) Geometry of helical LTCC chip antenna in the Ansoft HFSS. (b) Radiating element of helical LTCC chip antenna.

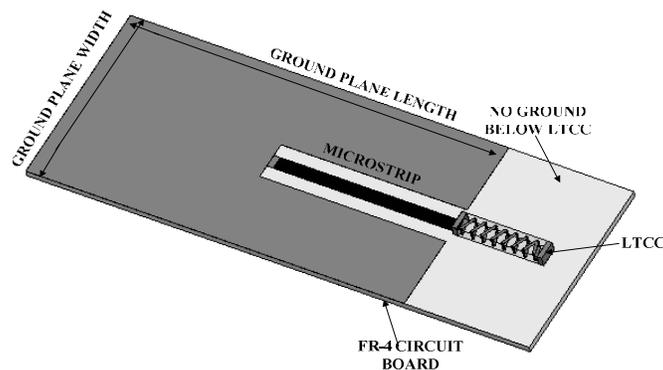


Figure 2: Geometry of the proposed antenna mounted FR-4 substrate.

line is utilized to excite the compact chip antenna, the impedance matching is obtained mainly by optimizing the structure of the helical line and junction part connected to the feed line.

4. SIMULATION RESULTS

Figure 3 shows the simulated return loss for various spacing between every radiating element, it is noted that as it increases, the resonant frequency decreases. It can be seen that as the spacing between every Radiating Element is 0.67 mm (the red line in Figure 3), the resonant frequency of the antenna is 1.575 GHz and a 10 dB return loss bandwidth is about 60 MHz (1.54 GHz–1.60 GHz). Figure 4 shows the simulated return loss for various via-hole height, as can be seen, the resonant frequency of the helical antenna decreases with the increment of the via-hole height. The radiation patterns of the antenna at the 1.575 GHz are shown in Figure 5, if the antenna is vertically oriented (x - z plane), the maximum gain of -1.201 dBi is obtained. Furthermore, the maximum gains are -1.321 dBi and -1.316 dBi, respectively, in x - y plane and y - z plane, these patterns are good enough to be used for GPS application. Figure 6 shows the simulated 3D radiation pattern of the proposed antenna for 1.575 GHz.

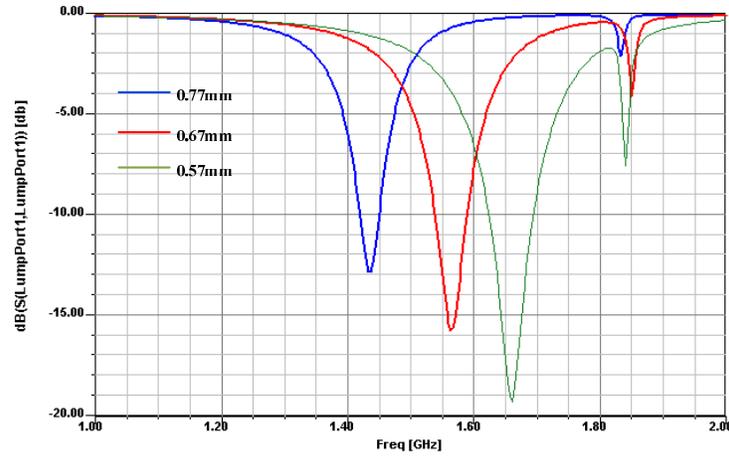


Figure 3: Simulated return loss for different values of spacing between radiating element.

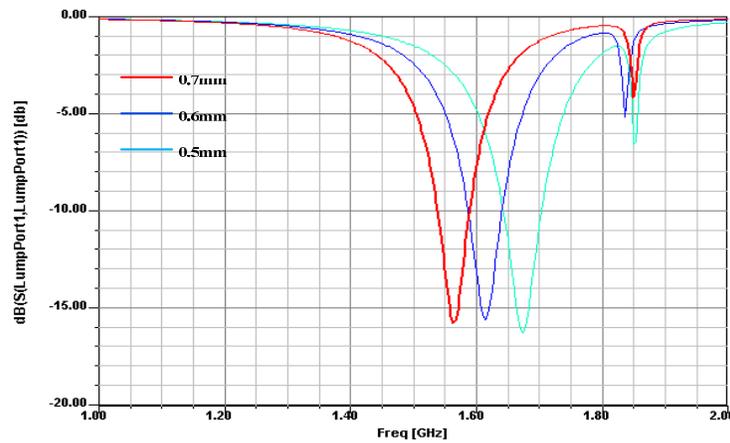


Figure 4: Simulated return loss for different values of the via-hole height.

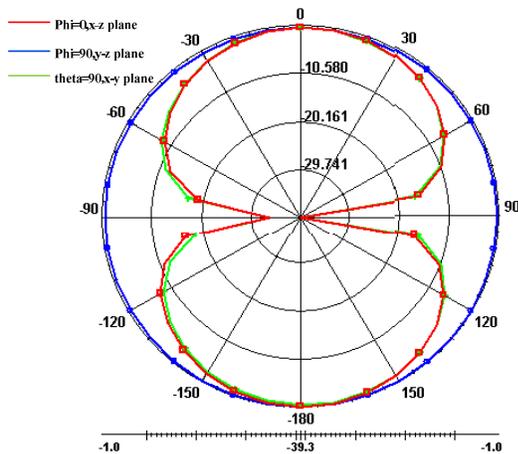


Figure 5: Radiation patterns of the proposed antenna at 1.575 GHz.

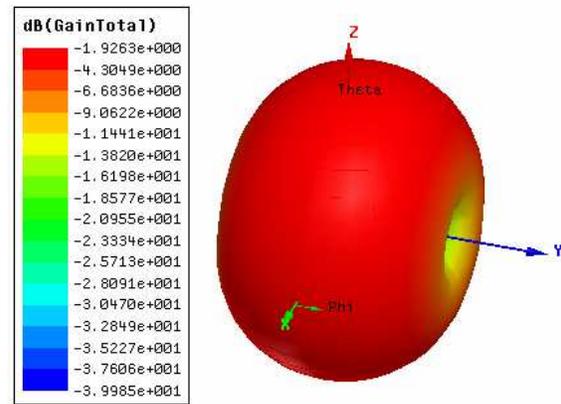


Figure 6: The simulated 3D radiation pattern of the proposed antenna at 1.575 GHz.

5. CONCLUSIONS

The design simulation of a 1.575 GHz helical LTCC chip antenna using LTCC multi-layer technology is presented. The required resonant frequency and return loss bandwidth was achieved by controlling the spacing between every radiating element and the height of via-hole, in order to meet the better requirements for GPS applications, increasing the bandwidth and gain are the goals for continuing the study.

REFERENCES

1. Sheen, J. W., "LTCC-MLC Duplexer for DCS-1800," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 47, No. 9, 1883–1890, Sept. 1999.
2. Seki, T., K. Nishikawa, and K. Cho, "Multi-layer parasitic microstrip array antenna on LTCC substrate for millimeter-wave system-on-package," *33rd European Microwave Conference*, Vol.3, 1393–1396, Oct. 2003.
3. Wang, H. Y. and M. J. Lancaster, "Aperture-coupled thin-film superconducting meander antennas," *IEEE Transactions on Antennas and Propagation*, Vol. 47, No. 5, 829–836, May 2002.
4. Haapala, P., P. Vainikainen, and P. Eratuuli, "Dual frequency helical antennas for handsets," *IEEE 46th Vehicular Technology Conference*, Vol. 1, 336–338, 1996.
5. Ansoft high frequency structure simulator (HFSS) Version 10.0, Ansoft Corporation, 2005.

A Compact Band Notched UWB Antenna for Mobile Applications

Nam-I Jo, Dang-Oh Kim, and Che-Young Kim

School of Electrical Engineering and Computer Science, Kyungpook National University
Sankyuk-dong Puk-gu, Daegu 702-701, South Korea

Abstract— In this paper, a compact circular UWB (ultra-wideband) antenna with WLAN 802.11a/n band (5.15~5.825 GHz) rejection characteristic is proposed for the mobile applications. The proposed UWB antenna was compact enough to be loaded into a mobile handset. The performance of the suggested antenna is mostly dependent on the gap between circular radiator and the ground plane. In order to have the band notched operation, the CSRR (complementary split ring resonator) is etched inside the circular patch of the UWB antenna.

MWS (Microwave Studio) of CST company was utilized in the design stage. The antenna was constructed on a substrate, Rogers 4003, with the thickness of 0.8 mm and relative permittivity of 3.38. A size of substrate is $24 \times 37.5 \text{ mm}^2$ and a radius of circular radiator is 5 mm long. The proposed antenna covers the frequency band from 3.2 to 10.6 GHz ($S_{11} \leq 10 \text{ dB}$), and the band rejection occurs about 5.7 GHz band. A simulation shows the maximum gain of 5.2 dBi, and radiation pattern is nearly omni-directional over the entire -10 dB return loss bandwidth. It also represents the group delay below 1nsec over UWB communication band (3.2 GHz ~ 10.6 GHz) except WLAN band. Based on the experimental results, the proposed antenna could be a good candidate for the handheld mobile handsets.

1. INTRODUCTION

Since the Federal Communications Commission (FCC) released the ultra-wideband communication band 3.1~10.6 GHz in 2002, the various type of UWB antenna has been developed for use in the wireless communication system [1]. One popular application of the UWB communication is the field of mobile multimedia system due to its high data rate. However, the high data rate demands the wide bandwidth, inevitably losing stringent protection from the interference such as the wireless LAN (WLAN) colocated band at 5.2 GHz/5.8 GHz [2]. Thus it is essential to have a wide bandwidth together with the specified band stop. Some published UWB antennas reveal a good behavior for a relaxed size. However, they are too cumbersome to be integrated into smaller sized mobile device. In other words, it necessitates extra space as a ground and circuit plane to be fitted into the real sized mobile handset. Hence, a key challenge in the design of mobile UWB antenna is adapting both to a wide bandwidth with rejection band and to a smaller dimension. In this paper, a compact mobile UWB antenna with band notched characteristic is proposed. Band notched characteristic is achieved by embedding complementary split ring resonator (CSRR) slots on the radiator. As compared with the published antennas in [3] and [4], the size of the proposed antenna is miniaturized by more than about 50% in the similar performance.

2. ANTENNA DESIGN

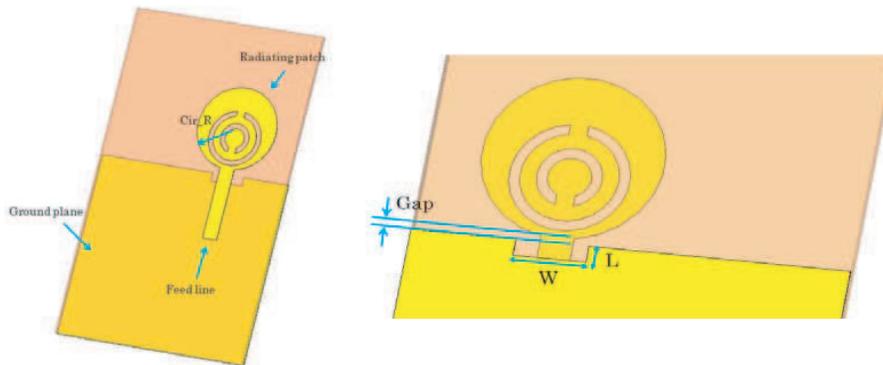
Figure 1(a) and the left antenna in Figure 1(b) show the geometry of the proposed UWB antenna. The proposed antenna has been fabricated on a substrate, Rogers 4003, with the thickness of 0.8 mm and relative permittivity of 3.38.

The width of the feed line at input of the proposed antenna is 1.8 mm long providing 50Ω characteristic impedance. As shown in the Figure 1, the antenna consists of a circular patch, CSRR, and modified partial ground plane. By using the Microwave Studio of CST company, the optimized parameters are found to be $C_{ir_R} = 5 \text{ mm}$, $W = 4 \text{ mm}$, $L = 1 \text{ mm}$, and $Gap = 0.58 \text{ mm}$. Figure 1(b) illustrates the size comparison between the proposed (left antenna) antenna and the traditional (right antenna) antenna published. The circular radiator of proposed antenna is about half smaller than that of the published UWB antenna. The size reduction becomes the crucial point in the implementation process of mobile handset.

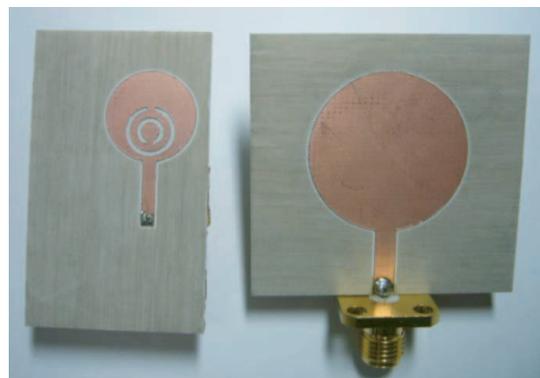
The geometry of the CSRR and the equivalent circuit of the CSRR are shown in Figure 2. The equivalent circuit of the CSRR is a sort of the LC resonant circuits with the inductance and the distributed capacitance [5].

The equivalent circuit model of the Figure 2(b) is the same as the resonant circuit of the band rejection filter connected L, C in parallel. Its resonance frequency can be expressed as [6]

$$\omega_o = \sqrt{\frac{2}{\pi r_o LC}}$$

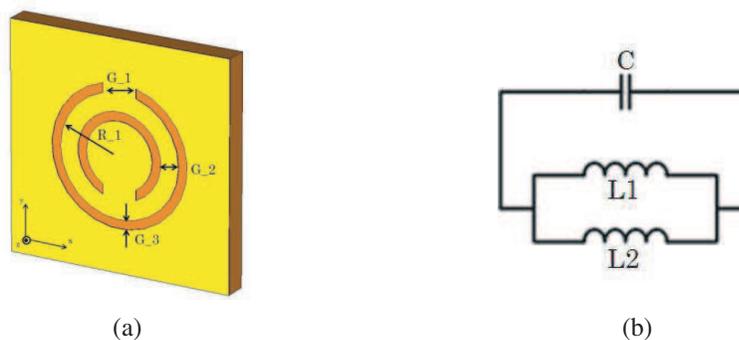


(a)



(b)

Figure 1: The geometry of the proposed UWB antenna, (b) photograph of the proposed UWB antenna (left), and traditional UWB antenna (right).



(a)

(b)

Figure 2: (a) The geometry of the CSRR, (b) the equivalent circuit of the CSRR.

where L is the inductance per unit length between the annular slots, C is the total capacitance of the CSRR, and r_o is the average radius of the two annular slots [6]. The parameters of the CSRR used in proposed antenna are $R_{-1} = 3.3$ mm, $G_{-1} = 1$ mm, $G_{-2} = 1.5$ mm and $G_{-3} = 0.5$ mm.

3. RESULT AND DISCUSSION

Figure 3 represents the simulated and measured return loss curve. The measured return loss is very close to those of simulated results. The proposed antenna has an impedance bandwidth (reference level $S_{11} < -10$ dB) from 3.2 to 10.6 GHz rejecting the 5–6 GHz band. By adjusting the gap between the circular radiator and the ground plane, the wideband could be achieved. In particular, by adjusting W and L of the ground plane, the band above 9 GHz was duly obtained.

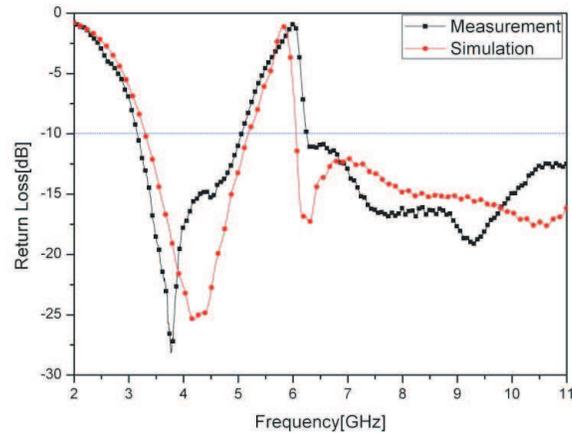


Figure 3: Simulated and measured return losses.

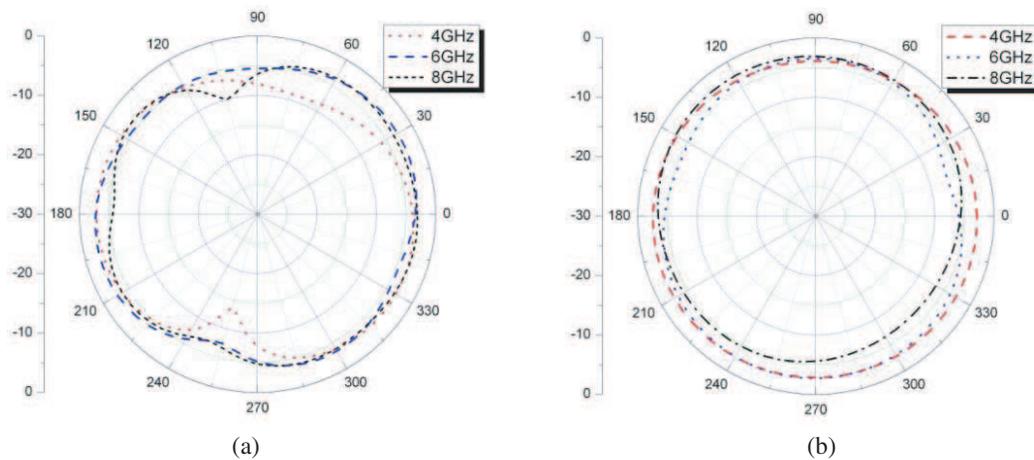


Figure 4: Simulated radiation patterns (a) E -plane, (b) H -plane.

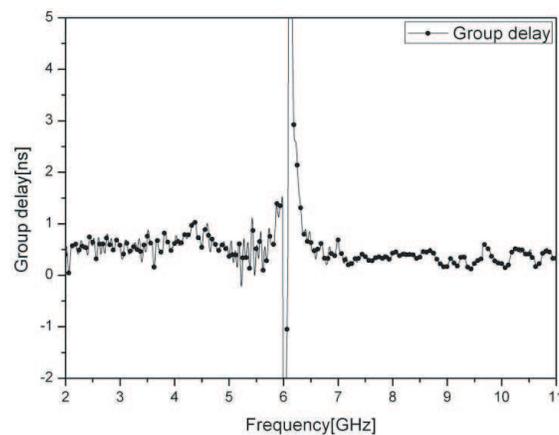


Figure 5: The measured group delay of the proposed antenna.

Figure 4 is the radiation patterns simulated at 4, 6, and 8 GHz, respectively. It can be seen that the antenna exhibits a nearly omni-directional radiation pattern in the H -plane (xz -plane) and a dipole like radiation pattern in the E -plane (yz -plane), which possesses the suitable pattern as a mobile antenna.

The group delay measured by Anritsu 38397C vector network analyzer is plotted in Figure 5. A group delay is an important parameter carrying information on a pulse dispersion. Except for the rejection band, spreading over the range of UWB communication, the group delay variation is less than 1 ns ensuring a proper operation as the mobile UWB antenna.

4. CONCLUSION

This article described the compact band notched UWB antenna for the mobile handsets. The CSRR is used to achieve a band-notched characteristic and the gap between the circular radiator and the ground plane has been introduced to provide the wideband performance. With the noted techniques, the proposed antenna offers wideband (3.2~10.6 GHz) operation and band (5~6 GHz) notched characteristic together with size reduction essential to the mobile phones. These benefits are confirmed by the simulation and the measurement on the fabricated UWB antenna.

REFERENCES

1. Revision of Part 15 of the Communication's Rules Regarding Ultra-Wideband Transmission Systems, Federal Communications Commission, ET-Docket 98-153, FCC 02-48, 2002.
2. Sadat, S., M. Fardis, F. Geran, and G. Dadashzadeh, "A compact microstrip square-ring slot antenna for UWB applications," *Progress In Electromagnetics Research*, PIER 67, 173–179, 2007.
3. Jang, J. W. and H. Y. Hwang, "An improved band-rejection UWB antenna with resonant patches and a slot," *IEEE Antennas and Wireless Propag. Lett.*, Vol. 8, 299–302, 2009.
4. Song, Y, Y. C. Jiao, T. L. Zhang, J. B. Jiang, X. Zhang, and F. S. Zhang, "Frequency notched UWB slot antenna with a fractal-shaped slot," *Journal of Electromagnetic Waves and Applications*, Vol. 23, 321–327, 2009.
5. Baena, J. D., J. Bonache, F. Martin, R. M. Sillero, F. Falcone, T. Lopetegi, M. A. G. Laso, J. Garcia-Garcia, I. Gil, M. F. Portillo, and M. Sorolla, "Equivalent-Circuit models for split-ring resonators and complementary split-ring resonators coupled to planar transmission lines," *IEEE Trans. Microwave Theory Tech.*, Vol. 53, No. 4, 1451–1461, 2005.
6. Marques, R., F. Mesa, J. Martel, and F. Medina, "Comparative analysis of edge and broadside-coupled split ring resonators for metamaterial design — Theory and experiments" *IEEE Trans. Antennas Propag.*, Vol. 51, 2572–2581, Oct. 2003.

Numerical Modeling a Microwave and Detection of Partial Discharge inside of HV Transformer

P. Fiala, E. Gescheidtová, and T. Jirku

Brno, FEEC BUT, UTEE, Kolejní 2906/4, Brno 612 00, Czech Republic

Abstract— The aim of this paper is to present the particulars of new research in the special measurement method modeling of starting process partial discharge inside of high voltage transformer. The numerical analysis of this effects can help to set up conditions to decrease it and get of information for preparing detection apparatuses and measurement methods.

1. INTRODUCTION

One of the problematic conditions in the field of high-voltage technology, apparatuses and devices (machines) consists in the emergence of partial discharges [1]. At this point, let us also note that several other effects have combined with this notion over time [2–4]. In consequence of these effects there emerge short electromagnetic pulses with a defined and measurable spectre in the characteristic frequency band [5]. The group of end products attributable to the emergence of interfering signals involves, for example, displacement current in a dielectric, pulse current on the interface between dielectrics, or the dielectric/metal interface owing to high electric field intensity and structure of the dielectric. In HV and VHV transformers (Fig. 1) the dielectric is mineral or synthetic oil.

Large distribution transformers are constructed in such a manner as to have structural measures facilitating oil purification. Also, these transformers are equipped with sensors indicating the initial stage of increase in pulse activity. In the course of this activity, as is well-known, there occurs an increase in the boundary value of the of the applied dielectric breakdown value. As referred to in the above text, oil is the dielectric. Under certain conditions, however, the separation of chemical compounds incurred by decomposition of the dielectric does not have to occur. Thus, free atoms of carbon, hydrogen and oxygen develop from hydrocarbons, and there also generates a certain percentage of water, other organic compounds, and semiconductive carbon. All of these elements decrease the quality of the dielectric; in addition to that, rapid increase in pulse activity may cause the formation of a hazardous explosive compound of oxygen and hydrogen. Then, this situation may result in a local explosion, damage to the device and reduction of its ability to perform the respective functions (Fig. 2).

This work deals with an analysis of electromagnetic field distribution in a transformer dielectric region. The structural parts enable the placement of sensors, whose structure and concept must be adapted to the characteristics of the configuration in such a manner that, from all components of the device, there is a measurable (indicable) electromagnetic pulse signal.

The analysis will be realized for the minimum required level of an electromagnetic pulse for the discrete values of frequencies from the desired spectral interval. An example will be evaluated of electromagnetic field distribution in the region of critical parts of the device.



Figure 1: Example of a distribution VHV transformer.



Figure 2: Explosion of a VHV transformer: an example.

2. MATHEMATICAL MODEL

It is possible to carry out an analysis of an MG model as a numerical solution by means of the Finite element method (FEM). The electromagnetic part of the model is based on the solution of full Maxwell's equations

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \quad \nabla \times \mathbf{H} = \sigma \mathbf{E} + \frac{\partial \mathbf{D}}{\partial t} + \mathbf{J}_s, \quad \nabla \cdot \mathbf{D} = \rho, \quad \nabla \cdot \mathbf{B} = 0^{v \text{ oblast} \Omega}. \quad (1)$$

where \mathbf{E} and \mathbf{H} are the electrical field intensity vector and the magnetic field intensity vector, \mathbf{D} and \mathbf{B} are the electrical field density vector and the magnetic flux density vector, \mathbf{J}_s is the current density vector of the sources, ρ is the density of free electrical charge, γ is the conductivity of the material and Ω is the definition area of the model. The relationships between the electric and the magnetic field intensities and densities are given by material relationships

$$\mathbf{D} = \varepsilon \mathbf{E}, \quad \mathbf{B} = \mu \mathbf{H}. \quad (2)$$

The permittivity ε , the permeability μ and the conductivity γ in HFM are generally tensors with main axes in the direction of the Cartesian coordinates x, y, z . When all the field vectors perform rotation with the same angular frequency ω , it is possible to rewrite the first Maxwell equation

$$\nabla \times \underline{\mathbf{E}} = -j\omega \mu \underline{\mathbf{H}}, \quad \nabla \times \underline{\mathbf{H}} = (\sigma + j\omega \varepsilon) \underline{\mathbf{E}} + \underline{\mathbf{J}}_s \text{ in } \Omega, \quad (3)$$

where $\underline{\mathbf{E}}, \underline{\mathbf{H}}, \underline{\mathbf{J}}_s$ are field complex vectors. Taking into account boundary conditions given in (1) and after rearranging (3) we get

$$(j\omega)^2 \varepsilon \underline{\mathbf{E}} + \sigma \underline{\mathbf{E}} + \nabla \times \mu^{-1} \nabla \times \underline{\mathbf{E}} = -j\omega \underline{\mathbf{J}}_s. \quad (4)$$

We apply the Galerkin method with vector approximation functions \mathbf{W}_i and use the vector form of the Green theorem on the double rotation element [3]. After discretisation we get the expression

$$-k_0 [M] \{\mathbf{E}\} + jk_0 [C] \{\mathbf{E}\} + [K] \{\mathbf{E}\} = \{\mathbf{F}\}, \quad (5)$$

where $\{\mathbf{E}\}$ is the column matrix of the electric intensity complex vectors. The matrixes $[K], [C]$ and $[M]$ are in the form that is given in manual [4] and the vector $\{\mathbf{F}\}$ is evaluated from the expression

$$\{\mathbf{F}\} = -jk_0 Z_0 \int_{\Omega} [\mathbf{W}_i] \{\mathbf{J}_s\} d\Omega + jk_0 Z_0 \int_{\Gamma_0 + \Gamma_1} [\mathbf{W}_i] \{\mathbf{n} \times \mathbf{H}\} d\Gamma. \quad (6)$$

The vector approximation functions \mathbf{W} are given in manual [4]; k_0 is the wave number for vacuum, Z_0 is the impedance of free space. The set of Equations (5) is independent of time and gives $\underline{\mathbf{E}}$. For the transient vector \mathbf{E} we can write

$$\mathbf{E} = \text{Re}\{\underline{\mathbf{E}} e^{j\omega t}\}. \quad (7)$$

3. THE GEOMETRICAL MODEL

The basic model was built in a computer-aided system like as parametrical model. Then, it was imported to the finite element method (FEM) analysis system and solved mainly in the ANSYS system in order to facilitate finding the basic electromagnetic quantities. Basic geometrical parts of model are shown in the Fig. 3.

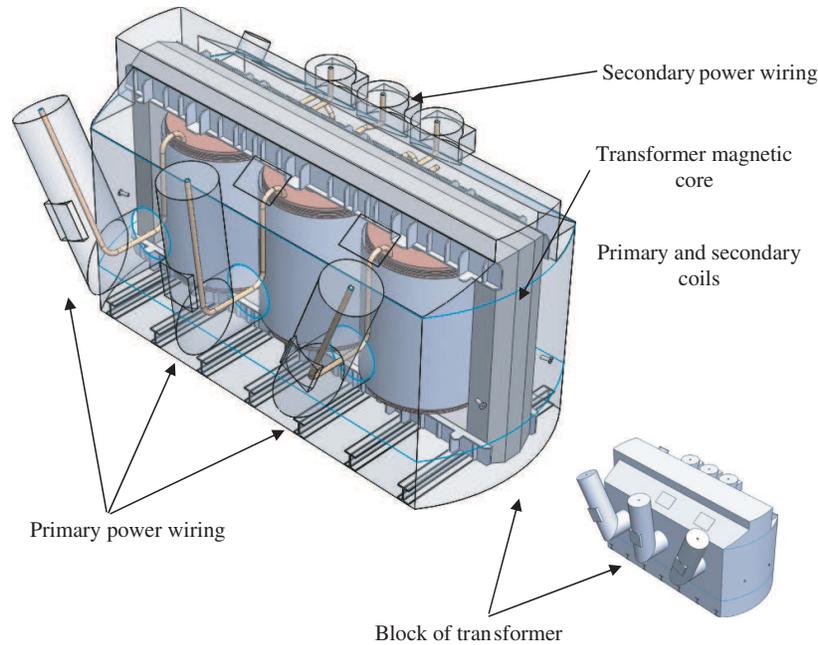


Figure 3: The geometrical tranformer model.

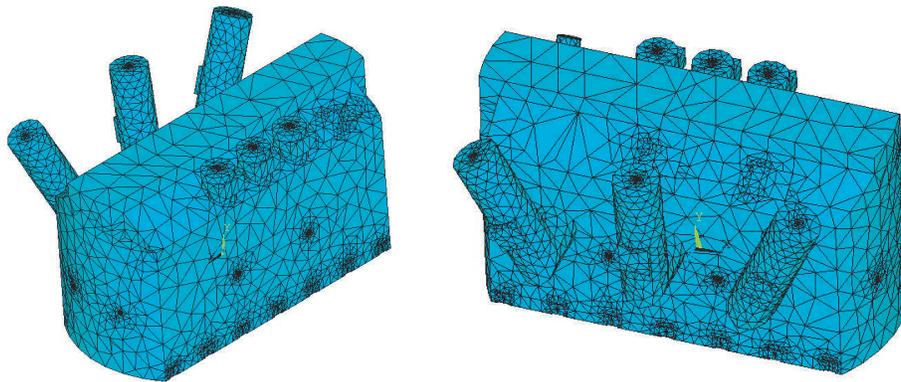


Figure 4: The geometrical/FEM model of the transformer.

4. FEM MODEL

A geometric model, Fig. 4 using HF119, HF120 in ANSYS software was built — Fig. 4. A solution of the HF field model was performed using the APDL program. According to the microwave model solution, HF Block of transformer Secondary power wiring Primary power wiring Transformer magnetic core Primary and secondary coils attenuation inside transformer part is evaluated. The analysis was performed for the frequency interval $f \in (0.5, 1.5)$ GHz, according to Fig. 5 from [5], and the analysis results were prepared for sensor design.

Numerical results can be interpreted in the form a module of vector electric field \mathbf{E} distribution, Fig. 6, in a details of transformer parts. It depends on frequency of analysis. For our model it was from 500 MHz to 3.0 GHz. On this results was chosen type and construction of partial discharge sensor for inside using.

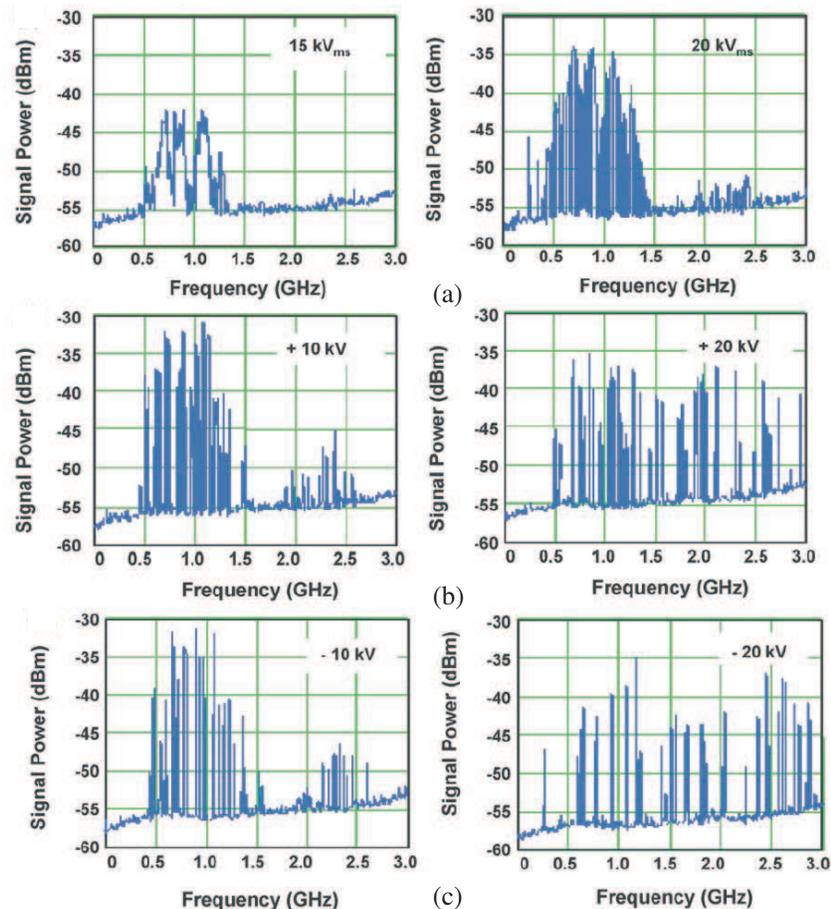


Figure 5: Typical spectrum analyzer display of UHF signals due to particle movement in a transformer oil under AC and DC voltages using a spectrum analyzer. (a) AC; (b) +DC; (c) -DC.

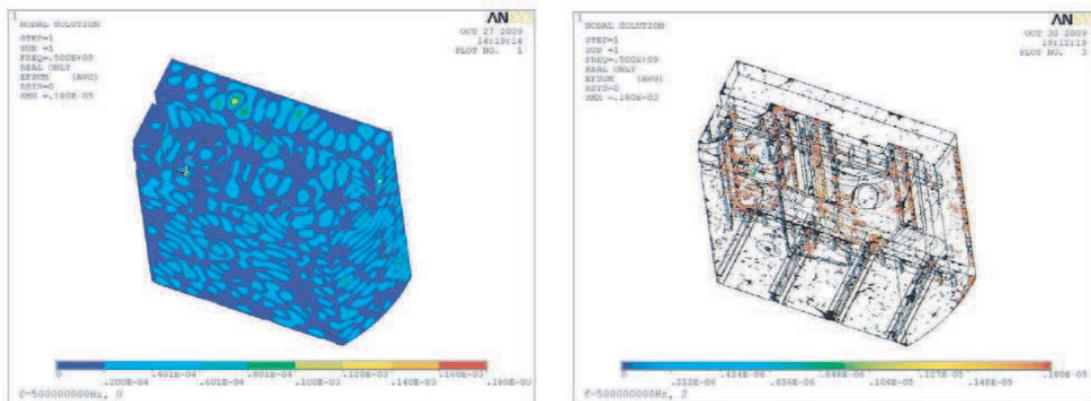


Figure 6: Typical electromagnetic field distribution, modul of electric field vector E , inside of part of transformer model, frequency $f = 500$ MHz,

5. CONCLUSION

The basic research of the numerical model HF wide band signals inside of a VHV transformer has brought a considerable sum of experience in the field of signals and possibilities of their measurement and detection. It was used for sensor design conception and construction.

REFERENCES

1. Kaneiwa, H., Y. Suzuoki, and T. Mizutan, "Characteristics of partial discharges, artificial simulated tree channels during in tree propagation," *IEEE Transactions on Dielectrics and*

- Electrical Insulation*, Vol. 8, No. 1, February 2001.
2. Fiala, P., “Secondary winding model of current transformer-switchable variant,” Research Report, Laboratory of Modeling and Optimization Field in Electromagnetic Systems, FEI VUT and ABB EJV a.s. Brno, No. 1/99, Brno, Czech Republic, Jan. 21, 1999.
 3. Fiala, P., “Transformer partial discharge modeling, minimal breakdown value set in a critical parts of transformer design,” Research Report, Laboratory of Modeling and Optimization Field in Electromagnetic Systems, FEI VUT and ABB EJV a.s. Brno, No. 2/99, Brno, Czech Republic, Mar. 18, 1999.
 4. Fiala, P., “Optimization of two-pole voltage transformer design $U_{max} = 25$ kV, new variant 2000,” Research Report, Laboratory of Modeling and Optimization Field in Electromagnetic Systems, FEI VUT Brno, No. 1/00, Brno, Czech Republic, Feb. 20, 2000.
 5. Sarathi, R., A. J. Reid, and M. D. Judd, “Partial discharge study in transformer oil due to particle movement under DC voltage using the UHF technique,” *Electric Power Systems Research*, Elsevier, Vol. 78, 1819–1825, 2008.
 6. Manuals ANSYS, v.8.1, USA.
 7. Fiala, P., “Coupled electromagnetic model of smoothing choke,” *Proceedings of the 5th AMTEE International Conference*, C05, Pilsen, Czech Republic, Dec. 9, 2001.

Progress in Studies of Radio Frequency Radiation of the Wireless Communication Device

Chaoqun Jiao and Lei Gao

School of Electrical Engineering, Beijing Jiaotong University, China

Abstract— The through reading a lot of literatures, the progress in Studies of Radio frequency radiation of the wireless communication device is presented in this paper. Then, the disadvantage and advantage of these methods are pointed out. Finally, with different situation, it is suggested that which method is chosen to resolve the practical problem.

1. INTRODUCTION

The wide application of wireless communication has attracted a lot of concern about safety of wireless devices [1–3]. Many investigations have been carried out to determine the effects of Radio Frequency (RF) radiation on human body using various exposure scenarios and models [1, 4–6]. Among these investigations, the Specific Absorption Rate (SAR) is considered as an index that quantifies the rate of energy absorption in biological tissues. Dimbylow et al. studied the SAR in human head exposed to a cell phone at 2 mm resolution with the frequency bandwidth ranging from 10 MHz to 3 GHz by using Finite-Difference Time-Domain (FDTD) method [5]. Martinez-Burdalo et al. further analyzed SAR depositions in different-aged human heads exposed to electromagnetic radiation resulted from a mobile phone [6]. Most of these studies have been focused on radiation effects from mobile phones and base stations. The work has been reported on the studies of radiation effects and radiation efficiency of wireless medical devices in interaction with human body [7]. Modern wireless technology is playing a more and more important role in telemedicine. Some wireless sensor platforms for pervasive healthcare monitoring have been designed in order to improve the quality of human life and minimize restrictions on daily activities [8–11]. Most of the related reports only analyzed the radiation efficiency and radiation effects of external sources such as wearable medical sensor devices. In fact, more and more implantable or ingestible devices are clinically employed, such as pacemakers, implantable defibrillators, capsule endoscope, implanted therapeutic devices and so on [12–14]. Recently, the research on biological effects of Ingestible Wireless Device (IWD) has been proposed particularly due to the persistence and high local energy deposition of antennas that are imbedded in human tissues [15–17]. Kim et al. found that the radiated power was the largest when a dipole antenna was located at the center of head model in comparison with other positions [18, 19]. The wireless multi-parameter monitoring of gastrointestinal (GI) tract [20–22], as well as wireless capsule endoscope [23], which allows high-resolution video being transmitted, has been widely applied. Among all the commercially available wireless capsule endoscopes, 434 MHz is the most often used frequency for video transmission [14, 23]. The frequencies of 315 MHz and 400–500 MHz are also used in some publications [24, 25]. At low transmission frequency, the antenna of the IWD is electrically small due to its compact size required for swallowing. Recently, 1.2 GHz and 2.4 GHz are also reported being used for ingestible capsule endoscopes in [26, 27]. Ito et al. evaluated the radiation characteristics of a helical antenna located in a cylindroid whole body phantom at 150 MHz and showed that the gain could be improved 5 dB or more by setting the target of local SAR and controlling the matching condition [28]. Chirwa et al. studied the radiation patterns and near fields of IWD in human body between 150 MHz and 1.2 GHz and suggested that the radiation characteristics of the sources in the GI tract should be optimized for a higher efficiency [29]. Chan et al. estimated the attenuation of a human body trunk at frequencies ranged between 100 MHz and 6 GHz from internal sources using a simplified experimental model [30]. However, the influence of frequency on biological effects and radiation efficiency were not studied. Chirwa et al. analyzed the ingestible sources in five positions at three orientations with the frequency ranging from 150 MHz to 1.2 GHz [29]. Lisheng Xu analyze the worst radiation intensity outside of human body model and, actually, the worst cases should be considered because the radiation characteristic of IWD is frequency, orientation and position dependent. The FDTD method is applied to analyze the electromagnetic interaction of IWD with human body and the bio-heat equation [31] is used to calculate the temperature rise resulting from the SAR deposition.

With rapid increase in the use of mobile phones in enclosed environments such as trains, cars, and elevators, public anxiety over the possibility of RF exposures in these environments exceeding

the basic restrictions of the International Commission on Non-Ionizing Radiation Protection (IC-NIRP) [32] and IEEE guidelines [33] has been growing. It should be noted that according to the guidelines mentioned above, the only scientifically established adverse health effect of RF energy in the human body is thermal, and its safety limits are given based on specific absorption rates (SARs). Discussions of the nonthermal effects of RF exposure in the human body and the interference with implantable medical devices are beyond the scope of this study. The public anxiety over the use of mobile phones in enclosed environments has risen partly because of some reports in academic journals suggesting that electromagnetic field (EMF) levels in these environments exceed the safety levels during mobile phone use [34–36]. Furthermore, the appearance of these reports in well-known international mainstream media such as the British Broadcasting Corporation (BBC) and others has probably fueled these concerns. On the other hand, there have been several attempts by researcher who do not agree with these reports to allay these fears. Several papers have been published on this concept to date [37–41]. After a thorough review of the studies on this subject, we have thus far found very little information, especially on the localized peak spatial-averaged 10-g SAR in a realistic human model in actual enclosed environments. This is by far the most important parameter in reaching any conclusion regarding the thermal effect of RF exposure. Some researchers have investigated this effect [42–44], but only on a passenger's head, while ignoring the rest of the body. The body in an enclosed environment like an elevator would also absorb electromagnetic energy. Furthermore, the head-and-metallic wall configuration does not represent a realistic case of a passenger in an elevator or other enclosed environment, especially when multireflections of the EMF from the walls are to be considered. The lack of spatial SAR information may be partly due to the large-scale computer resources required to perform numerical investigations of a realistic human model in an actual elevator. Recently, a numerical investigation of this problem using a finite-difference time-domain (FDTD) technique was reported [45]; however, the report gives limited information of only the whole-body average SAR. Considering the above factors, A. Y. Simba et al. have previously presented preliminary investigations of this study [46]. The nonuniform mesh FDTD technique [47], which minimizes computer resources by a factor of almost five, and a supercomputer were used to perform detailed calculations of the SAR in an anatomically realistic human model inside an actual elevator at the cellular phone operating frequencies used in Japan (i.e., 900, 1500, and 2000 MHz). Specifically, they address two issues, which are: 1) whether or not the safety guidelines are exceeded in the elevator, by comparing maximum values obtained with the basic restrictions and 2) to provide an understanding of the SAR characteristics of the mobile phone user in such environments under worst case conditions. The investigations in their paper included the effect of the user's position inside the elevator and the influence of elevator's structure, such as its walls and openings [48].

2. COMPATATION OF THESE STUDIES AND SUGGESTIONS

From the literatures, with regard to the different practical problem, the different model and method should be used. For example, as to the common problem, it is enough to consider the radiation effects from mobile phones and base stations. But as to the mobile phones used in enclosed environments such as trains, cars, and elevators, the effect of trains, cars, and elevators to the radiation must be considered. Furthermore, except for the the mobile phones, the Radiation of the other Wireless Communication Device must be considered. Finally, The “worst case” as to the practical problem also be considered such as in [48].

3. CONCLUSION

The progress in Studies of radio frequency radiation of the wireless communication device is presented in this paper. The disadvantage and advantage of these studies are pointed out. It is suggested that which studies is better with regard to the practical problem.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Fund (No: 50807003) and the Research Fund of Beijing Jiaotong University (No. 2007XM044).

REFERENCES

1. Poljak, D., *Human Exposure to Electromagnetic Fields*, WIT Press, 2004.
2. “IEEE standard for safety levels with respect to human exposure to radio frequency electromagnetic fields, 3 kHz to 300 GHz,” IEEE Standard C95.1-1999, Apr. 16, 1999.

3. “IEEE recommended practice for determining the peak spatial-average specific absorption rate (SAR) in the human head from wireless communications devices: Measurement techniques IEEE,” *IEEE Std 1528*, 2003.
4. Gandhi, O. P. and G. Kang, “Some present problems and a proposed experimental phantom for SAR compliance testing of cellular telephones at 835 and 1900 MHz,” *Physics in Medicine and Biology*, Vol. 47, No. 9, 1501–1518, 2002.
5. Dimbylow, P. J., “Fine resolution calculations of SAR in the human body for frequencies up to 3 GHz,” *Physics in Medicine and Biology*, Vol. 47, No. 16, 2835–2846, 2002.
6. Martinez-Burdalo, M., A. Martin, M. Anguiano, and R. Villar, “Comparison of FDTD-calculated specific absorption rate in adults and children when using a mobile phone at 900 and 1800 MHz,” *Physics in Medicine and Biology*, Vol. 49, No. 2, 345–354, 2004.
7. Alomainy, A., Y. Hao, C. G. Parini, and P. S. Hall, “Comparison between two different antennas for UWB on-body propagation measurements,” *Antennas and Wireless Propagation Letters*, Vol. 4, 31–34, 2005.
8. Yang, G. Z., *Body Sensor Networks*, Springer, New York, 2006.
9. Thomas, F. B., “Biomonitoring with wireless communications,” *Annu. Rev. Biomed. Eng.*, Vol. 5, 383–412, 2003.
10. Rosen, A., M. A. Stuchly, and A. V. Vorst, “Applications of RF/microwaves in medicine,” *IEEE Transactions on Microwave Theory and Techniques*, Vol. 50, No. 3, 963–974, Mar. 2002.
11. Hall, P. S. and H. Yang, *Antennas and Propagation for Body-centric Wireless Communications*, Artech House, London, 2006.
12. Virtanen¹, H., J. Keshvari, and R. Lappalainen, “The effect of authentic metallic implants on the SAR distribution of the head exposed to 900, 1800 and 2450 MHz dipole near field,” *Physics in Medicine and Biology*, Vol. 52, 1221–1236, 2007.
13. Mohamad, S., H. Yamu, and C. Jonathan, “Wireless smart implants dedicated to multichannel monitoring and microstimulation,” *IEEE Circuits and Systems Magazine*, Vol. 5, 21–39, 2005.
14. Iddan, G., G. Meron, A. Glukhovsky, et al., “Wireless capsule endoscopy,” *Nature*, Vol. 405, 417–420, 2000.
15. Werber, D., A. Schwentner, and E. M. Biebl, “Investigation of RF transmission properties of human tissues,” *Advance in Radio Science*, Vol. 4, 357–360, 2006.
16. Usui, H., M. Takahashi, and K. Ito, “Radiation characteristics of an implanted cavity slot antenna into the human body,” *Proceedings of IEEE Antennas and Propagation Society International Symposium*, 1095–1098, Jul. 9–14, 2006.
17. Yu, H., G. S. Irby, D. M. Peterson, M-T. Nguyen, G. Flores, N. Euliano, and R. Bashirullah, “A printed capsule antenna for medication compliance monitoring,” *IEE Electronic Letters*, Vol. 43, No. 22, 1179–1181, 2007.
18. Kim, J. and Y. Rahmat-Samii, “An implanted antenna in the spherical human head: SAR and communication link performance,” *Proceedings of IEEE Topical Conference on Wireless Communication Technology*, 202–203, Oct. 15–17, 2003.
19. Yahya, R. S. and J. H. Kim, “Implanted antennas in medical wireless communications,” *Synthesis Lectures on Antennas*, Vol. 1, No. 1, 1–82, 2006.
20. Johannessen, E. A., L. Wang, L. Cui, T. B. Tang, M. Ahmadian, A. Astaras, S. W. Reld, P. S. Yam, A. F. Murray, B. W. Flynn, S. P. Beaumont, D. R. Cumming, and J. M. Cooper, “Implementation of multichannel sensors for remote biomedical measurements in a microsystems format,” *IEEE Transactions on Biomedical Engineering*, Vol. 51, No. 3, 525–535, 2004.
21. Yoon, K. W., S. H. Woo, J. H. Lee, Y. H. Yoon, M. K. Kim, C. H. Won, H.-C. Choi, and J. H. Cho, “Telemetry capsule for pressure monitoring in the gastrointestinal tract,” *IEICE Trans. Fundamentals*, Vol. E89-A, No. 6, 1699–1670, Jun. 2006.
22. Pandolfino, J. E., et al., “Ambulatory esophageal pH monitoring using a wireless system,” *The American Journal of Gastroenterology*, Vol. 98, No. 4, 740–749, 2003.
23. Meron, G. D., “The development of the swallowable video capsule (M2A),” *Gastrointestinal Endoscopy*, Vol. 52, 817–819, Dec. 2000.
24. Kwak, S. I., K. Chang, and Y. J. Yoon, “Small spiral antenna for wideband capsule endoscope system,” *Electronics Letters*, Vol. 42, No. 23, 1328–1329, Nov. 9, 2006.

25. Park, H. J., H. W. Nam, B. S. Song, J. L. Choi, H. C. Choi, J. C. Park, M. N. Kim, J. T. Lee, and J. H. Cho, "Design of bi-directional and multi-channel miniaturized telemetry module for wireless endoscopy," *Proceedings of 2nd Annual International IEEE-EMBS Special Topic Conference on Microtechnologies in Medicine and Biology*, 273–276, Madison, Wisconsin, USA, May 2–4, 2002.
26. Kim, Y., G. Lee, S. Park, B. Kim, J. O. Park, and J. H. Cho, "Pressure monitoring system in gastro-intestinal tract," *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*, 1321–1326, Barcelona, Spain, Apr. 2005.
27. Han, S. G., B. Y. Chi, and Z. H. Wang, "A 2.4 GHz low power ASK transmitter for wireless capsule endoscope applications," *Chinese Journal of Semiconductors*, Vol. 27, No. 6, 988–993, 2006.
28. Koyangi, Y., H. Kawai, K. Ogawa, and K. Ito, "Consideration of the local SAR and radiation characteristics of a helical antenna using a cylindroid whole body phantom at 150 MHz," *Electronics and Communications in Japan*, Part 1, Vol. 87, No. 1, 48–59, 2004.
29. Chirwa, L. C., P. A. Hammond, S. Roy, and D. R. S. Cumming, "Electromagnetic radiation from ingested sources in the human intestine between 150 MHz and 1.2 GHz," *IEEE Transactions on Biomedical Engineering*, Vol. 50, No. 4, 484–492, Apr. 2003.
30. Xu, L., M. Q.-H. Meng, H. Ren, and Y. Chan, "Radiation characteristics of ingestible wireless devices in human intestine following radio frequency exposure at 430, 800, 1200 and 2400 MHz," *IEEE Transactions on Antennas and Propagation*, Accepted for Future Publication.
31. Pennes, H. H., "Analysis of tissue and arterial blood temperature in resting forearm," *J. Appl. Physiol.*, Vol. 1, 93–122, 1948.
32. ICNIRP, "Guidelines for limiting exposure to time-varying electric, magnetic and electromagnetic fields (up to 300 GHz)," *Health Phys. Soc.*, Vol. 74, No. 4, 494–522, Apr. 1998.
33. "IEEE standards for safety levels with respect to exposure to radio frequency electromagnetic fields, 3 kHz to 300 GHz," IEEE Standard C95.1-2005, 2005.
34. Hondou, T., "Rising level of public exposure to mobile phones: Accumulation through additivity and reflectivity," *J. Phys. Soc. Jpn.*, Vol. 71, No. 2, 432–435, Feb. 2002.
35. Hondou, T., "Physical validity of assumptions for public exposure to mobile phones," *J. Phys. Soc. Jpn.*, Vol. 71, No. 12, 3101–3102, Dec. 2002.
36. Hondou, T., T. Ueda, Y. Sakata, N. Tanigawa, T. Suzuki, T. Kobayashi, and K. Ikeda, "Passive exposure to mobile phones: Enhancement of intensity by reflections," *J. Phys. Soc. Jpn.*, Vol. 75, No. 8, 084801-1–08481-5, Aug. 2006.
37. Kramer, A., J. Frohlich, and N. Kuster, "Towards danger of mobile phones in planes, trains, cars, and elevators," *J. Phys. Soc. Jpn.*, Vol. 71, No. 12, 3100–3100, Dec. 2002.
38. Chou, C. H., "Trapped' radiation," *New Sci. Mag.*, Vol. 174, No. 2348, Jun. 2002, [Online], Available: <http://space.newscientist.com/article/mg17423486.000-trapped-radiation.html>.
39. Tropicainen, A., "Human exposure by mobile phones in enclosed areas," *Bioelectromagnetics*, Vol. 24, No. 1, 63–65, Jan. 2003.
40. Hikage, T., T. Nojima, S. Watanabe, and T. Shinozuka, "Electric-field distribution estimation in a train carriage due to cellular radios in order to assess the implantable cardiac pacemaker EMI in semi-echoic environments," *IEICE Trans. Commun.*, Vol. E88-B, No. 8, 3281–3286, Aug. 2005.
41. Ferrer, J., et al., "On the exposure to mobile phone radiation in trains," *Appl. Phys. Lett.*, Vol. 86, 224101-1–224101-3, 2005.
42. Martinez-Burdalo, M., L. Nonidez, A. Martin, and R. Villar, "FDTD analysis of the maximum SAR when operating a mobile phone near a human eye and a wall," *J. Phys. Soc. Jpn.*, Vol. 71, No. 12, 3101–3102, Dec. 2002.
43. Cooper and V. Hombach, "The specific absorption rate in a spherical head model from a dipole with metallic walls nearby," *IEEE Trans. Microw. Theory Tech.*, Vol. 40, No. 4, 377–382, Nov. 1998.
44. Dominguez, H., A. Raizer, and W. P. Carpes, Jr., "Electromagnetic fields Radiated by a cellular phone in close proximity to metallic walls," *Magnetics*, Vol. 38, No. 2, 793–796, Mar. 2002.
45. Siegbahn, M., H. Gradin, B. Thors, and C. Tornevik, "RF exposure from multiple mobile phones in an enclosed environment," *Abstracts Bioelectromagn. Soc. Annual Meeting*, 499–502, Jun. 2006.

46. Simba, A. Y., T. Hikage, S. Watanabe, and T. Nojima, “The dependence of SAR upon position of a mobile phone user in enclosed environments,” *29th Bioelectromagn. Annu. Abstract Collection Meeting*, 256–258, Jun. 2007.
47. Jiang, H. and H. Arai, “Analysis of computational error in antenna’s simulation by using non-uniform mesh FDTD,” *IEICE Trans. Commun.*, Vol. E83-B, No. 7, 1544–1553, Jul. 2000.
48. Simba, A. Y., T. Hikage, T. Nojima, and S. Watanabe, “Specific absorption rates of anatomically realistic human models exposed to RF electromagnetic fields from mobile phones used in elevators,” *IEEE Transactions on Microwave Theory and Techniques*, Vol. 57, No. 5, 1250–1259, May 2009.

Behavioral Models for Power Amplifier Using a Difference-frequency Dual-signal Injection Method

Hui Wang and Peiguo Liu

School of Electronic Science and Engineering, National University of Defense Technology, Changsha, China

Abstract— In this paper a difference-frequency dual-signal injection method is developed to modeling of RF power amplifiers (PAs) in nonlinear microwave systems. In this method, the dual-signal is with different frequency spacing. By varying the frequency spacing of dual-signal and the power level, behavioral models for PAs can be extracted without memory effects and with memory effects. For verification, the difference-frequency dual-signal is implemented with a 10 W class-AB PA and tested using a dual sinusoidal signal with 10 MHz frequency spacing from 0.9 GHz \sim 1 GHz. From the experiment results, it is found that the model improves adjacent channel power ratio (ACPR) prediction accuracy by over 8 dB and 2 dB, compared to the conventional memory polynomial model and with conventional Volterra-based behavioral model respectively.

1. INTRODUCTION

Radio frequency Power amplifiers (PAs) are indispensable components in a microwave system and are inherently nonlinear [1]. It is well know that there is an approximate inverse relationship between the PAs efficiency and their linearity. To increase their efficiency, PAs are sometimes driven into their nonlinear operating range. For such reasons, behavioral modeling for RF PAs has grown to become a topic of great interest for all those involved in microwave engineering [3–9].

In this paper, we propose a difference-frequency dual-signal injection method to modeling of RF power amplifiers (PAs) in nonlinear microwave systems. This paper is organized as follows. In Section 2, dual-signal injection method is described. The IM distortion (IMD) and AM/AM, AM/PM are extracted, and the models without memory effects and with memory effects are described in Section 4. The experiments are given in Section 5, and with a conclusion in Section 6.

2. DUAL-SIGNAL INJECTION METHOD

Figure 1 shows the block diagram for IMD and AM/AM, AM/PM measurement and model extraction. A power synthesizer is used to synthesize a dual sinusoidal signal into the power amplifier under test.

The dual-signal can be described as

$$x(t) = x_1(t) + x_2(t) = A_1 \cos(\omega_1 t + \theta_1) + A_2 \cos(\omega_2 t + \theta_2) \quad (1)$$

where $A_1 = A_2 = A$, and θ_1 and θ_2 are two independent stochastic variables.

Using the trigonometric identity

$$\begin{aligned} x(t) &= 2A \cos \left[\frac{(\omega_1 + \omega_2)t}{2} + \frac{(\theta_1 + \theta_2)}{2} \right] \cdot \cos \left[\frac{(\omega_1 - \omega_2)t}{2} + \frac{(\theta_1 - \theta_2)}{2} \right] \\ &= 2A \cos[\omega_c t + \theta_c] \cdot \cos[\omega_m t + \theta_m] \end{aligned} \quad (2)$$

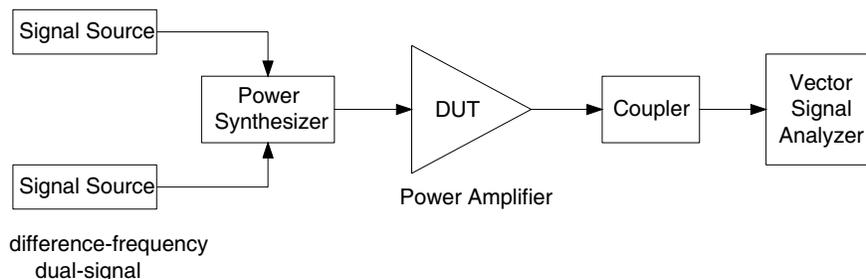


Figure 1: Block diagram for model extraction.

where

$$\omega_c = \frac{(\omega_1 + \omega_2)}{2} \tag{3}$$

$$\omega_m = \frac{(\omega_1 - \omega_2)}{2} \tag{4}$$

The first cosine term in (2) corresponds to the suppressed carrier at the frequency between the two input signals, while the second cosine term modulates the carrier at the baseband frequency ω_m .

3. IM DISTORTION AND AM/AM, AM/PM

This paper is devoted to providing a comprehensive study of the PAs' in-band and out-of-band distortion: intermodulation distortion and AM/AM and AM/PM characteristics.

3.1. IM Distortion

In [11], it is said “When an amplifier is excited by multiple frequencies and driven into its nonlinear operating range, it generates numerous mixing products.”

The input signal $x(t)$ can be described as (1), and the output would be given by

$$y(t) = \sum_{i=1}^N a_i [x(t)]^i = \sum_{i=1}^N a_i [A_1 \cos(\omega_1 t + \theta_1) + A_2 \cos(\omega_2 t + \theta_2)]^i \tag{5}$$

which shows that the output would be composed of a very large number of mixing terms involving all possible combinations. The new frequency components could be described as

$$\omega_{m,n} = m\omega_1 + n\omega_2 \tag{6}$$

In (6), m and n are integers; the sum of the magnitudes of these integers is defined as the order of IMD. Typically only third order IMD (the frequency components at $2\omega_1 - \omega_2$ and $2\omega_2 - \omega_1$) is studied because it is the greatest magnitude and most likely to fall into an adjacent receive band.

3.2. AM/AM, AM/PM

AM/AM and AM/PM measurements are consisted in acquiring the input-output relations of the PAs in both amplitude and phase. We describe (2) as

$$x(t) = r(t) \cos(\omega_c t + \varphi(t)) \tag{7}$$

where $r(t)$ and $\varphi(t)$ are the amplitude and phase of $x(t)$.

When

$$y(t) = g(r(t)) \cos(\omega_c t + \varphi(t) + f(r(t))) \tag{8}$$

where $g(\cdot)$ and $f(\cdot)$ represent the AM/AM and AM/PM distortions, respectively, and $r(t)$ and $\varphi(t)$ are the envelope and phase of the input signal, respectively.

4. BEHAVIORAL MODELS FOR POWER AMPLIFIER

A memoryless PA can be modeled with AM/AM and AM/PM [9]. The input signal is rewritten as

$$x_{RF}(t) = r(t) \cos(\omega_c t + \varphi(t)) \tag{9}$$

4.1. Memoryless Models

The memoryless model has been used for many years because of its easier computational implementation, its relative efficiency and acceptable level of accuracy can be achieved in narrow band systems.

Then the output signal can be written as

$$y_{RF}(t) = g(r(t)) \cos(\omega_c t + \varphi(t) + f(r(t))) \tag{10}$$

$$y_{RF}(t) = g(r(t)) \cos(f(r(t))) \cos(\omega_c t + \varphi(t)) - g(r(t)) \sin(f(r(t))) \sin(\omega_c t + \varphi(t)) \tag{11}$$

4.2. Memory Polynomial Model

The memory polynomial model is widely used to describe nonlinear effects in PAs with memory effects. The general form of this model can be written as

$$y(n) = \sum_{q=0}^Q h_1(q) x(n-q) + \sum_{q_1, q_2, q_3=0}^Q h_3(q_1, q_2, q_3) x(n-q_1) x(n-q_2) x^*(n-q_3) \quad (12)$$

$$y(n) = \sum_{k=1}^K \sum_{q=0}^Q a_{kq} x(n-q) |x(n-q)|^{k-1} \quad (13)$$

where $x(n)$ and $y(n)$ are the input and output, respectively. a_{kq} are the model coefficients. K is the polynomial function order, and Q is the memory depth.

4.3. Volterra-based Behavioral Model

The Volterra-based behavioral model assumes that the response of a nonlinear system with memory, having input $x(n)$ and output $y(n)$, can be expressed as

$$y(n) = \sum_{k=1}^K \sum_{q_1=0}^Q \dots \sum_{q_m=0}^Q h_m(n-q) \prod_{j=1}^m x(n-q_j) \quad (14)$$

A Volterra-based behavioral model is combination of linear convolution and a nonlinear power series; it provides a general way to model a nonlinear system with memory. However, high computational complexity and great number of coefficients makes this method impractical in some real applications, because of the degree of nonlinearity and the length of the system.

Recently, to overcome these problems, pured Volterra series [4], modified Volterra series [5], dynamic deviation reduced Volterra series [6–8] have been used for reducing the complexity and simplifying classical Volterra series based behavioral models for RF power amplifiers.

4.4. The Proposed Model

In the conventional memory polynomial model, the baseband output (13) is given as a function of the baseband complex input samples $[x(n), x(n-1), \dots, x(n-Q)]$. Here, we use the envelope of the input to instead the q samples, then the output $y(n)$ is written as

$$y(n) = \sum_{k=1}^K \sum_{q=0}^Q a_{kq} x(n) |x(n-q)|^{k-1} = x(n) \sum_{k=1}^K \sum_{q=0}^Q a_{kq} |x(n-q)|^{k-1} \quad (15)$$

5. MODEL VALIDATION

To analyze the performances of different behavioral models, a 450 ~ 1500 MHz, 10 W LDMOS PA is simulated, which is the commercial amplifier MW6S010N manufactured by Freescale Semiconductor Inc. The input is a dual sinusoidal signal with 10 MHz frequency spacing from 0.9 GHz ~ 1 GHz.

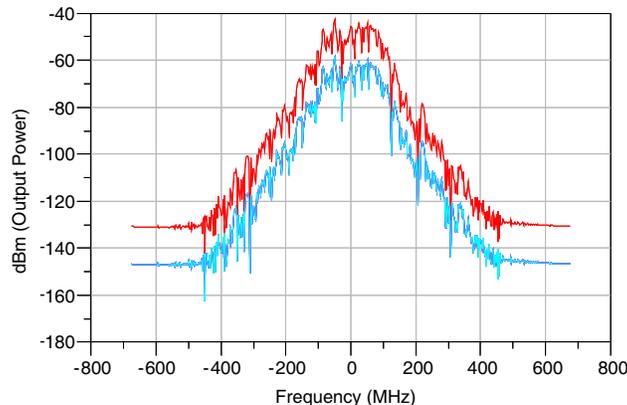


Figure 2: Comparison of the power spectrums of the behavioral models of the PA.

This was calculated using harmonic-balance simulator in Agilent's Advanced Design System (ADS). The model is identified by minimizing the mean square error (MSE), and method given by M. Isaksson [2] is used for the synchronization of input and output signals.

Figure 2 shows the power spectrum density (PSD) of both models of the PA. It is clear that the spectrums of the models track the output spectrum of the PA well, both in-band and in the adjacent channels.

6. CONCLUSION

We have proposed a difference-frequency dual-signal injection method to modeling of RF power amplifiers (PAs) in nonlinear microwave systems. The proposed model can be used to simulate the IMD, and AM/AM and AM/PM that depend on the history of the input signal.

To validate the proposed modeling technique, a 450 ~ 1500 MHz, 10 W LDMOS PA is simulated. It is found that the model improves ACPR prediction accuracy by over 8 dB and 2 dB, compared to the conventional memory polynomial model and with conventional Volterra-based behavioral model respectively.

REFERENCES

- Schreurs, D., et al., *RF Power Amplifier Behavioral Modeling*, The Cambridge RF and Microwave Engineering Series, 2008.
- Isaksson, M., D. Wisell, and D. Ronnow, "A comparative analysis of behavioral models for RF power amplifiers," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 54, 348–359, 2006.
- Crespo-Cadenas, C., J. Reina-Tosina, and M. J. Madero-Ayora, "Volterra behavioral model for wideband RF amplifiers," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 55, 449–457, 2007.
- Zhu, A. and T. J. Brazil, "Behavioral modeling of RF power amplifiers based on pruned volterra series," *IEEE Microwave and Wireless Components Letters*, Vol. 14, 563–565, 2004.
- Mirri, D., G. Luculano, F. Filicori, G. Pasini, G. Vannini, and G. P. Gabriella, "A modified Volterra series approach for nonlinear dynamic systems modeling," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, Vol. 49, 1118–1128, 2002.
- Zhu, A., J. C. Pedro, and T. J. Brazil, "Dynamic deviation reduction-based volterra behavioral modeling of RF power amplifiers," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 54, 4323–4332, 2006.
- Isaksson, M. and D. Ronnow, "A kautz-volterra behavioral model for RF power amplifiers," *IEEE MTT-S International Microwave Symposium Digest*, 485–488, 2006.
- Zhu, A. and T. J. Brazil, "RF power amplifier behavioral modeling using Volterra expansion with Laguerre functions," *2005 IEEE MTT-S International Microwave Symposium Digest*, 963–966, 2005.
- Hyunchul, K., M. D. McKinley, and J. S. Kenney, "Extraction of accurate behavioral models for power amplifiers with memory effects using two-tone measurements," *2002 IEEE MTT-S International Microwave Symposium Digest*, 139–142, 2002.
- Cabral, P. M., J. Pedro, and N. B. Carvalho, "Modeling nonlinear memory effects on the AM/AM, AM/PM and two-tone IMD in microwave PA circuits," *International Journal of RF and Microwave Computer-Aided Engineering*, Vol. 16, 13–23, 2006.
- Remley, K. A., D. F. Williams, D. M. M. P. Schreurs, and J. Wood, "Simplifying and interpreting two-tone measurements," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 52, 2576–2584, 2004.
- Clark, C. J., C. P. Silva, A. A. Moulthrop, and M. S. Muha, "Power-amplifier characterization using a two-tone measurement technique," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 50, 1590–1602, 2002.

Analysis and Design for High-gain Antenna with Periodic Structures

Han-Nien Lin and Chun-Chi Tang

Department of Communications Engineering, Feng-Chia University
100 Wen-Hua Rd., Taichung 40724, Taiwan

Abstract— Unlike metal as conventional Perfect Electric Conductor (PEC) for unidirectional beam reflector, the periodic structures are now attracting more attention from academia and industrial as High Impedance Surface (HIS) or Artificial Magnetic Conductor (AMC) reflector. Because quarter wavelength ($\lambda/4$) separation between antenna and PEC reflector is needed for unidirectional beam applications, it inevitably increases the profile and stringently limits the operation bandwidth of the antenna system. However, the utilization of AMC of periodic structures as reflector not only significantly reduces the antenna profile with phase enhancement, but it will also enhance its gain and directivity performance. In this paper, we use monopole antenna with center frequency at 2.45 GHz and design periodic structure operated in frequency range between 2.4 GHz and 2.483 GHz to investigate the applications of EBG structure on antenna. We will finally compare the simulated result and measured result to further predict the optimization of the antenna-EBG system design parameters.

1. INTRODUCTION

With the increasingly popular and wide-spreading of portable wireless communications devices in recent years. The mono-pole antenna discussed here is designed for applications of WLAN (802.11 b/g) and Bluetooth systems operating in 2.4 ~ 2.484 GHz frequency range. On the other hand, the electromagnetic band-gap (EBG) structure [1] has gained the most attention in the microwave community. The most attractive characteristic of EBG material is the band-stop and slow-wave response from the periodic arrangement of structure or component. It thus results in high surface impedance and energy band-gap (stop-band or pass-band) due to impedance discontinuity between air and high-impedance surface.

Due to periodic dimension of EBG structure commonly being about half-wavelength of the stop-band, the large size has limited the applications of EBG structure in low frequency range. In this paper, we introduce folded bends to lengthen the electric current path and thus increase equivalent inductance. We also utilize the longer edge-coupling effect to lengthen the magnetic current path and thus increase equivalent capacitance. Therefore, we are able to design compact EBG structure for low frequency applications using the above techniques.

Electric conductor is usually used as shielding or reflecting substrate to minimize radiation hazard or generate unidirectional antenna beam in portable wireless communication device [2, 3]. When low-profile antenna is placed close to electric conductor, the image current induced from conductor will degrade the performance of antenna system. It is therefore required the separation between low-profile antenna and electric conductor should be at least quarter-wavelength ($\lambda/4$) [4]. The induced surface current on electric conductor will also radiate while traveling toward edges and affect the field distribution captured by antenna. Opposite to the reflection coefficient -1 with 180° phase shift from perfect electric conductor (PEC), the electromagnetic wave impinging upon AMC material will result in in-phase reflection coefficient $+1$. We thus utilized the EBG structure as antenna reflector and design the band-gap to match antenna's resonant frequency in this paper. The low-profile requirement and performance improvement of antenna is achieved by suppressing the surface wave from the application of EBG structure.

2. DESIGN OF EBG STRUCTURE

The completely constructed reflection phenomena from radiation would result from periodic structure with half-wavelength or its integral multiple as period dimension and further result in energy band-gap characteristics. The stop-band is usually determined by periodic separation λ EBG between metallic components. The stop-band and center frequency of EBG structure presented in this paper are 2.4 ~ 2.483 GHz and 2.45 GHz respectively. When FR4 substrate ($\epsilon_{eff} = 4.26$ and thickness 0.8 mm) is used in this design, the resonant wavelength in dielectric and unit cell dimension λ EBG can be calculated from Equations (1) and (2) to be 76 mm and 38 mm respectively. While conventional EBG structure is capacitive itself between periodic metallic components, we introduced folding bend metallic components to reduce dimension of metallic patch L according

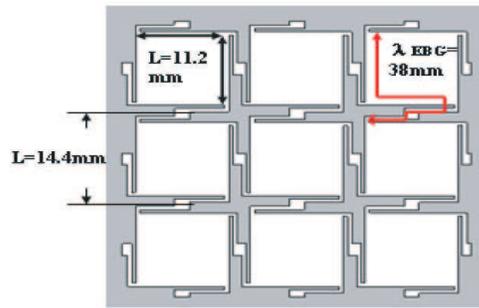


Figure 1: Dimensions of periodic structure.

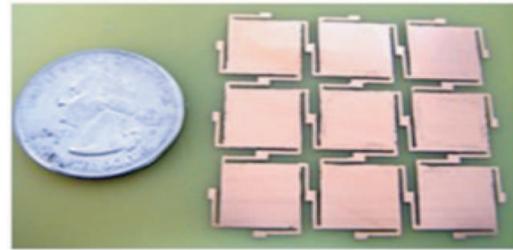


Figure 2: Prototype of periodic structure.

to Equation (3) due to effective inductance of metallic component itself. The effect of increased equivalent inductance and capacitance is able to shift stop-band or energy band-gap to lower frequency. The original dimension L is reduced from half-wavelength (38 mm) to 11.2 mm as shown in Figure 1. The EBG structure investigated in this paper is 3×3 Unit cells as shown in Figure 2.

$$fr = \frac{c}{\lambda_0 \sqrt{\frac{\epsilon_{eff} + 1}{2}}} \tag{1}$$

$$\lambda_{EBG} = \frac{\lambda_0}{2} = \frac{1}{2} \left(\frac{2\pi}{\beta} \right) \tag{2}$$

$$fr = \frac{1}{2\pi\sqrt{LC}} \tag{3}$$

where c is the speed of light in free space, ϵ_{eff} the equivalent dielectric constant, λ_{EBG} the wavelength of metallic component periodic dimension, λ_0 the wavelength of light in free space, fr the resonant frequency, β the wave-number in free space, and η is the free space wave impedance.

We first calculated the basic design parameters from the above equations, and then utilized the full-wave FDTD (Finite-Difference Time-Domain) electromagnetic simulation software for the optimal geometric structure design of periodic pattern. The resulted transmission characteristic is shown in Figure 3, where the bandwidth of stop-band is 220 MHz (2.31 ~ 2.53 GHz) to meet -20 dB attenuation specification.

3. SIMULATION AND MEASUREMENT COMPARISON OF MONOPOLE ANTENNA

The monopole antenna investigated is designed on FR4 substrate with equivalent dielectric constant $\epsilon_{eff} = 4.26$ and 0.8 mm thickness. The dimension of antenna is shown in Figure 4 when monopole antenna equals quarter-wavelength with center frequency $fr = 2.45$ GHz. Figure 5 shows the good characteristic agreement between Simulation and measurement results. Figure 6 shows maximum antenna gain 2.34 dBi occurring at center frequency 2.45 GHz and Figure 7 shows its efficiency equals 58%. The measured 3-dimensional radiation pattern of the monopole antenna is shown in Figure 8.

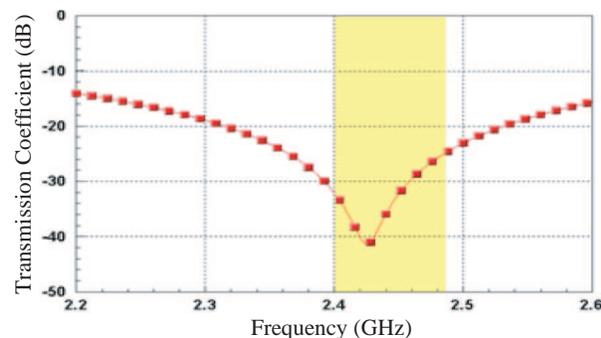


Figure 3: Simulated transmission coefficient of periodic structure.

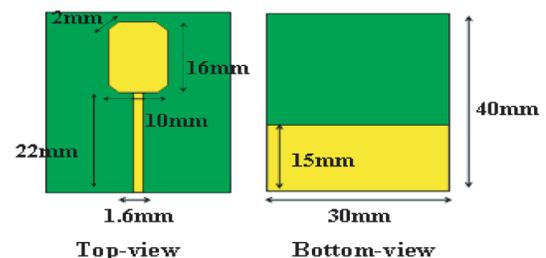


Figure 4: Dimension of monopole antenna.

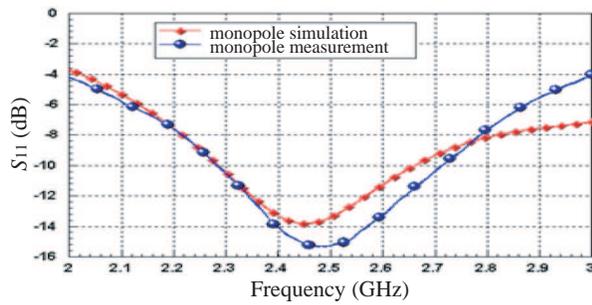


Figure 5: Simulation and measurement comparison of monopole antenna.

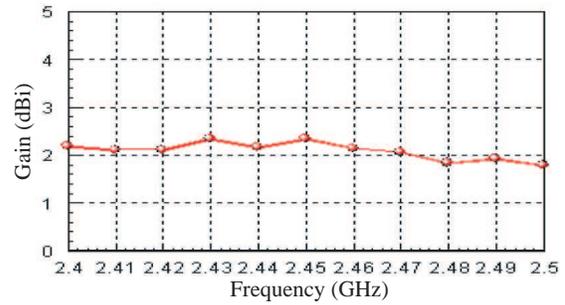


Figure 6: Measured gain of monopole antenna.

4. ANALYSIS OF PEC AND EBG STRUCTURE EFFECT ON MONOPOLE ANTENNA PERFORMANCE

We first placed the EBG structure 0.12λ ($h\text{-air} = 15\text{ mm}$) beneath the monopole antenna to work as reflector. The physical arrangement of the monopole antenna with EBG structure as reflector is shown in Figure 9. Figure 10 shows the arrangement of the monopole antenna with different structures as reflector, and Figure 11 shows the simulated results for PEC and EBG cases. It is usually required quarter-wavelength separation between antenna and PEC reflector to obtain best performance, however the image current induced by PEC would usually result in energy dissipation, impedance mismatching, and performance degradation of antenna. It would also produce a lower resonant frequency due to capacitive coupling effect, and thus result in impedance matching degradation in operation frequency as shown in Figure 11. Figure 12 shows the maximum antenna gain enhanced from 2.34 dBi to 3.28 dBi (almost 1 dB improvement) at center frequency 2.45 GHz with EBG structure implementation and Figure 13 shows the antenna efficiency be increased from

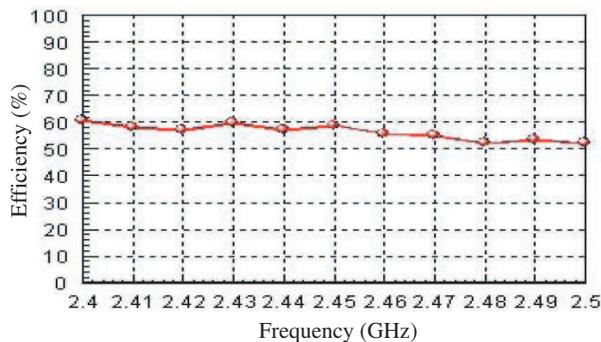


Figure 7: Measured efficiency of monopole antenna.

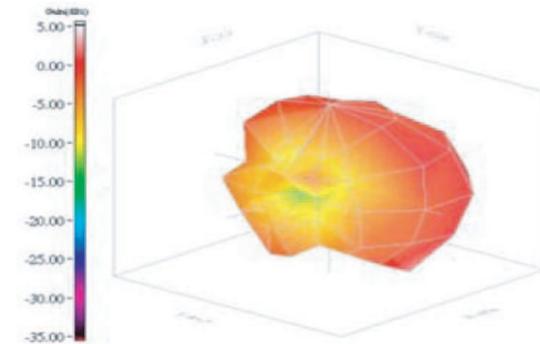


Figure 8: Measured 3D radiation pattern of monopole antenna (@ 2.45 GHz).

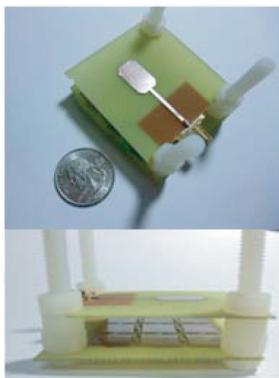


Figure 9: Physical arrangement of monopole antenna and EBG combination.

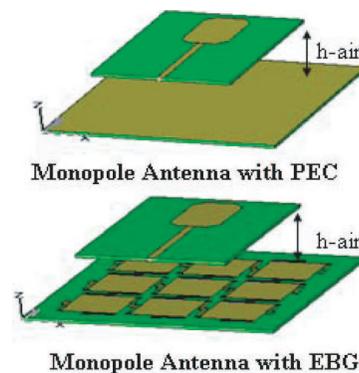


Figure 10: Arrangement of the monopole antenna with different structures as reflector.

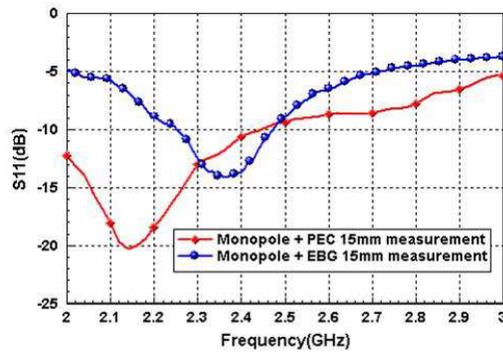


Figure 11: S_{11} comparison of the monopole antenna with PEC and EBG as reflector.

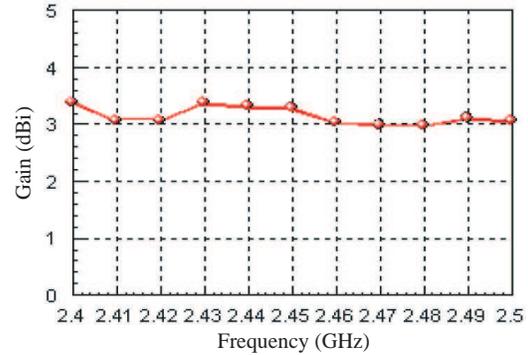


Figure 12: Measured gain of the monopole antenna and EBG combination.

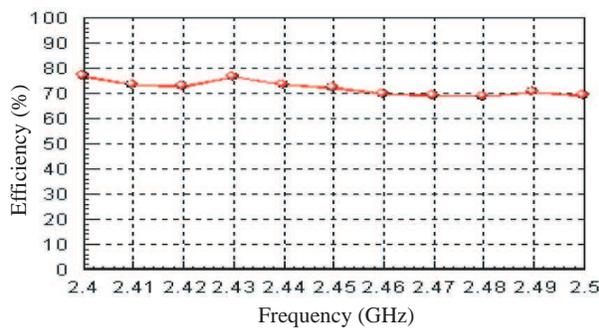


Figure 13: Measured efficiency of the monopole antenna and EBG combination.

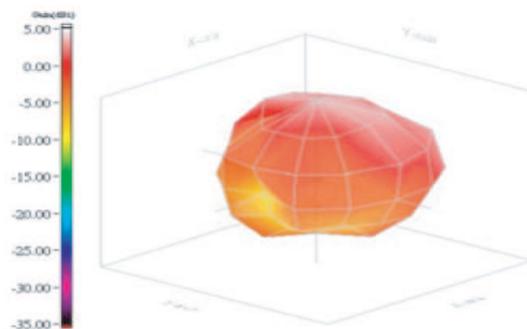


Figure 14: Measured 3D radiation pattern of the monopole antenna and EBG combination (@ 2.45 GHz).

58% to 72% (14% improvement). The measured 3-dimensional radiation pattern of the monopole antenna and EBG combination is shown in Figure 14.

5. CONCLUSION

In this paper we have successfully reduced the size of EBG structure about 30% (periodicity reduced from 38 mm to 11.2 mm) to extend its application to lower frequency. We also implemented EBG structure as antenna reflector to obtain 50% separation reduction (from 30 mm to 15 mm) between antenna and reflector compared to PEC. In the meantime, the antenna gain has improved 1 dB (from 2.34 dBi to 3.28 dBi) and the efficiency increased 14% (from 58% to 72%) both measured at center frequency 2.45 GHz. Therefore not only low-profile construction is achieved, but antenna performance is also apparently improved from the simulation and measurement verification.

REFERENCES

1. Yang, F., “Electromagnetic band gap structures in antenna engineering,” *The Cambridge RF and Microwave Engineering Series*, New York, 2009.
2. Yang, F. and Y. Rahmat-Samii, “A low profile circular polarized curl antenna over electromagnetic-band-gap (EBG) surface,” *Microwave Optical Tech. Lett.*, Vol. 31, No. 4, 264–267, November 2001.
3. Li, Z. and Y. Rahmat-Samii, “PBG, PMC and PEC ground planes: A case study of dipole-antenna,” *IEEE Trans.*, Los Angeles, CA, 674–677, 2000.
4. Bell, J. M. and M. F. Iskander, “A low-profile archimedean spiral antenna using an EBG ground plane,” *IEEE Antennas Wireless Propagat. Lett.*, Vol. 3, 223–226, 2004.

High Frequency Parameters of a Hermetic Motor and Their Effects on Conducted Emission

Ming Chen, Xudong Sun, and Lipei Huang

Country State Key Lab of Power Systems, Department of Electrical Engineering
Tsinghua University, Beijing 100084, China

Abstract— The high frequency (HF) equivalent circuit model of the typical three phase hermetic motor used in a variable frequency air conditioner is established for analyzing the conducted emission to the distribution network, and its parameters are extracted from the impedance characteristics of the motor by applying a least-squared data fitting procedure. The impedance characteristics are measured in two states: the fluorocarbon in the motor is liquid and exhausted, and the difference between the parameters in the two states is analyzed. The HF circuit model of the conducted emission propagation path is established by connecting the models of the motor and the related passive elements together. The effect of the motor parameters is investigated by simulation, and the results indicate that the equivalent winding-to-ground capacitance is the most sensitive parameter. The emission spectrum when the motor operates normally is measured to verify the above model and analysis. The spectrum obtained by simulation with the parameters in the latter state matches with the measured one in general, which shows that the motor model and the extraction method of its HF parameters are effective, and the analysis results are credible.

1. INTRODUCTION

In a modern PWM variable frequency AC drive consisting in a motor and an inverter, power semiconductor devices such as insulated gate bipolar transistors (IGBTs) are widely used as power switches generally operating at high frequency. The rapid transition of voltage and current when the switches turn on or off would cause electromagnetic emission through the motor, cable, the passive devices and other related parts in the circuit, and result in electromagnetic interference (EMI) problems, such as terminal over voltage in motor with long cables and high frequency leakage currents flowing in the stray capacitances of the cables and the motor. In order to study the characteristics of the propagation path and its influence on the level of conducted emissions, it is necessary to use precise models of the motor, its PWM inverter and cables [1]. Besides, designers and users of the drives are very interesting in predicting the conducted EMI level with an acceptable accuracy. Therefore, the high frequency (HF) models of the whole inverter system have to be investigated [2]. In any case, the HF behavior of the motor plays an important role in determining the conducted emission [3].

Variable frequency air conditioners are used more and more widely nowadays. In such an air conditioner, a drive consisting in a three phase hermetic motor and an inverter is adopted. The conducted emission caused by the inverter is strictly limited by the relative standards, so the manufacturers have to try to make them satisfy the requirements. The motor is an important part of the propagation path of the emission. In order to evaluate the emission level, a modeling method of the typical three phase hermetic motor is developed based on experiments in this paper. The HF parameters of the hermetic motor are extracted when the fluorocarbon in it is liquid and exhausted respectively, and the difference between them is examined to find out the one suitable for the evaluation. Then the effects of the parameters are analyzed and a comparison is made between the spectrum resulting from simulation and the measured one to verify the motor model and its parameters.

2. THE HF CIRCUIT MODEL AND PARAMETERS OF THE HERMETIC MOTOR

In order to establish the HF equivalent circuit model of the propagation path of the conducted emission to the distribution network, the models of the hermetic motor with certain accuracy should be set up at first. Such a motor has inductive couplings and stray capacitances which are more complex than those of a simple inductive load [4, 5]. Obviously, the whole high frequency parameters of the motor are very complicated, which are related to its windings, core and structure. Therefore, simplification should be made for establishing the model and extracting its parameters.

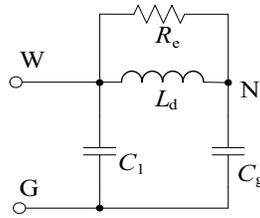


Figure 1: The HF circuit model for one phase.

2.1. The HF Equivalent Circuit Model of the Motor

The HF parameters of the motor are inherently distributed ones. But, for the convenience of analysis, some lumped parameters have to be used to set up the model.

In most cases, the high frequency lumped equivalent circuit with parasitic capacitances for an AC motor will provide an adequate approximation of the impedance characteristics of the motor [6]. Such a model, after a slight modification, is adopted as the high frequency equivalent circuit model of the hermetic motor. The equivalent circuit for a single phase of the hermetic motor is shown in Fig. 1, which has the following lumped parameters: the phase leakage inductance L_d , the capacitance C_1 representing the effect of the distributed capacitive coupling between the input leads to the ground, the capacitance C_g representing the winding-to-ground distributed capacitances, and the equivalent resistance R_e representing the iron loss in the core and frame of the motor. These parameters do not include the skin effect for the stator and rotor leakage inductances. It should be underline that all the values of these parameters have to be referred to a single star-connected phase.

This high frequency model of the hermetic motor based on lumped parameters should be accurate over the frequency range from a few kilohertz up to several million hertz. By adopting such a simplified model, using more parasitic parameter is avoided, and consequently any network simulation program can be used to analyze it without any problem. This model will be used in simulation so that the conducted emission can be optimally analyzed.

2.2. The Extraction Method

The parameters in the above model are very difficult to calculate because the motor is hermetic and no concrete information on its structure is available. So an extraction method based on experiment has to be used.

Firstly, an impedance frequency characteristic of the motor can be obtained by connecting the three phase terminals together and measuring between the terminals and the grounding, since the motor under research has a star connection, and its frame is grounded whereas the neutral point of the three phase windings is not connected to the ground. The measured characteristic of the motor are considered as the impedance-frequency characteristic of the equivalent circuit in Fig. 1. Then, the parameters in the circuit can be obtained from the measured curve by applying a least-squared data fitting procedure.

Since the frequency components higher than 1MHz are not able to deeply penetrate in the motor windings, and the parasitic capacitances play a dominant role at lower frequency [3], so the impedance-frequency characteristic can be measured over the frequency range from 1 kHz to 1MHz by using an impedance analyzer.

2.3. Extracting the HF Parameters of the Motor in Two States

2.3.1. Measurement of the Impedance $Z_{WG}(f)$

The impedance $Z_{WG}(f)$ (where f is the frequency) between the nodes W and G in Fig. 1 can be obtained by measuring between the three phase terminals connected together and the grounding terminal with an impedance analyzer, with a floating motor neutral terminal.

2.3.2. A Least-squared Data Fitting

A least-squared data fitting procedure is applied to the measured curve to obtain the parameters of the motor. The calculation needs to be carried out in complex domain, and an adequate choice of the initialization values of these parameters is very important to obtain a good fitting of the measured data. The parameters obtained are shown in Table 1. The measured impedance-frequency

Table 1: The HF parameters of the hermetic motor.

Parameter	C_1 (nF)	C_g (nF)	R_e (Ω)	L_d (μ H)
Value	0.11	22.2	1242.5	13.0

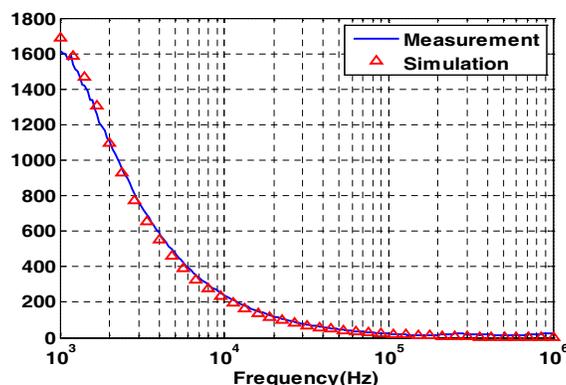


Figure 2: The measured and simulation impedance frequency characteristics curves.

characteristic of the motor and the one obtained by simulation with the parameter values in Table 1 are shown in Fig. 2. An agreement between them is well evident.

2.3.3. Comparison of the Parameters in Two States

The above impedance-frequency curve of the motor is measured when the air conditioner stops working, and the fluorocarbon in the motor is liquid. But the fluorocarbon is gaseous when the air conditioner is working normally. Since the dielectric constant of the gas is usually smaller than that of the liquid, the value of the parameter C_g when the air conditioner is working should be smaller than that extracted by measurement when the fluorocarbon is liquid. So, in order to obtain accurate parameter values, it is expected to measure the impedance-frequency curve of the motor when the air conditioner is working together with the motor.

However, it is impossible to measure the characteristic in such a condition, because high voltages exist on the motor and the instrument cannot endure them. An alternate extraction method is adopted. The impedance-frequency curve is measured when the fluorocarbon is entirely exhausted and the air fills in the motor, which can be considered as quite close to the situation that the fluorocarbon is gaseous. The same method as mentioned above is used to extract the parameters of motor in this state, and the results are shown in Table 2.

It can be seen from Tables 1 and 2 that there are some differences between the parameters measured in the above two states. The most distinguished one occurs in the parameter C_g : its value in the state that the fluorocarbon is exhausted is much smaller than the one in the state that the fluorocarbon is liquid, whereas the difference of the other parameters is no more than 10%.

Table 2: The parameters of the hermetic motor (without fluorocarbon).

Parameter	C_1 (nF)	C_g (nF)	R_e (Ω)	L_d (μ H)
Value	0.10	5.89	1180.4	13.8

3. ANALYSIS AND VERIFICATION OF THE HF PARAMETERS OF THE MOTOR

It is difficult to verify the effectiveness of the motor model as an analysis tool of the conducted emission by examining itself alone, so an alternative method should be used. The former paper [7] has constructed the model of the propagation path in the same three phase inverter (without taking into account the propagation path of the motor), and the model has been proved to be effective and accurate by simulations and experiments. Hence the model of the motor developed above can be added to the existing circuit model of the propagation path to form an equivalent circuit, and then the motor model can be verified by comparing the simulation results of the emission with the experiment one for the whole equivalent circuit. The circuit model of the whole propagation path is shown in Fig. 3, where the HF equivalent circuit model of the motor is shown in the dashed-line.

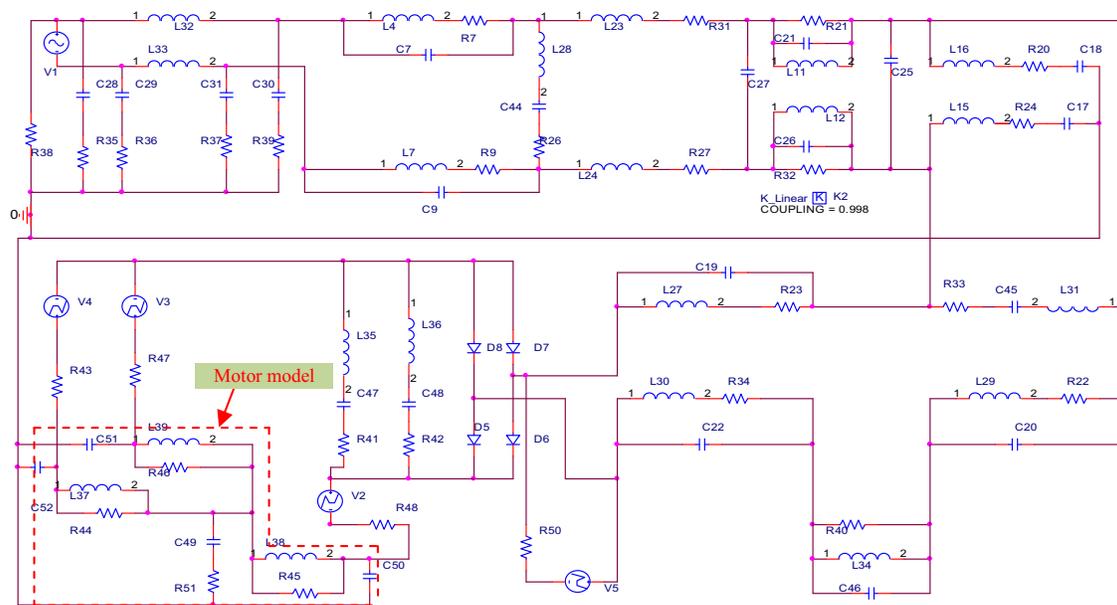


Figure 3: The HF equivalent circuit model of the propagation path.

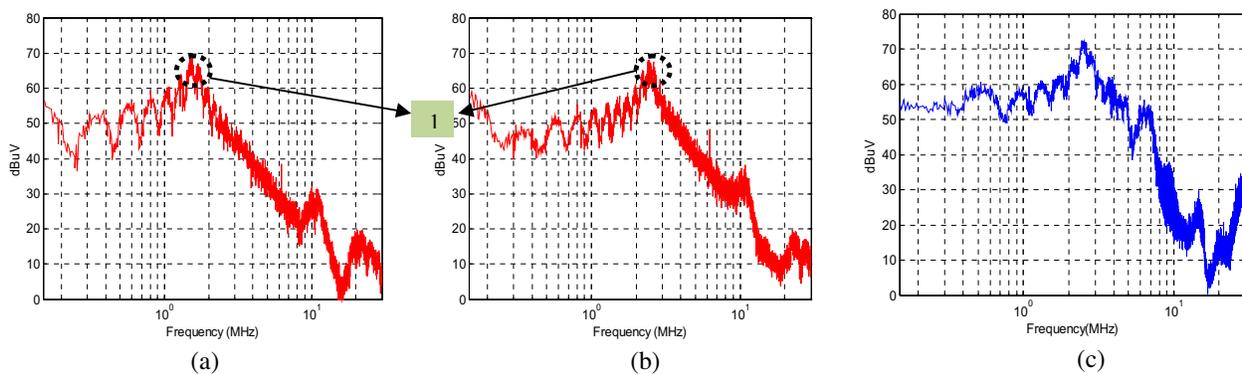


Figure 4: The spectrums obtained by simulation and measurement, (a) fluorocarbon in liquid state, (b) without fluorocarbon, (c) measurement result.

Using the model shown in Fig. 3, the effects of the motor parameters on the emission are analyzed by simulation. The results indicate that the parameter C_g is the essential parameter of the motor, and also the important parameter to affect the characteristics of the propagation path. The overall magnitude of the emission spectrum changes a little with the value of C_g , whereas the frequency of the peak point of the spectrum varies a lot, and becomes higher when the value gets smaller. The parameter L_d has a little effect on the propagation path, and the parameters R_e and C_1 nearly have no effect.

The simulation results of the conducted emission spectrum to the distribution network are shown in Figs. 4(a) and (b) with the parameters of the motor shown in Tables 1 and 2, respectively. It can be seen that the shape and tendency of the two spectrums approximately match with each other. There is a peak point in each spectrum (indicated by the dashed circle in Fig. 4), and the magnitudes of the two peak points are almost the same, but the frequencies at which the peak occurs are quite different: one is about 1.55 MHz, and the other is about 2.4 MHz.

The corresponding spectrum obtained by measurement when the motor is working normally is shown in Fig. 4(c). It can be seen that the simulation spectrum with the parameters of the motor without fluorocarbon (i.e., entirely exhausted) approximately matches with the measurement one in general, and the frequencies at which the peak occurs are almost the same, about 2.4 MHz and 2.45 MHz in the simulated and measured spectrums, respectively. This indicates that the HF model of the motor and its parameters extracted in the state that the fluorocarbon is entirely exhausted are effective, and the circuit model of the whole propagation path of the three phase converter unit are valid.

4. CONCLUSIONS

The HF equivalent circuit model of the hermetic motor is established, and its parameters are extracted from the impedance–frequency characteristics of the motor obtained by measurement in two states: the fluorocarbon in the motor is liquid and it is completely exhausted. The model is added to the equivalent circuit of the propagation path of the inverter so that a model of the whole propagation path is set up. The effect of the motor parameters on the conducted emission is investigated by simulation, which indicates that the capacitance of the winding to ground is the essential parameter. A comparison is made between the emission spectrums obtained by simulation and the measured spectrum. The simulated spectrum with the parameters in the state that the fluorocarbon is entirely exhausted approximately match with the measured one in general, and the peaks in the two spectrums have almost the same magnitude and frequency, which shows that the model and the parameters of the hermetic motor developed in this paper are effective and credible, and the measurement undertaken when the fluorocarbon is exhausted is an adequate substitute for that when the fluorocarbon is gaseous.

REFERENCES

1. Ogasawara, S. and H. Agaki, “Modeling and damping of high frequency leakage currents in PWM inverter-fed AC motor drive system,” *IEEE Transactions on Industry Applications*, Vol. 32, 1105–1114, May 1996.
2. Moreira, A. F. and T. A. Lipo, “High frequency modeling for cable and induction motor overvoltage studies in long cable drives,” *IEEE Transactions on Industry Applications*, Vol. 38, No. 5, 1297–1306, 2002.
3. Ran, L., S. Gokani, J. Clare, K. J. Bradley, and C. Christopoulos, “Conducted electromagnetic emission in induction motor drive system — Part I: Time domain analysis and identification of dominant modes,” *IEEE Transactions on Power Electronics*, Vol. 13, No. 4, 757–767, 1998.
4. Vermilion, C., F. Costa, A. Lacarroy, and M. Arpillre, “Modeling PWM inverter-fed AC motor drive system,” *Proceedings of the European on Power Electronics Conference (EPE)*, 2003.
5. Zare, F., “A high frequency model of an electric motor used for high speed switching applications,” *Proceedings of the European on Power Electronics Conference (EPE)*, 2003.
6. Boglietti, A. and E. Carpaneto, “Induction motor high frequency model,” *Industry Applications Conference (IAS)*, Vol. 3, 1551–1558, 1999.
7. Chen, M., X. Sun, and L. Huang, “A method to model the conducted EMI propagation path for PFC,” *IPEMC'09*, 674–677, 2009.

Using Grey Decision Making Approach to Improve FPGA Performance

Jan-Ou Wu¹, Yang-Hsin Fan², and San-Fu Wang³

¹Department of Electronic Engineering, De Lin Institute of Technology, Tu-cheng, Taipei, Taiwan, R.O.C.

²Department of Computer Science and Information Engineering, National Taitung University Taitung, Taiwan, R.O.C.

³Graduate Institute of Computer and Communication Engineering National Taipei University of Technology, Taipei, Taiwan, R.O.C.

Abstract— In this paper, we study LUT-based Field programmable gate arrays (FPGAs) architectural parameters association with the LUT cluster size N and input number k for FPGA targeting at delay, area, and power to improve performance. This performance problem has been proved that the different pair of LUT cluster size N and input number k can lead to different trade-off the FPGA in delay, area, and power consumption previously. Hence, we present an efficient methodology based on the grey decision making technology to apply LUT-based FPGA physical design of the electronic design automation (EDA) tools. This methodology can cluster association with the LUT cluster size N and input number k to FPGA optimal performance architecture. Experimental results demonstrate that our proposed can get optimal performance architecture to minimum area, energy, and delay in association with the different pair of LUT cluster size N and input number k decision for MCNC benchmark circuits, respectively. Furthermore, our results show that the min-delay logic block architecture has cluster size 6 and input number k size 7, the min-area logic block architecture has cluster size 8 and input number k size 4, and the min-energy logic block architecture has cluster size 8 and input number k size 4.

1. INTRODUCTION

FPGA have become an attractive stem from the ASIC due to effective issues such as short implementation series, short time to market, and low manufacturing. An FPGA chip consists of the programmable logic blocks, programmable interconnections, and programmable I/O pads, which are placed on a two dimensional array in the FPGA chip.

In 2002, Lai and Kao [1] proposed that the peculiar FPGA can be superior to the identical FPGA for various objectives. For example, Figures 1(a) and 1(b) are two different mappings using 4-LUT and 3-LUT identical FPGAs to implement the same circuit. It can be seen that the 4-LUT mapping uses fewer pins than the 3-LUT mapping. However, the number of memory bits in the 4-LUT mapping is greater than the number of bits in the 3-LUT mapping. Figure 1(c) illustrates the circuit mapped into a peculiar FPGA that provides 3-LUT and 4-LUTs. At the same time, the area efficiency in the peculiar FPGA is better than the identical FPGAs to either 4-LUT or 3-LUTs.

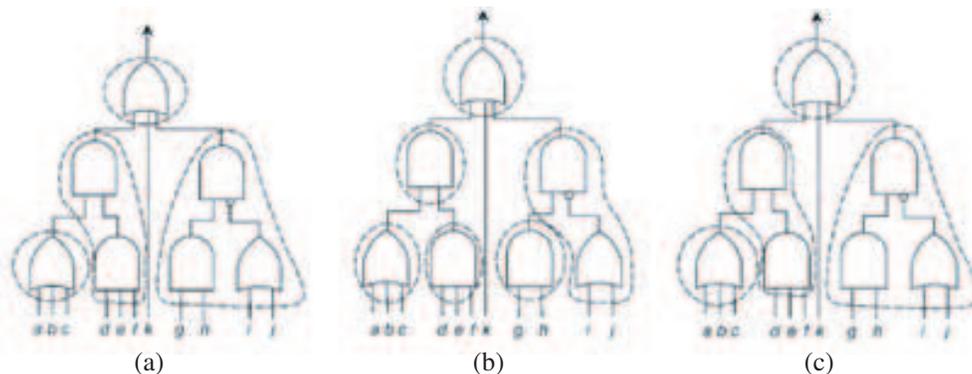


Figure 1: Three mappings of a given circuit with various FPGAs.

Previous literature [1, 2] has demonstrated that different pair of LUT input number k and cluster size N leads to different trade-off the FPGA in delay, area and, power consumption. It will

be interesting to investigate how the pairs LUT cluster size N and input number k affect FPGA delay, area, and power consumption. In this paper, we address grey decision technology for FPGA physical design flow synthesis. This methodology can cluster association of the LUT cluster size N and input number k to FPGA optimal performance architecture in delay, area, and power.

The remainder of this paper is organized as follows. Section 2 discusses the proposed grey decision model. Section 3 shows the Example of application. Finally, a conclusion is presented in Section 4.

2. GREY DECISION MODEL

Since Deng [3] first investigated grey theory, the theory has been utilized for various applications [4, 5]. The theory includes five major parts: grey relational analysis, grey predication, grey decision making, grey programming and grey control. In particular, grey decision theory has an effective method used to solve uncertainty problems with discrete data and incomplete information.

The system decision making is a work of engineering development. Decision making includes the events, strategies and targets. We use this method to decide the pairs of LUT clustering size N and input number k to FPGA optimal performance architecture. The general procedures of Grey decision making theory are as follows.

First, define the events set $A = \{a_i\} (i = 1, 2, \dots, n)$, strategies set $B = \{b_i\} (i = 1, 2, \dots, n)$ and make $S = A \times B = \{S_{ij}(a_i b_j) | a_i \in A, b_j \in B\}$ to be the situation. A strategy indicates any kind of result of a decision making process. One event and strategy are making up the specific situation pairs.

Second, define the target sequence as $\{1, 2, \dots, p\}$. The target refers to the indicators when evaluating the strategies. Let $u_{ij}^{(p)}$ be the sample effect value of situation S_{ij} on target p . Let $r_{ij}^{(p)}$ be the effect measure of situation S_{ij} on the target p . Effect measure $r_{ij}^{(p)}$ and sample effect value $u_{ij}^{(p)}$ are a pair of mapping. $\gamma : u_{ij}^{(p)} \rightarrow r_{ij}^{(p)} \in [0, 1]$. The value of the sample effect will be transformed into a value between 0 and 1.

Table 1: Experimental data of the strategies vs. area, energy power and delay.

Strategies Sets (N, k)	Target 1 Area (1E+06) (number of min-width transistor area)	Target 2 Power (nano joule/cycle)	Target 3 Delay (nano second)
(6, 3)	8.35	5.49	15.40
(6, 4)	6.75	4.10	12.10
(6, 5)	7.35	5.51	11.80
(6, 6)	9.25	5.55	11.25
(6, 7)	9.95	5.85	10.75
(8, 3)	9.40	6.10	14.75
(8, 4)	6.74	3.95	12.05
(8, 5)	8.69	5.09	11.55
(8, 6)	9.60	5.65	11.30
(8, 7)	11.00	6.51	10.87
(10, 3)	11.20	6.30	14.87
(10, 4)	6.90	3.90	12.20
(10, 5)	9.25	5.01	11.35
(10, 6)	10.70	6.31	11.25
(10, 7)	12.25	7.10	11.10
(12, 3)	10.1	6.25	14.75
(12, 4)	8.59	4.51	11.70
(12, 5)	9.60	5.49	11.75
(12, 6)	10.1	6.20	11.80
(12, 7)	13.25	9.90	12.00

Third, define the effect measure. Three methods [6] of effect measure's analysis are upper effect measure, lower effect measure and medium effect measure. The upper effect measurement for the maximum target is used, the lower effect measurement for the minimum target is used and the medium effect measurement for the nominal effect weighting of effect sample is used.

Finally, define the decision-making matrix. A decision-making cell is made up of a situation S_{ij} and its effect measure r_{ij} . With the same event a_i , strategies b_1, b_2, \dots, b_n can establish situations $S_{i1}, S_{i2}, \dots, S_{in}$. With the same strategy b_j , events a_1, a_2, \dots, a_n can constitute situations $S_{1j}, S_{2j}, \dots, S_{mj}$. The S_i row of decision-making matrix is $S_i = [r_{i,1}/S_{i1}, r_{i,2}/S_{i2}, \dots, r_{i,n}/S_{in}]$ and decision-making line S_j is $S_j = [r_{1j}/S_{1j}, r_{2j}/S_{2j}, \dots, r_{mj}/S_{mj}]$. Make $M = (r_{ij}/S_{ij})_{m \times n}$ to be the decision-making matrix. $M^{(p)}$ is the decision-making matrix of target p .

$$M^{(\Sigma)} = \begin{bmatrix} \frac{r_{11}^{(\Sigma)}}{S_{11}} & \frac{r_{12}^{(\Sigma)}}{S_{12}} & \cdots & \frac{r_{1n}^{(\Sigma)}}{S_{1n}} \\ \frac{r_{21}^{(\Sigma)}}{S_{21}} & \frac{r_{22}^{(\Sigma)}}{S_{22}} & \cdots & \frac{r_{2n}^{(\Sigma)}}{S_{2n}} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{r_{m1}^{(\Sigma)}}{S_{m1}} & \frac{r_{m2}^{(\Sigma)}}{S_{m2}} & \cdots & \frac{r_{mn}^{(\Sigma)}}{S_{mn}} \end{bmatrix} \tag{1}$$

where

$$r_{ij}^{(\Sigma)} = \sum_{p=1}^p w_p \cdot r_{ij}^{(p)}, \quad \sum_{p=1}^p w_p = 1 \tag{2}$$

In the above formula w_p is the weight of target. If b_j^* is the best strategy, then $r_{ij^*}^{(\Sigma)} = \max_j \{r_{ij}^{(\Sigma)}\}$.

If $a_{i^*}^*$ is the best strategy, then $r_{i^*j}^{(\Sigma)} = \max_i \{r_{ij}^{(\Sigma)}\}$.

Table 2: Effect measure of area, power, delay.

Strategies Sets (N, K)	Effect Measure r_1 of Target 1 Area	Effect Measure r_2 of Target 2 Power	Effect Measure r_3 of Target 3 Delay
(6, 3)	0.8072	0.7104	0.6981
(6, 4)	0.9985	0.9512	0.8884
(6, 5)	0.9170	0.8647	0.9110
(6, 6)	0.7286	0.7027	0.9556
(6, 7)	0.6774	0.6667	1.0000
(8, 3)	0.7170	0.6393	0.7288
(8, 4)	1.0000	0.9873	0.8921
(8, 5)	0.7756	0.7662	0.9307
(8, 6)	0.7021	0.6903	0.9513
(8, 7)	0.6127	0.5991	0.9890
(10, 3)	0.6608	0.6190	0.7229
(10, 4)	0.9768	1.0000	0.8811
(10, 5)	0.7286	0.7784	0.9471
(10, 6)	0.6299	0.6181	0.9556
(10, 7)	0.5502	0.5493	0.9685
(12, 3)	0.6673	0.6240	0.7288
(12, 4)	0.7846	0.8647	0.9188
(12, 5)	0.7021	0.7104	0.9149
(12, 6)	0.6673	0.6290	0.9110
(12, 7)	0.5087	0.3939	0.8958

3. EXAMPLE OF APPLICATION

In 2005, Fei Li et al. [7] have been shown that architectural parameters such as cluster and LUT sizes affect the total FPGA power, delay and, area. This study combines grey decision model (Section 2) and Fei Li et al. [7] to LUT-based FPGA optimal performance in association with the different pair of N and k . The event is defined the FPGA optimal performance. The strategies are $S_{11}, S_{12}, \dots, S_{NK}$ in $(6, 3), (6, 4), \dots, (12, 7)$ for different pair of cluster size N and LUT input number k , respectively. The targets to measure the effect are delay, area, and power minimization. From Fei Li et al. [7], the strategies vs. area, energy power and delay data are showed in Table 1.

Target 1, Target 2, and Target 3 are expressed the representation of area, power, and delay, so is suitable for lower effect measure. For the first area target $\{u_{ij}^{(1)}\} = (8.35, 6.75, 7.35, \dots, 10.1, 13.25)$, so $u_{\min} = \min_i \min_j \{u_{ij}^{(p)}\} = 6.74r_{1,1}^{(1)} = \frac{\min\{u_{ij}^{(p)}, u_0\}}{\max\{u_{ij}^{(p)}, u_0\}} = \frac{6.74}{8.35} = 0.8072, r_{2,1}^{(1)} = 0.9985, \dots, r_{19,1}^{(1)} = 0.6673, r_{20,1}^{(1)} = 0.5087$. The effect measure of area, power, and delay can be calculated for lower effect measure in Table 2.

The effect measure r_1, r_2 , and r_3 of Target 1, Target 2, and Target3 represent the area, energy, and delay in FPGA all the logic block architectures, respectively. The weight of them is given in average. Using Formula (2), $r_{1-1}^{(\Sigma)}$ can be calculated $r_{1-1}^{(\Sigma)} = \sum_1^k w_k \cdot r_{1-1}^{(K)} = (0.8072 + 0.7104 + 0.6981)/3 = 0.7385$. So, integrated effect measure of area, energy, and delay can be seen in Table 3.

From the above table, we can see that the integrated effect measure sizes are $r_{1-7} > r_{1-12} > r_{1-2} > \dots r_{1-6} > r_{1-11} > r_{1-20}$. So, when FPGA work in optimal performance architecture, integrated effect measure of FPGA system in the min-area logic block architecture has cluster size 8 and input number k size 4. Using the same method as operating the weight of them is given inequality in area, energy, and delay. Our results show that the min-delay logic block architecture has cluster size 6 and input number k size 7 in weight $w = (0.05, 0.05, 0.9)$. The min-energy logic block architecture has cluster size 8 and input number k size 4 in weight $w = (0.05, 0.9, 0.05)$.

Table 3: Integrated effect measure of area, energy, and delay.

Strategies Sets (N, K)	Integrated Effect Measure	
(6, 3)	r_{1-1}	0.7385
(6, 4)	r_{1-2}	0.9461
(6, 5)	r_{1-3}	0.8976
(6, 6)	r_{1-4}	0.7956
(6, 7)	r_{1-5}	0.7814
(8, 3)	r_{1-6}	0.6951
(8, 4)	r_{1-7}	0.9598
(8, 5)	r_{1-8}	0.8242
(8, 6)	r_{1-9}	0.7812
(8, 7)	r_{1-10}	0.7336
(10, 3)	r_{1-11}	0.6676
(10, 4)	r_{1-12}	0.9527
(10, 5)	r_{1-13}	0.8181
(10, 6)	r_{1-14}	0.7345
(10, 7)	r_{1-15}	0.6893
(12, 3)	r_{1-16}	0.6734
(12, 4)	r_{1-17}	0.8561
(12, 5)	r_{1-18}	0.7758
(12, 6)	r_{1-19}	0.7358
(12, 7)	r_{1-20}	0.5995

4. CONCLUSION

This study used grey decision making theory of one event, twenty strategies and three targets. Experimental data of twenty strategies operating of the FPGA work was employed to make decision. Min-area logic block, min-energy logic block and min-delay logic block are prominent indicators of optimal performance architecture. The descriptions of research results are as follows:

- 1) By using grey decision making theory the research result showed that the min-delay logic block architecture has cluster size 6 and input number k size 7, the min-area logic block architecture has cluster size 8 and input number k size 4, and the min-energy logic block architecture has cluster size 8 and input number k size 4.
- 2) By comparing different working time of strategies operating, the research result showed that the weights of targets play a very important part in grey decision making theory.
- 3) By using grey decision making theory approach to all integrated effect measure of the three targets, the research results showed that the best extraction way is conformed to the [7] results.

ACKNOWLEDGMENT

The authors would like to thank the National Science Council of the Republic of China, Taiwan, for financially supporting this research under Contract No. NSC 98-2221-E-237-010.

REFERENCES

1. Lai, Y.-T. and C.-C. Kao, "Technology mapping algorithm for heterogeneous field programmable gate arrays," *IEE Proc. — Computer. Digit. Tech.*, Vol. 149, No. 6, 249–255, Nov. 2002.
2. Ahmed, E. and J. Rose, "The effect of LUT and cluster size on deep-submicron FPGA performance and density," *Proc. ACM Int. Symp. Field-Programmable Gate Arrays*, 3–12, Monterey, CA, Feb. 2000.
3. Deng, J. L., "Introduction to grey system theory," *The Journal of Grey System*, Vol. 1, 1–24, 1989.
4. Wu, J.-O., C.-S. Cheng, and C.-C. Tsai, "Application of grey relational analysis to minimal clock skew routing in SoC," *The Journal of Grey System*, Vol. 16, No. 3, 221–234, 2004.
5. Tsai, C.-C., J.-O. Wu, and T.-Y. Lee, "GDME: Grey relational clustering applied to a clock tree construction with zero skew and minimal delay," *IEICE Trans. on Fundamentals of Electronics, Communications and Computer Sciences*, Vol. E91-A, No. 1, 365–374, Jan. 2008.
6. Luo, Y. X., *Grey System Theory and Approach to Mechanical Engineering*, National University of Defense Technology Press, 152–174, Changsha, 2001.
7. Li, F., Y. Lin, L. He, D. Chen, and J. Cong, "Power modeling and characteristics of field programmable gate arrays," *IEEE Trans. on CAD*, Vol. 24, No. 11, 1712–1724, 2005.

Finite Element Analysis of Electromagnetic Valve Actuation for Engine

Shizuo Li

College of Electrical Engineering, Guangxi University, Nanning, Guangxi 530004, China

Abstract— A new electromagnetic structure is introduced in the research of electromagnetic valve actuation (EVA) for engine, it uses air as damp and has compact volume. The magnetic field of EVA has been calculated accurately by finite element method, the electromagnetic energy of different gap can be got and the electromagnetic force can be calculated by virtual displacement method. Calculation shows that the electromagnetic force of this new structure can meet the dynamic requirements of EVA. Research result is beneficial to optimization of structure parameter and achievement of control method for EVA.

1. INTRODUCTION

Electromagnetic valve actuation (EVA) is a new type of valve actuation, it uses electromagnet to drive and control valve instead of camshaft. EVA can simplify the structure of engine, and can control every valve separately.

EVA must meet following performance: (1) Compact volume in order to install on the surface of cylinder with several valves. (2) Provide enough electromagnetic force. (3) Open and close in high speed to fit 500–6000 r/s of engine. (4) Operate normally in high temperature which may reach 200°C. (5) Soft landing to avoid rigid collision.

This paper presents a new structure of EVA, and finite element method is used to calculate magnetic field of it to get the electromagnetic force under different gap.

2. CALCULATION METHOD

2.1. EVA Structure

A new structure of EVA is introduced and illustrated in Fig. 1, it is rotational symmetry. There are two symmetrical cores and windings on the two sides, the current in the two windings are equal in magnitude but opposite in direction. The central core and winding is fixed on axis, and can move to right or left side which can drive valve open or off. If the current of central winding has same direction with that of right winding, the axis moves to right side, and vice versa, so we can control the current of central winding to drive valve.

This structure of EVA is very compact in volume, and uses air as damp to solve the problem of soft landing instead of spring which is effected by temperature easily and has big volume. Only the current of central winding is need to control and control circuit becomes simple.

2.2. Finite Element Method

Finite element method is used to calculate the magnetic field of EVA. According electromagnetic field theory, the magnetic field of EVA satisfies:

$$\nabla \times H = J \quad (1)$$

$$\nabla \cdot B = 0 \quad (2)$$

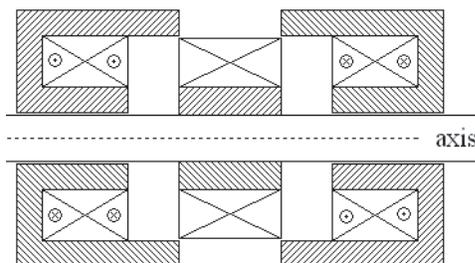


Figure 1: Structure of EVA.

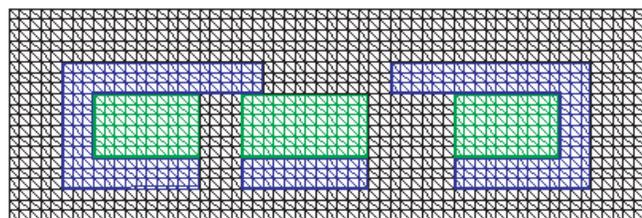


Figure 2: Triangular element dissection.

Here J is current density, H is magnetic density, B is magnetic flux density, the relation of H and B is $H = B/\mu$, μ is magnetic conductivity.

Vector potential A is introduced, and has

$$B = \nabla \times A \tag{3}$$

Partial differential equation of A is:

$$\nabla^2 A = -\mu J \tag{4}$$

Due to rotational symmetry, cylindrical coordinates system (r, α, z) is convenient. Current density J only has α^0 component, so A only has α^0 component too. A is independent of variable α and only varies with variable r and z , so A can be treated as a scalar function of variable r and z . The relation of B and A is:

$$B = -\frac{\partial A}{\partial z}r^0 + \left(\frac{A}{r} + \frac{\partial A}{\partial z}\right)z^0 \tag{5}$$

Partial differential Equation (4) is corresponding to functional extremum problem as following:

$$P(A) = 2\pi \int_S \frac{1}{2\mu} B^2 r dr dz - 2\pi \int_S J A r dr dz = \min \tag{6}$$

Here S is the solving zone on the meridional plane.

For simpleness, equivalent triangle element dissection is used. Fig. 2 shows the triangular element dissection, there are 2400 elements and 1281 nodes.

2.3. Electromagnetic Energy and Force

Electromagnetic energy distributes continuously in the whole solving zone with a density of $HB/2$. There are two calculation methods of electromagnetic energy:

$$W = \frac{1}{2} \int_V HB dV = 2\pi \int_S \frac{1}{2\mu} B^2 r dr dz \tag{7}$$

$$W = \frac{1}{2} \int_V J A dV = 2\pi \int_S \frac{J A}{2} r dr dz \tag{8}$$

Two calculation methods are same theoretically, but are different in actual numerical calculation. Formula (7) needs differential operation of vector potential A before integral operation and cause more error, formula (8) integrates to vector potential A directly without differential operation and has good accurateness.

According to virtual displacement method, electromagnetic energy differentiates with respect to displacement is equal to electromagnetic force, so:

$$F = \frac{\partial W}{\partial z} \tag{9}$$

Here z is the displacement of axis.

3. CALCULATION RESULT

Current density of left winding J_1 is 17 MA/m², current density of right winding J_2 and current density of central winding J_3 are -17 MA/m² respectively. The current of central winding has same direction with that of right winding, so the axis moves to right side.

Supposed that the gap between left winding and central winding is δ with the range from 1 mm to 11 mm. In order to get the accurate movement process, magnetic field under different δ has been calculated by finite element method.

3.1. Distribution of Magnetic Field

Figure 3 shows the distribution of magnetic field under different δ . Magnetic field of left and right winding which have long core is much bigger than that of central winding with short core.

With the increasing of the gap δ , magnetic field of right winding becomes bigger and that of left winding becomes smaller.

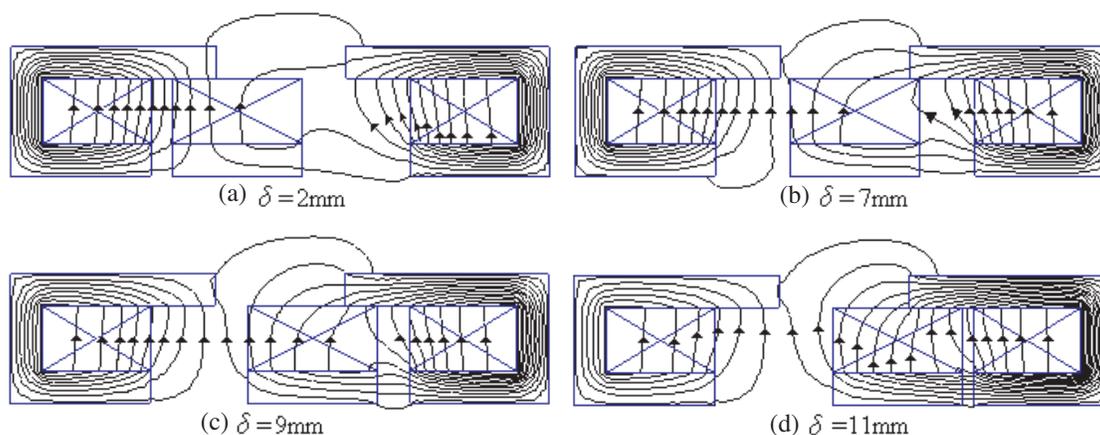


Figure 3: Distribution of magnetic field.

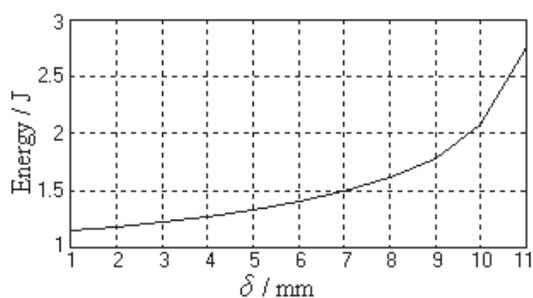


Figure 4: Curve of electromagnetic energy.

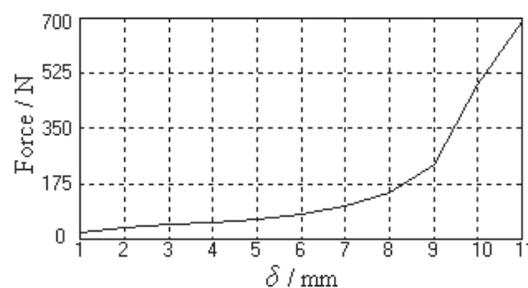


Figure 5: Curve of electromagnetic force.

3.2. Electromagnetic Energy and Force

The curve of electromagnetic energy with respect to δ is shown in Fig. 4. Analysis illustrates that magnetic field produced by central winding will enhance magnetic field produced by right winding and reduce magnetic field produced by left winding. When the gap δ is small, the effect of counteraction is larger than that of enhancement, so the energy is little. With the increasing of the gap δ , the energy becomes bigger more and more.

Using numerical differentiation method, electromagnetic force of formula (9) can be calculated approximatively by:

$$F \approx \frac{\Delta W}{\Delta \delta} = \frac{W(\delta_1) - W(\delta_2)}{\delta_1 - \delta_2} \quad (10)$$

Here $W(\delta)$ is the electromagnetic energy with respect to δ .

Figure 5 shows the curve of the electromagnetic force with respect to δ . Inertial mass of axis is less than 300 g, time to move 10 mm from left side to right side for axis is less than 10 ms, and this can meet the dynamic requirements of EVA.

4. CONCLUSIONS

This new structure of EVA has compact volume and good dynamic performance, it uses air damp instead of spring.

Finite element method is fit to calculate magnetic field of EVA, and electromagnetic force can be determined by virtual displacement method. Calculation shows that the electromagnetic force of this new structure can meet the dynamic requirements of EVA. Research result is beneficial to optimization of structure parameter and achievement of control method for EVA.

REFERENCES

1. Giglio, V., B. Iorio, G. Police, et al., "Analysis of advantages and of problems of electromechanical valve actuators," SAE Paper [C], 2002-01-1105, 2002.
2. Chihaya, S., S. Hisao, U. Atsushi, et al., "Study on variable valve timing system using electromagnetic mechanism," SAE Paper [C], No. 2004-01-1869, 2004.

3. Salber, W., H. Kemper, F. Staay, et al., “The electro-mechanical valve train — A system module for future powertrain concepts,” *MTZ*, Vol. 61, 12, 826–829, 2000 and *MTZ*, Vol. 62, No. 1, 44–55, 2001.
4. Tai, C., *Modeling and Control of Camless Engine Valvetrain Systems*, University of California at Los, Los Angeles, 2002.
5. Park, S. H., J. H. Lee, J. S. Yoo, et al., “A developing process of newly developed electromagnetic valve actuator — Effect of design and operating parameters,” SAE Paper [C], No. 2002-01-2817, 2002.
6. Wang, Y., T. Megli, M. Haghgoie, et al., “Modeling and control of electromechanical valve actuator,” SAE paper [C], No. 2002-01-1106, 2002.

Seismic Traveling Macroscale Irregularities at Ionospheric F2-region on Data of Distance Sounding

U. K. Kalinin¹, N. P. Sergeenko², and M. V. Rogova²

¹Fedorov Institute of Applied Geophysics, Moscow, Russia

²Pushkov Institute of Terrestrial Magnetism, Ionosphere and Radiowaves Propagation Troitsk, Moscow, Russia

Abstract— The properties of travelling ionospheric macroscale irregularities (their horizontal sizes are 1–4 thousand kms, excesses of a background are 15–30%), formed in F2 layer as a result of strong earthquake in time of heliogeophysical perturbations are explored. The irregularities arise 10–15 h before earthquakes and move horizontally with a transonic speed on distances of a few thousand kms up to round — the-world trajectories focused approximately along an arc of a major circle, transiting above epicenter region. Information of spent examination was caused by spatial-time differences of dynamics of F2-layer of an ionosphere in time of the heliogeophysical disturbances from dynamics of quasi-causative seismic macroscale ionospheric irregularities.

The complex analysis of the arrays of critical frequencies of F2-layer of an ionosphere (foF2) of a world network of automatic ground ionospheric stations of vertical sondage was carried out. Also the ionograms of board sondage of space station “Mir” were used for the analysis. Global TEC maps was used to detect the ionospheric irregularities associated with earthquake precursors. Analyses of all these data confirmed a hypothesis about occurrence of macroscale positive ionospheric irregularities in the vicinity of strong earthquakes ($M > 5$) epicenters also.

1. INTRODUCTION

This work is devoted to the study of the possibilities of detection of traveling macroscale irregularities (TMIs) of seismic origin in the main peak of ionosphere at the time of heliogeophysical disturbances. The sizes of irregularities reach $1 \sim 3$ thousand kms, lifetimes exceed 10^4 sec, and moving distances are $10 \sim 15$ thousand kms, the speeds are approximate to sonic velocity. It was also established that the formation of TMI precedes $10 \sim 15$ hrs to the moment of catastrophic earthquakes with magnitude $M \geq 6$ and localization ~ 1 thousand kms from the epicenter of the earthquake [1–3]. Perturbations of electron concentration, commensurable in sizes and contrast range with seismic TMIs, but completely different character of temporal and spatial changes, can appear on the conditions of ionospheric storms and substorms. Below we review features of occurrence of TMI in quiet and disturbed heliogeophysical conditions.

2. DATA ANALYSIS

Files of values of relative variation of critical frequencies of F2 layer of the ionosphere by the global network of ionosphere stations of vertical sounding δfoF2 were used in the analysis as a main source of information about seismic TID: $\delta\text{foF2} = (\text{foF2} - \text{foF2}_m) / \text{foF2}_m$, foF2_m — moving median [4]. $\{\delta\text{foF2}\}$ are used for detection and numerical characterization of ionosphere disturbances [5]. It does not depend on multiplicative errors of representation foF2 due to the use of various scales at stations of vertical sounding, gives “protection” against inadequate influence of major “emissions” and excludes regular seasonal and daily variability.

The seismic TMIs are selected as a local irregularity with extreme value $(\delta\text{foF2})_{\max} \geq 15\%$ at the middle latitudes and $(\delta\text{foF2})_{\max} \geq 10\%$ at low latitudes. It is also supposed that during of $2 \sim 3$ hours diversion should be more than 10%. Kp, Dst, and AE indexes were used for the description of geomagnetic situation.

2.1. The Quite Conditions

In Fig. 1(a) we can see an example of variations $\delta\text{foF2}(t)$ for different ionosphere stations for the earthquake, which occurred at Alyaska on 09.03.1985 ($\varphi = 66.6\text{N}$; $\lambda = 150.5\text{W}$; $t_0 = 14^{\text{h}}08^{\text{m}}\text{UT}$; $M = 6.2$), a vertical line marks the moment of the earthquake, an oblique line represents variant of propagation of signals most likely related to a developing earthquake. The appearing disturbances before the earthquake travel to Asia, Indian ocean and up to La Renion ionosphere station.

The first signals are registered $\Delta t = -13$ hours ahead of time. Amplitudes of signals are $\delta\text{foF2}(t) > 20\%$. According to the data in Fig. 1(c), positive impulses appeared on quiet geomagnetic background: $K_p \leq 2$, $\text{AE} < 200\text{ nT}$, Dst did not fall below -20 nT .

The Mercator projection of world map is represented in Fig. 1(b), the location of epicenter is shown by cross. Dark circles show the locations of ionosphere stations, where positive disturbances were observed. The dotted curve shows the trajectory of moving of this disturbance. The trajectory is focused approximately along the part of an arc of a major circle, transmitted above the epicenter region. The light circles show location of the other ionosphere stations, where there were either negative disturbances or no disturbances or no observations in this time for some reasons.

2.2. The Disturbance Conditions

2.2.1. Ionosphere Substorm

Figure 2 represents the data for the Caribbean earthquake that took place on 16.03.1985 ($\varphi = 17.5\text{ N}$; $\lambda = 62.5\text{ W}$; $t_0 = 14^{\text{h}}53^{\text{m}}\text{ UT}$, $M = 6.8$). Fig. 2(c) shows there was a substorm ($AE_{\text{max}} \sim 400\text{ nT}$) at 14 h UT, 1 hour before the earthquake. This substorm took place on a quiet geomagnetic background, Dst did not fall below -20 nT . The $\delta foF2(t)$ data (Figs. 2(a) and (b)) allows us to draw

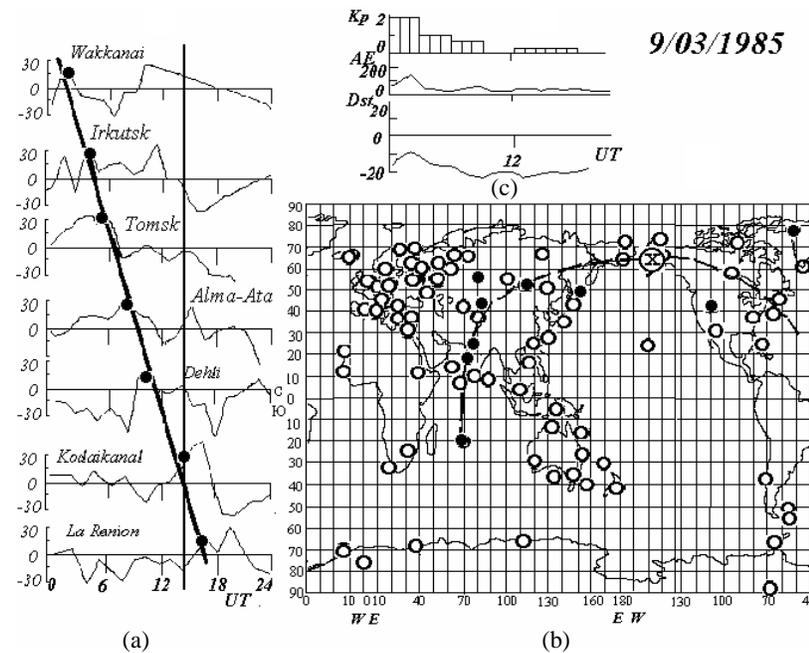


Figure 1: $\delta foF2(t)$ variations at different ionosphere stations for the Alyaska earthquake on 09.03.1985.

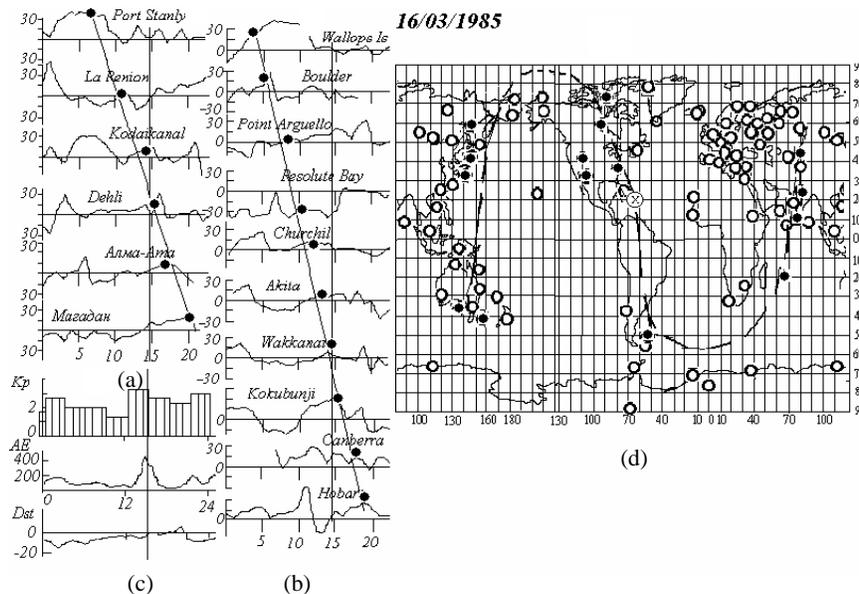


Figure 2: $\delta foF2(t)$ variations at different ionosphere stations for the Caribbean earthquake on 16.03.1985.

a common line of delays for all stations altogether. This line gives $\Delta t = -13$ hrs as the estimate time of the appearance of irregularity above the epicenter.

The interval of distances is significant here: $-6 \sim 17 \cdot 10^3$ km. Positive impulses appeared on a quiet geomagnetic background. The substorm began approximately 4 hrs later. The trajectory of a motion of seismic disturbances corresponds to a site of an arc of a major circle, directional in opposite sides from the epicenter.

2.2.2. Ionosphere Storm

Figure 3 represents data for the Indonesian earthquake on 01.03.1985 ($\varphi = 1.4$ S; $\lambda = 119.6$ E; $t_0 = 17^h 11^m$ UT; $M = 6.4$). $\delta f_oF_2(t)$ variations are showed for the various stations located in Asia, Europe, Australia, Northern America at the left part of this figure. The variations of geomagnetic indices are presented in Fig. 3(c) which testify that 01.03.1985 the moderate geomagnetic disturbance ($Dst_{max} \sim -50$ nT, $Kp_{max} = 6$) was observed.

The substorm has also taken place at ~ 5 h UT on disturbance background. The $\delta f_oF_2(t)$ data in Fig. 3(b) testify a biphasic ionosphere disturbance in the F2 layer. All diagrams contain clear positive perturbations with amplitude $15 \sim 30\%$ and duration $5 \sim 7$ hrs and in day local time beginning (5 h UT) on the first day of the geomagnetic disturbance. Approximately at 18 h LT the perturbation has transformed in a negative phase at all stations.

It is obvious that when observed simultaneously on ionosphere stations located at various latitudes and longitudes the positive phase of ionosphere storm cannot be identified as seismic precursor, though they have happened in “the necessary” time.

However, Fig. 3(a) illustrates the $\delta f_oF_2(t)$ data on a chain of stations from Manila and Vanimo, posed in a region of forthcoming earthquake, through domestic stations in the Asia and Europe (Irkutsk, Tomsk, Sverdlovsk, Gorky, Moscow), posed approximately 5 thousands kms from the earthquake epicenter, up to north Africa station Ouagadougou. The positive impulses — harbingers of earthquake have appeared in a region of epicenter ~ 13 hrs before the beginning earthquake. The time delays of their occurrence at the relevant stations specify an apparent velocity of their travel ~ 1000 kms/hrs. In Fig. 3(c) dotted line indicated a possible direction of a motion of these TIDs.

In Fig. 3(a) the chain of $\delta f_oF_2(t)$ impulses with duration $3 \sim 4$ hrs was observed both in day time, and in evening local time, while the positive phase of a biphasic ionosphere disturbance usually occurs only in day time (06 \sim 18 hrs LT) and last not less than $7 \sim 8$ hrs [5].

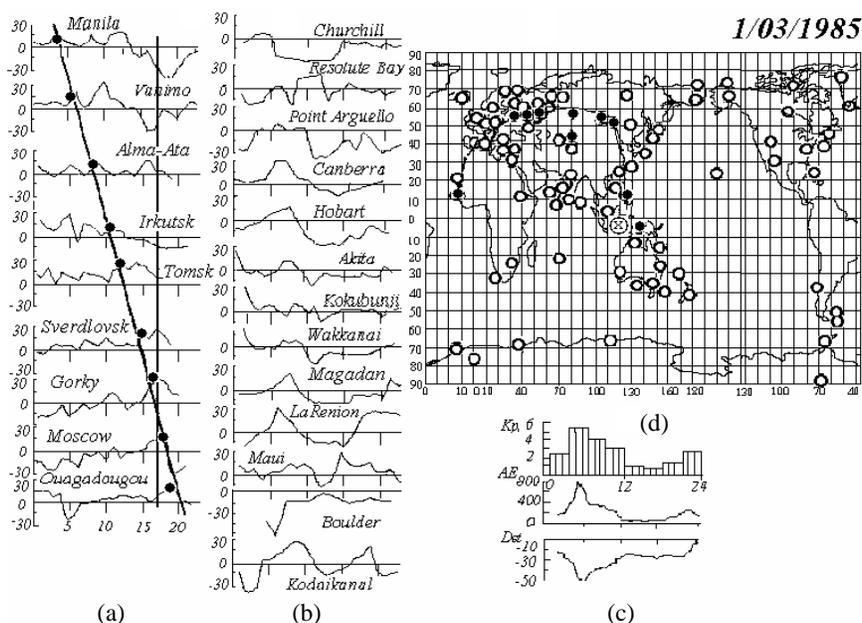


Figure 3: $\delta f_oF_2(t)$ variations at different ionosphere stations for the Indonesian earthquake on 1.03.1985.

3. DISCUSSION

Situations are presented when apparent TIDs in F2 layer of the ionosphere are presumably linked with developing earthquake. We have given examples when they arose in quiet conditions, in

conditions of isolated substorm on the backgrounds of quiet ionosphere and during ionosphere storm with positive initial phase. Undoubtedly, the above given examples do not cover entire spectrum of ionosphere variations.

The changes of electron concentration of F2 layer from storm to storm are significant in the sense that two storms never existed with the same behavior of ionosphere parameters. The basic approach to select macro-scale irregularities consists of the use of differences in dynamics of ionosphere variations, characteristic for disturbed heliogeophysical conditions and in dynamics of TIDs, inconvertible under the shape and velocity of a motion, but with the casual moments of occurrence. Identifications of the dynamics of TIDs require processing data for the territories of hundreds of thousand square kms.

Basic differences between TIDs and effects of magnetosphere perturbations in F2 layer are:

- ◆ During an ionosphere substorm electron concentration is incremented as a result of activity of zonal and meridian electric fields as well as under IGW in the day time after preliminary decrease [6]. The trajectories of movement of TIDs coincide with the arcs of the major circle, whereas the perturbations, which arise as a result of IGW generation during a substorm, are usually spread from high latitudes to low in a meridian direction. IGW generation is delayed by 1 hr with respect to the AE-index;
- ◆ Ionosphere storm represents a continuous process lasting from several hours to several days. Parameter variations have global character. In the absence of AE outbursts, positive impulses recorded in $\delta f_oF_2(t)$ on the background of negative storm phase can be related to seismic TIDs. Positive perturbations of electron concentration can be observed during ionosphere storms at the transitional and low latitudes; as well as at the midlatitudes in the early stages ionosphere storm, if this storm begins at a day time.

4. CONCLUSION

The above reviewed situations demonstrate that diagnostics of seismic TIDs is quite possible in many cases, as there are essential differences in the character of time and spatial changes allowing us to distinguish these disturbances from perturbations of solar and magnetosphere origin. However data configurations that are submitted in Figs. 1–3, are not possible for all earthquakes. One of the reasons is in actual properties of a network of ionosphere stations — nonuniform spatial arrangement and, naturally, only on land. Other difficulties are the detection of TIDs on the background of other disturbances in the ionosphere. It is represented that this short-term harbinger could be the essential block in blanket algorithm of the seismic-ionosphere forecast.

REFERENCES

1. Kalinin, U. K. and N. P. Sergeenko, "Mobile solitary macroinhomogeneities appearing in the ionosphere several hours before catastrophic earthquakes," *Doklady of Russian Academy of Sciences / Earth Science Section*, Vol. 387, No. 1, 105–107, 2002.
2. Kalinin, U. K., A. A. Romanchuk, N. P. Sergeenko, and V. N. Shubin, "The large-scale isolated disturbances dynamics in the main peak of electronic concentration of ionosphere," *Journal of Atmospheric and Solar Terrestrial Physics*, Vol. 65, No. 11–13, 1175–1177, 2003.
3. Sergeenko, N. P. and A. L. Kharitonov, "Short-time magnetosphere — ionosphere predictors of catastrophe earthquakes," *Investigations of Earth from Space*, No. 6, 61–68, 2005 (Russian).
4. Gaivoronskaya, T. V., N. P. Sergeenko, and L. A. Udovich, "Deviations of critical frequencies of F layer from median values," *Ionospheric Disturbances and Their Influence on a Radio Communication*, 55–73, Moscow, Nauka, 1971 (Russian).
5. Zevakina, R. A., N. P. Sergeenko, E. M. Zgulina, and G. N. Nosova, *Text-book of Short-time Forecast of Ionosphere*, 71, Materials of the World Data Center B, 1990 (Russian).
6. Brunelly, B. E. and A. A. Namgaladze, *The Physics of Ionosphere*, 527, 1988 (Russian).

Based on the Coherent Point Target Monitoring Urban Subsidence in Beijing

Hong-Li Zhao^{1,2}, Jian-Ping Chen^{1,2}, Jing-Hui Fan³, Xiao-Fang Guo³, and Huan-Huan Liu^{1,2}

¹School of the Geosciences and Resources, China University of Geosciences (Beijing), Beijing 100083, China

²Beijing Land Resources Information Development Research Laboratory, Beijing 100083, China

³China Aero Geophysical Survey & Remote Sensing Centre for Land & Resources, Beijing 100083, China

Abstract— Beijing has suffered from serious land subsidence caused by ground water over-extraction. DInSAR can monitor movements on the earth surface at centimeter or even millimeter levels. Acquiring and processing SAR data are relatively cheaper than ground work of leveling or GPS. DInSAR has been considered as a cost-efficient and precise tool.

In this paper, Coherent Point Targets interferometry method was applied to extract the linear deformation information and retrieve the deformation history. The phase regression model is used to estimate the linear deformation rate and height correction. The standard deviation from the regression function serves as a quality measure for the regression function and the derived parameters. Based on 20 ENVISAT ASAR images covering the period from June 2003 to March 2007 the linear deformation rate maps of eastern Beijing area are generated. At the same time, Beijing' subsidence result based on the multi-baseline is compared with the one based on the single-baseline.

1. INTRODUCTION

Land subsidence caused by ground water over-extraction in Beijing is a seriously hazard threatening people and urban infrastructure. Although Beijing's in-situ subsidence monitoring system is necessary and reliable, its sparse data grid in the area is a limitation which makes it difficult for us to understand the whole deformation field. Compared with the traditional monitoring method, differential SAR interferometry (DInSAR) is more cost-efficient and accurate as well. DInSAR can monitor the movements on the earth surface at centimeter or even millimeter level. Therefore, many earth physical phenomena can be discovered and analyzed, such as earthquake, volcano, atmosphere meteorology, ice sheet movement, subsidence, landside, and so on.

However, temporal decorrelation and atmosphere disturbance degrade the accuracy of DInSAR for slow deformation monitoring [1–5]. Atmospheric inhomogeneities produce an atmospheric phase screen (APS) on every SAR image, which can contaminate the results of the deformation monitoring [6, 7]. As the APS can't be eliminated by SAR data processing, many researchers introduce some auxiliary data such as MERIS, MODIS and GPS PWV data into DInSAR data processing. In this article, we apply the Coherent Point Targets interferometry method, which regards the atmosphere disturbance as residual phase and can greatly weaken them, and select those interferograms without obviously affected atmosphere disturbance.

2. METHOD

Coherent point targets' analysis key steps mainly include [8]: (1) Input coregistered and resampled SLC, DEM data; (2) Coherent points selection; (3) Coherent points' data extraction, such as their SLC, DEM, baseline data and so on; (4) Coherent points' differential interferometric phase generation; (5) Phase analysis using iterative model; (6) Result output.

The method of Coherent Point Targets interferometry can retrieve the subsidence phase history, extract the linear deformation information from interferometry phase and greatly weaken atmosphere disturbance. During the process, the interferograms are only interpreted for the selected coherent points [8]. Considering efficiency and data storage, vector format data structures are used instead of the raster data format used in conventional interferometry. The point targets must remain stable over the time period to permit analysis of the phase history.

In addition, a phase regression model is used to estimate the linear deformation rate and height correction and residual phase of the coherent point targets over an area. It is a very step-wise improvement that the information can be extracted from iteration model. The standard deviation from the regression function serves as a quality measure for the regression function and the derived parameters. Besides, for phase unwrapping, the minimum cost flow (MCF) algorithm is

used to minimize the total cost associated with phase discontinuities in the scene. And Delaunay triangulation is used to generate an optimized irregular triangular network of the points to be unwrapped.

3. DATA PROCESSING

For Coherent Point Targets interferometry processing, 20 ASAR images acquired between 2003 and 2007 in Beijing area have been selected. Both single-baseline and multi-baseline method are used to retrieve the deformation information. In the data stack, each SAR image is named as the observation date. In Fig. 1(a), the image 20050309 is selected as the master image, which can optimize the distribution of the baselines. As for the 19 interferometric pairs, the absolute values of perpendicular baselines are less than 900 m, and the time baselines vary from -630 to 735 d. For the data process based on the multi-baseline, certain interferometric pairs are selected, whose perpendicular baselines are less than 300 m and time baselines are less than 600 d. At the same time, the interferometric pairs affected by the atmosphere were picked out.

To minimize spatial distances, it is usually adequate to select a reference point near the center of the area of interest [8]. The starting point for unwrapping should be the ones whose phase is smooth and without deformation. The slopes of the regression functions are directly related to the relative height correction, and the linear deformation rate. For example, we selected a coherent point based in the multi-baseline result and show the regression functions for the height correction and linear deformation rate in Fig. 2. In Fig. 2(a), near the red line the selected point's height correction phases is shown and the slope of the red line means the height correction value. In Fig. 2(b), scattered points near the red line are the selected point's deformation phases and the slope of the red line means the linear subsidence rate.

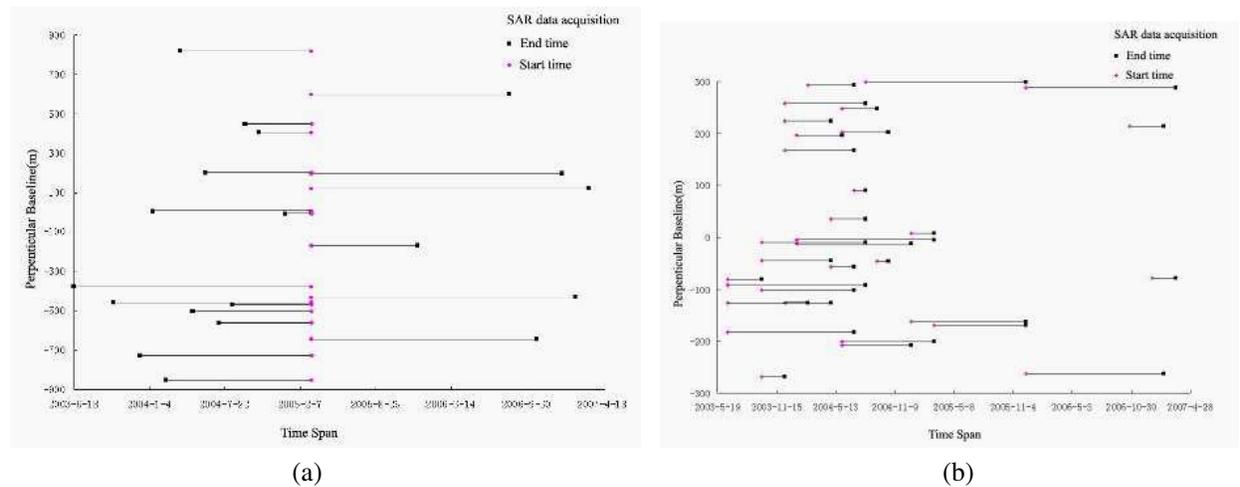


Figure 1: The interferometric pairs combination. (a) For single-baseline method. (b) For multi-baseline method

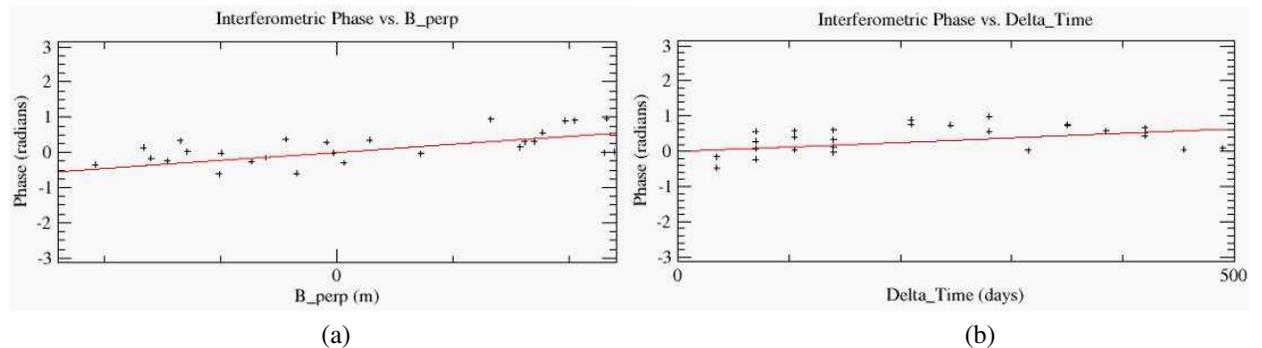


Figure 2: The plots of regression function related to height corrections and subsidence rate for a selected point, (a) related to height correction, (b) related to linear deformation rate.

4. RESULTS AND ANALYSIS

Classification and visualization of the ground subsidence data illustrate the subsidence trend and the boundary between the center and the eastern area of Beijing, which are respectively corresponding to mainly stable ground and increasingly land subsidence-affected ground. The results of coherent point targets are shown in Fig. 3. The single-baseline method and result about Fig. 3(a) have been shown in [9]. The result shown in Fig. 3(b) is calculated by multi-baseline method. In order to analyze the land subsidence, two points were selected in the subsidence center. The deformation histories of the two points are shown in Fig. 4.

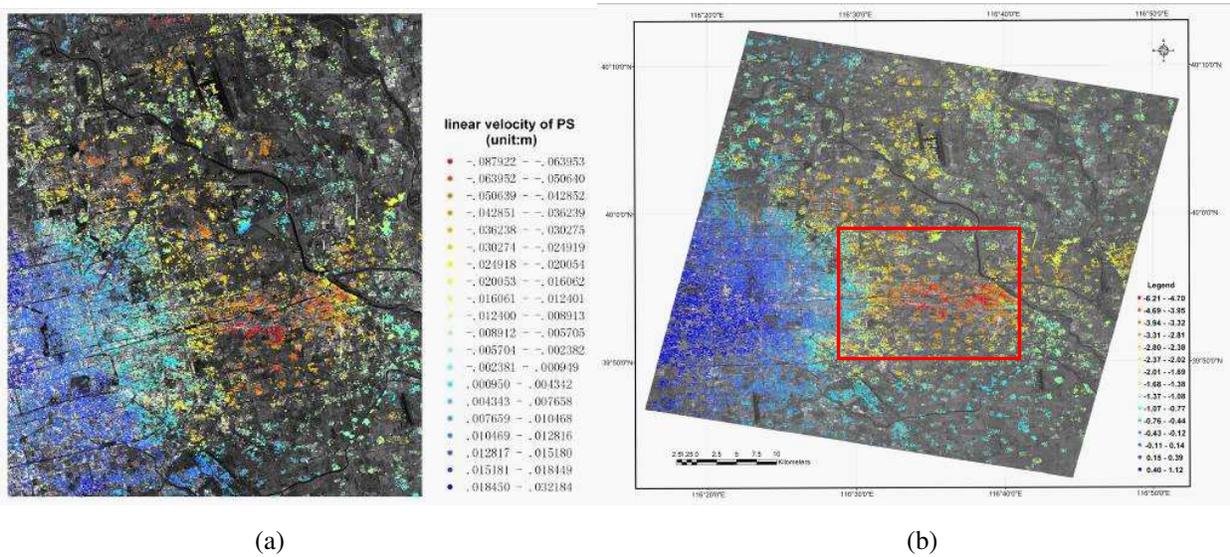


Figure 3: The land subsidence velocity map of test site, (a) based on single-baseline [9], (b) based on multi-baseline

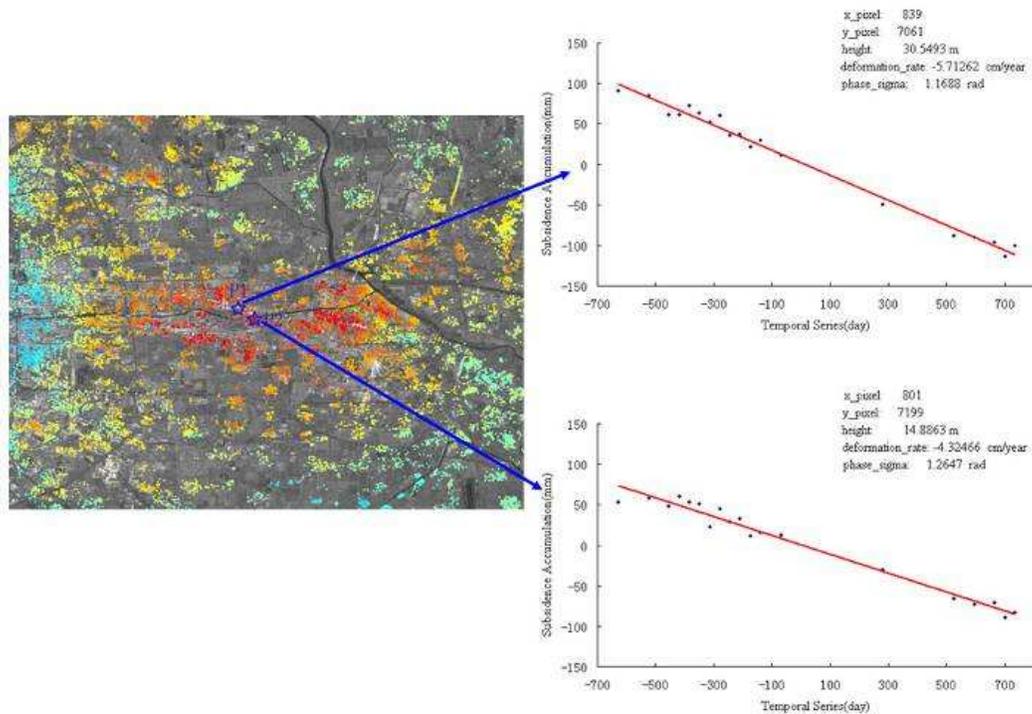


Figure 4: Two plots of land subsidence accumulation of selected points.

From the two results shown in Fig. 3, similar deformation trend can be easily recognized. We can safely draw a conclusion that both single-baseline and multi-baseline method used in this paper can generate a valuable deformation trend.

However, from the comparison between the two results, we can see that the subsidence values based on the multi-baseline (from -6.21 cm to 1.12 cm) are smaller than that based on the single-baseline (from -8.8 cm to 3.2 cm). According to previous ground work [10], the former are closer to the ground truth than the latter. Two reasons can account for the better result of multi-baseline method:

(1) More interferometric pairs than the ones from single-baseline method could be chosen, which can shorten the temporal interval and increase the sampling frequency for the deformation history.

(2) While determine the points whether should be selected by multi-baseline method, not only the amplitude dispersion index and the temporal coherence value are considered, which have been applied in single-baseline method [6, 7, 9]. The standard deviation of the difference between the differential interferometric phase and the calculated phase from regression model is also chosen as a threshold. By the quality controlling threshold, the accuracy of the result should be enhanced.

5. CONCLUSIONS

The concept of coherent point targets technique is briefly introduced. The results of single-baseline and multi-baseline methods are shown and compared. The subsidence deformation field of the eastern Beijing could be sketched by both the two kind of methods. As for the values of the linear deformation rate, the result of multi-baseline method is closer to the ground truth. The deformation values given by this paper should still be validated by ground measurements.

ACKNOWLEDGMENT

This work was supported by the Ministry of Land and Resources (Project No.: 200811053). And we wish to thank Ge Daqing, Hu Qiong and Zhang Ling for their kind helps.

REFERENCES

1. Ge, D., et al., "Surface deformation monitoring with multi-baseline dInSAR based on coherent point target," *Journal of Remote Sensing*, Vol. 11, No. 4, 574–580, July 2007 (in Chinese).
2. Ge, D., et al., "Surface deformation rate derivation based on differential interferograms stack," *Remote Sensing for Land & Resources*, Vol. 1, 25–31, Mar. 2007 (in Chinese).
3. Ge, D., et al., "Surface deformation field monitoring by use of small-baseline differential interferograms stack," *Journal of Geodesy and Geodynamics*, Vol. 28, No. 2, 61–66, Apr. 2008 (in Chinese).
4. Hanssen, R. F., *Radar Interferometry-data Interpretation and Error Analysis*, Kluwer Academic Publishers, New York, 2002.
5. Zebker, H. A., P. A. Rosen, S. Hensley, "Atmospheric effects in interferometric synthetic aperture radar surface deformation and topographic maps," *Geophysical Research Letters*, Vol. 102, 547–563, 1997.
6. Ferretti, A., C. Prati, and F. Rocca, "Nonlinear subsidence rate estimation using permanent scatterers in differential SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, Vol. 38, 2202–2212, 2000.
7. Ferretti, A., C. Prati, and F. Rocca, "Permanent scatterers in SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, Vol. 39, 8–20, 2001.
8. GAMMA Software: Users Guide, 2004.
9. Zhao, H.-L., et al., "Application of PSInSAR for monitoring urban subsidence in Beijing," *PIERS Proceedings*, 739–743, Moscow, Russia, Aug. 18–21, 2009.
10. Sun, C., et al., "Analyses of ground settling in the eastern suburbs of Beijing," *Geotechnical Engineering Field*, Vol. 5, No. 11, 2005 (in Chinese).

A Study of the High Resolution COSMO-SkyMed SAR Data for Ground Subsidence

Hong-Li Zhao^{1,2}, Jing-Hui Fan³, Zhen-Chao Wang¹, Jian-Ping Chen^{1,2},
Huan-Huan Liu^{1,2}, and Xiao-Fang Guo³

¹School of the Geosciences and Resources, China University of Geosciences (Beijing), Beijing 100083, China

²Beijing Land Resources Information Development Research Laboratory, Beijing 100083, China

³China Aero Geophysical Survey & Remote Sensing Centre for Land & Resources
Beijing 100083, China

Abstract— COSMO SkyMed satellites are very suitable for natural disaster monitoring, because they can acquire high quality images with very high resolution. DInSAR can map ground deformation phenomena over tens-of-kilometers-wide area with centimeter-scale accuracy level, and has been considered as a powerful tool. In this paper, COSMO-SkyMed SAR and ASAR data are respectively processed by DInSAR for land subsidence derivation in Tianjin. The comparison between the two preliminary results shows that similar deformation trend can be identified based on the two kinds of SAR data and the deformation values of COSMO's DInSAR result are closer to truth. With the characteristics of high resolution and short wavelength, COSMO data seem more robust for extract high-precision ground deformation information than ASAR data in case of no obvious orbit error existing in both the 2 kinds of data. While COSMO data are suitable for high-precision information extraction in a relatively small and important area, ASAR data are fit for large area, regional surface deformation survey.

1. INTRODUCTION

The satellites of COSMO-SkyMed, which are an Earth observation satellite system funded by the Italian Ministry of Research and Ministry of Defense and conducted by the Italian Space Agency (ASI) [1], have some significant advantages [2]: an area of interest is unmatched revisited 4–6 hours with 4 satellites which can fast response in 18 hours; observations of the area will be repeated several times a day in all-weather conditions; COSMO SkyMed is equipped with a polarimetric X-band SAR; the satellites can get large daily area collection (500 km² in month at 5 m resolution); and satellites can acquire high quality images with very high resolution (1 m for civil applications). So they are very suitable for natural disaster monitoring.

DInSAR can map ground deformation phenomena over tens-of-kilometers-wide area with centimeter-scale accuracy level [3], and has been considered as a powerful tool. Because of the data availability, previous DInSAR applications mainly focusing on ERS, ENVISAT, RADARSAT-1 and JERS SAR data with resolution coarser than 10m. Nowadays, several kinds of high resolution SAR data are available for DInSAR and there are high expectations for the application of these data. In this paper, COSMO-SkyMed SAR data are preliminary processed by DInSAR for land subsidence derivation of Tianjin. The DInSAR results of COSMO-SkyMed and ASAR data are preliminary compared as well.

2. METHOD

In the process of DInSAR, a reference interferogram should be simulated based on the DEM. Then, based on the reference SAR geometry, the interferometric baseline model, the height map in radar geometry and the unwrapped interferometric phase corresponding exclusively to topography is calculated. The main process of traditional two-pass DInSAR method involves: co-registration, interference, removal of flat phase and DEM phase, filtering, unwrapping, conversion of unwrapped phase to elevation change and resampling to map coordinates. This method is easy to be implemented, but residual phase will be left when the interferometric space baseline is not accurate enough. To deal with the possible errors from orbit data, the process of improved DInSAR two-pass method [4–6] is shown in Fig. 1. In case of residual phase is obvious, four key steps in Fig. 1 will be applied, that is: removal of residual fringes from differential interferogram using fringe rate; refinement of differential interferogram using a baseline model; compensation of unwrapped interferogram for residual quadratic phase; refinement of the differential interferogram with improved quadratic phase model.

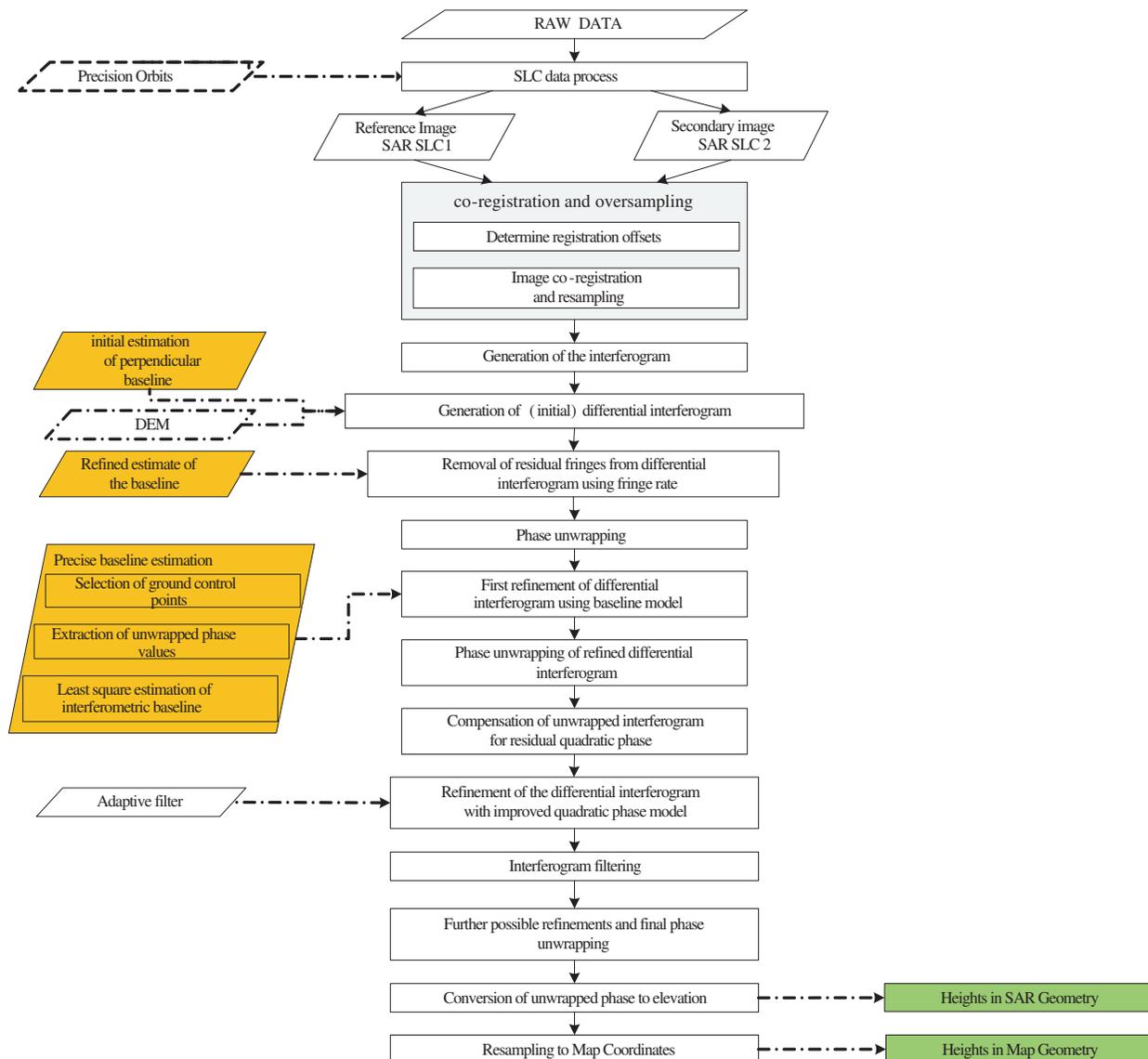


Figure 1: Two-pass DInSAR processing chain.

3. DATA PROCESSING AND RESULTS ANALYSIS

While calculating the land subsidence using differential interferometric phases, correlation of the interferometric pairs is an important factor. In case of serious decorrelation, the phases with almost pure deformation information are difficult to be obtained. To obtain accurate interferometric phase, a proper numbers of pixels with high coherence are necessary. If spatial perpendicular baseline is too long, the difference of look angle will increase, which may cause an offset between radar echo signal data spectrum and target spectrum. When the offset reaches a certain level, the completely loss of coherence will appear [6]. In Table 1, 1 ENVISAT ASAR and 3 COSMO InSAR pairs are listed with their space and time baseline. As for the space perpendicular baseline of COSMO InSAR pairs, I3's should have better coherence than I1 and I2. In addition, I3's the time baseline of is more similar to I4's than I1 and I2. So, in order to compare the deformation DInSAR results between COSMO data and ASAR data, the authors mainly analyze I3 and I4. The COSMO data are acquired in stripmap mode and cover a 40 km-width swath with resolution about 3 m and wavelength about 3.1 cm. The parameters of ASAR data are IS2 mode, 100 km-width swaths, about 20 m resolution, 5.6 cm wavelength.

The two kinds of data all both processed by the improved two-pass DInSAR method shown in Fig. 1. While comparing the results of ASAR and COSMO data, both the whole area of the 2 kinds of images are considered (see Fig. 2 and Fig. 3). From Fig. 2, the difference of the coverage of

Table 1: List of interferometric SAR pairs chosen for processing.

InSAR pairs	Satellite's name	Reference image (Capture time)	Interferometric image (Capture time)	Perpendicular baseline (m)	Space of time (days)
I1	COSMO-SkyMed	20080409	20080425	639.0458	16
I2	COSMO-SkyMed	20080409	20080628	497.6697	80
I3	COSMO-SkyMed	20080425	20080628	-141.3761	64
I4	ENVISAT	20080418	20080627	-113.3749	70

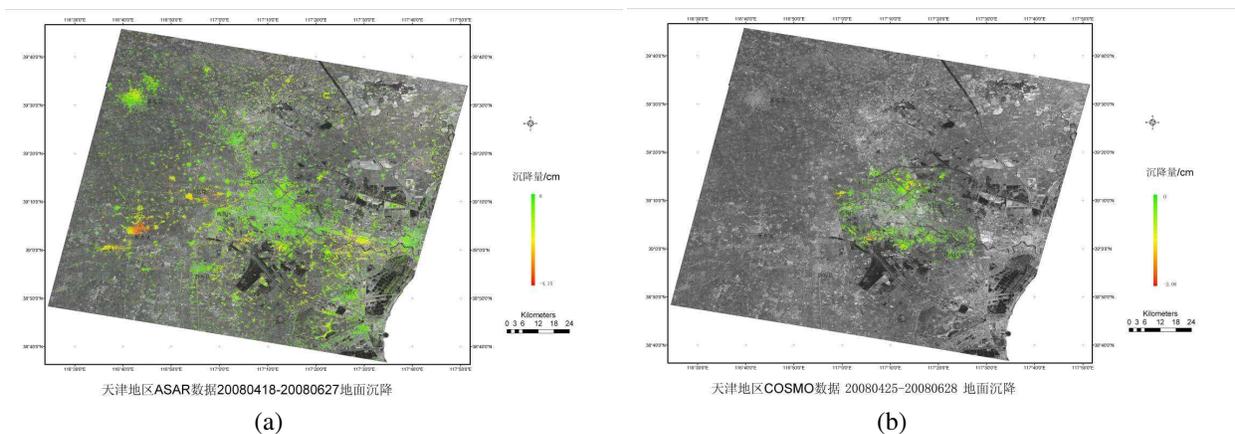


Figure 2: The results of ASAR and COSMO data on the whole area of ASAR image. (a) The land subsidence field from ASAR data. (b) The land subsidence field from COSMO data.

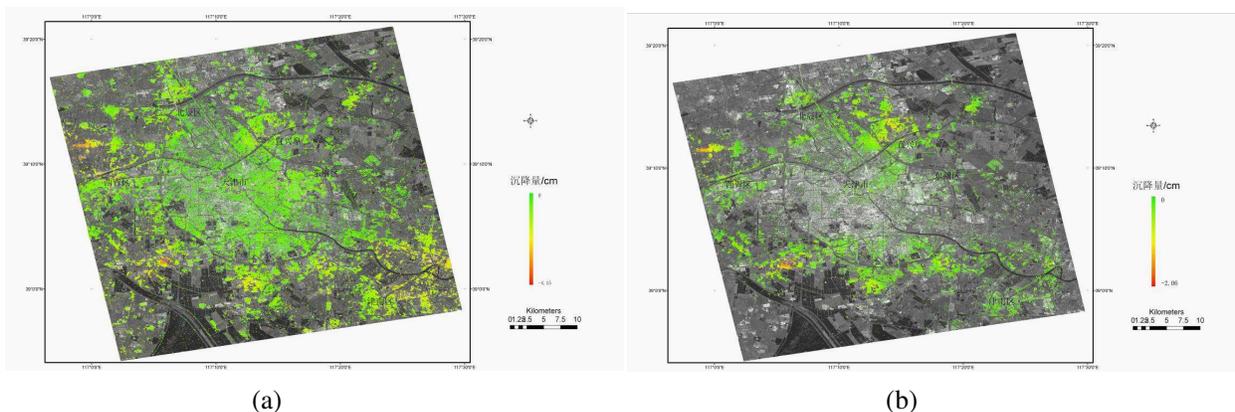


Figure 3: The results of ASAR and COSMO data on the whole area of COSMO image. (a) The land subsidence field from ASAR data. (b) The land subsidence field from COSMO data.

ASAR and COSMO data is obvious. From Fig. 3, it can be found that the similar deformation trend can be identified based on the two kinds of SAR data.

In the area shown in Fig. 3, the statistics are compared between the DInSAR results of ASAR and COSMO data (see Table 2). It should be pointed out that the ASAR and COSMO have been to the same resolution as soon as the statistics are calculated. From Table 2, the deformation values from ASAR pairs (20080418-20080627) range from 0 cm to -4.15 cm, while those of COSMO pairs (20080425-20080628) vary from 0 cm to -2.06 cm. According to previous works [7, 8], the 2 sets of deformation values are both higher than the ground data and the COSMO results are closer to ground truth.

Each of the two deformation results comes from just one differential interferogram, the DEM error, atmospheric disturbance and thermal noise can't be effectively reduced. As for the orbit error, though the complex data processing algorithms could remove most of the residual phases, it

Table 2: The deformation statistics comparison between ASAR and COSMO in the same area.

Statistics	Min (cm)	Max (cm)	Mean (cm)	Stdev (cm)
COSMO	-2.0571	0.0000	-0.0219	0.1101
ASAR	-4.1548	0.0000	-0.0648	0.3012
difference value	-3.7298	1.4384	-0.0033	0.0618

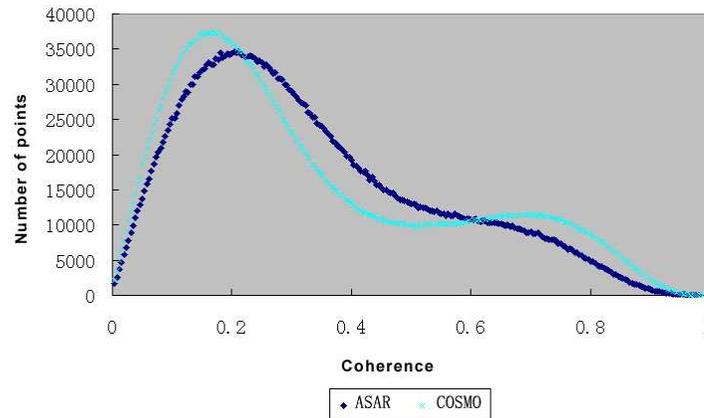


Figure 4: The coherence statistic comparison between ASAR and COSMO in the same area.

may at the same time lead to new errors caused by the mismatches between the algorithm model and the real orbit situation. If a large volume of SAR data, for example, more than 20 SAR data are used, precise deformation result can be retrieve by some PS-like technique [9, 10].

To analyze the characteristic of the coherence of COSMO high resolution data, the authors calculated the coherence map of InSAR pairs I3 and I4 on the same area and with same resolution. From Fig. 4, it is obvious that in the range of high coherence (0.62~1) and low coherence (0~0.22), the pixel number of COSMO pairs is larger than those of ASAR pairs. In the value scope from 0.22 to 0.6, there are more pixels in ASAR data than in COSMO data.

Both COSMO and ASAR deformation values are calculated from pixels with coherence higher than 0.4. In Fig. 2 and Fig. 3, it's shown that there are more areas in ASAR results with coherence from 0.4 to 1 than in COSMO data. But from Fig. 4 we can safely conclude that there are more high coherent areas in COSMO data with coherence higher than 0.62. These areas are mainly corresponding to artificial buildings. Maybe these high coherent areas in COSMO data account for the COSMO deformation values closer to ground truth.

4. CONCLUSIONS

From the results of this paper, preliminary conclusions could be made as below.

(1) Given proper temporal baseline, spatial baseline and artificial targets, a couple of COSMO InSAR pairs is more suitable to extract high-precision ground deformation information than a couple of ASAR InSAR images.

(2) The high-resolution COSMO image covers a smaller area compared with an ASAR image mode image. So, high resolution COSMO data are suitable for deformation monitoring in a relatively small and important area, while ASAR data are fit for regional survey.

(3) Although the deformation extreme values obtained from the single interference data pairs are higher than ground data, the deformation trend in Tianjin urban area is preliminary validated by two kinds of SAR data and two-pass method. As for more accurate deformation results, time series SAR data are required.

ACKNOWLEDGMENT

This work was supported by the Ministry of Land and Resources (Project No.: 200811053). The COSMO images are supplied by Beijing Earth Observation Inc. And the authors wish to thank Ge Daqing, Hu Qiong, Zhang Ling and Yan Dapeng for their kind helps.

REFERENCES

1. Krebs, G. D., COSMO 1, 2, 3, 4, http://space.skyrocket.de/index_frame.htm?http://space.skyrocket.de/doc_sdat/cosmo-skymed-1.htm. Sep. 27, 2009.
2. Maranesi, M., COSMO-SkyMed.Company Confidential, Sep. 25–27, 2007.
3. Fan, J.-H., “Research on coherent targets based dinsar technique and its applications to surface deformation monitoring,” Institute of Remote Sensing Applications, Chinese Academy of Sciences, Beijing, China, Jan. 2008.
4. GAMMA Remote Sensing AG, Interferometric SAR Processing, Gamma Remote Sensing, Nov. 2007.
5. Kampes, B., R. Hanssen, and P. Zbigniew, “Radar interferometry with public domain tools,” *Proceedings of Fringe*, Frascati, Italy, Nov. 1–5, 2003.
6. Hanssen, R. F., *Radar Interferometry — Data Interpretation and Error Analysis*, Kluwer Academic Publishers, New York, 2002.
7. Ren, L.-S., L.-R. Huang, J. Zhou, et al., “Results of settlement monitoring by GPS in Tianjin during the last ten years,” *Science of Surveying and Mapping*, Vol. 31, No. 4, 17–19, 2006.
8. Jiang, Y.-X., J.-T. Yang, K.-G. Dong, et al., “Precision of subsidence measurement with GPS technique,” *Science of Surveying and Mapping*, Vol. 31, No. 5, 63–65, 2006.
9. Ferretti, A., C. Prati, and F. Rocca, “Nonlinear subsidence rate estimation using permanent scatterers in differential SAR interferometry,” *IEEE Trans. Geosci. Remote Sens.*, Vol. 38, 2202–2212, 2000.
10. Berardino, P., F. Casu, G. Fornaro, et al., “Small baseline DIFSAR techniques for earth surface deformation analysis,” *Fringe03 Workshop*, Frascati, Italy, 2003.

Inversion of Vegetation Parameters Based on Polarimetric SAR Interferometry

Linxi Zhang^{1,2}, Jie Ren², Xingzhao Liu², and Chufeng Hu^{1,2}

¹National Key Laboratory of UAV Specialty Technique, Northwestern Polytechnical University
Xi'an, Shaanxi 710072, China

²Electronic Information Department, Northwestern Polytechnical University
Xi'an, Shaanxi 710072, China

Abstract— Polarimetric SAR interferometry (PolInSAR) is a well known technique that can be applied to inversion of vegetation parameters. In this paper, a two-layer coherent scattering model composed of a random volume over the ground is established firstly to describe the vegetation parameters, and then the polarization and interference are unified by defining a complex interferometric coherence which can be gained through the model. Finally the measured SAR data is substituted into the expression of the complex interferometric coherence to come true the inversion of vegetation parameters. In the process of the inversion of vegetation parameters, three-stage method is used. By employing the method, ground topography can be received accurately and other vegetation parameters such as vegetation height and mean extinction coefficient also can be estimated. However, the random model is based on the assumption that volume interferometric coherence is independent of polarization. Under normal circumstances, the propagation of the electromagnetic waves is anisotropic because the extinction coefficients are highly polarization dependent, namely, that volume interferometric coherence is dependent of polarization, so a new vegetation model composed of an oriented volume over the ground needs to be created. For the inversion of the oriented model, three-stage method is still used and the inversion procedure is similar with the random model. The difference between them is the interferometric coherence of random is a line inside the unit circle of the complex coherence plane, while oriented is a triangular region. The simulation results demonstrated that an accurate inversion of the vegetation parameters is still possible.

1. INTRODUCTION

PolInSAR is a new advanced technique recently based on measurement techniques of polarimetric SAR (PolSAR) and interferometric SAR (InSAR), including the characteristics that PolSAR is sensitive to the shape and direction of the scattering volume and InSAR is sensitive to the spatial distribution and height of it. Therefore, PolInSAR has greater superiority and higher accuracy than other PolSAR or InSAR on retrieving the vertical structure parameters of the vegetation. In this paper, a two-layer coherent scattering random model proposed by Treuhaft [1] is employed and three-stage method [2] is used for the inversion of vegetation parameters. Moreover, the relationship between the polarimetry and extinction coefficient is taken into account and an oriented model [3, 4] is also employed for parameter inversion, the simulation results demonstrate the effectiveness of three-stage method.

2. VEGETATION MODEL

2.1. Two-layer Random Volume Model

Two-layer random volume model is first derived by Treuhaft, and then expanded to full polarization interference [1]. The full polarization interference scattering characteristics of forest can be represented by a two-layer coherent scattering model (see Fig. 1). Where, vegetation layer can be regarded as a random volume and the propagation of the electromagnetic waves does not depend on polarization, the polarization is affected mostly by the ground scattering echo.

In this case, the complex interferometric coherence has the form shown in Equation (1) [1]:

$$\gamma(\omega) = e^{j\phi_0} \left[\gamma_v + \frac{\mu(\omega)}{1 + \mu(\omega)} (1 - \gamma_v) \right] \quad (1)$$

where, Φ_0 is topographic phase, γ_v represents volume scattering coherence, $\mu(\omega)$ is the ground-to-

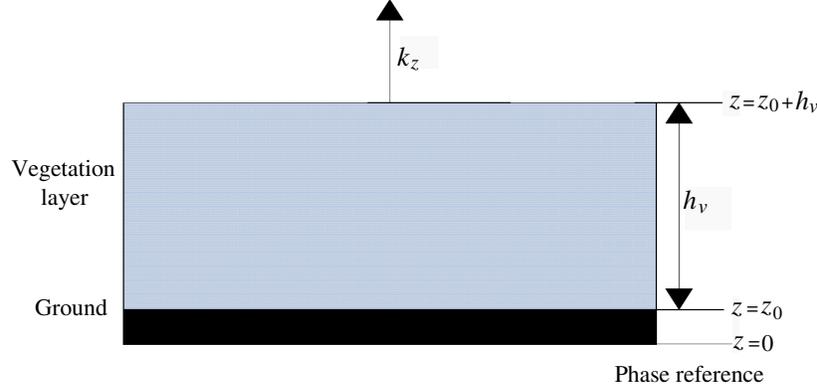


Figure 1: A two-layer coherent scattering model.

volume scattering ratio, and parameters γ_v and $\mu(\omega)$ are defined in Equation (2).

$$\gamma_v = \frac{2\sigma \left[e^{(jk_z + \frac{2\sigma}{\cos\theta})h_v} - 1 \right]}{\cos\theta \left(jk_z + \frac{2\sigma}{\cos\theta} \right) \left(e^{\frac{2\sigma h_v}{\cos\theta}} - 1 \right)}, \quad \mu(\omega) = \frac{2\sigma}{\cos\theta \left(e^{\frac{2\sigma h_v}{\cos\theta}} - 1 \right)} \frac{\omega^{*T} T_g \omega}{\omega^{*T} T_v \omega} \quad (2)$$

where, σ is the extinction coefficient, θ is the mean angle of incidence, k_z represents the vertical wave number, h_v is tree height, T_g and T_v [5] are coherency matrix associated to the contributions of the ground and volume, respectively.

2.2. Two-layer Oriented Volume Model

In contrast with the random volume case, the propagation of the electromagnetic waves is anisotropic in oriented model for the extinction coefficients are highly polarization dependent ($\sigma = \sigma_\omega$), which means volume interferometric coherence γ_v is also a function of polarization. In the model, the scattering distribution of ground and volume is changed with the polarization state. So the volume scattering coherence and ground scattering coherence are both polarization dependent, and the coherence can be written as shown in Equation (3) [6]:

$$\gamma = e^{j\phi_0} \left[\gamma_v(\sigma_\omega) + \frac{\mu(\sigma_\omega, \omega)}{1 + \mu(\sigma_\omega, \omega)} (1 - \gamma_v(\sigma_\omega)) \right] \quad (3)$$

where,

$$\gamma_v(\sigma_\omega) = \frac{2\sigma_\omega \left[e^{(jk_z + \frac{2\sigma_\omega}{\cos\theta})h_v} - 1 \right]}{\cos\theta \left(jk_z + \frac{2\sigma_\omega}{\cos\theta} \right) \left(e^{\frac{2\sigma_\omega h_v}{\cos\theta}} - 1 \right)}, \quad \mu(\sigma_\omega, \omega) = \frac{2\sigma_\omega}{\cos\theta \left(e^{\frac{2\sigma_\omega h_v}{\cos\theta}} - 1 \right)} \frac{\omega^{*T} T_g \omega}{\omega^{*T} T_v \omega} \quad (4)$$

3. SIMULATION PARAMETERS AND THREE-STAGE INVERSION

3.1. Simulation Parameter

Table 1: Parameters of random volume.

Parameters	values
frequency	$f = 1.3$ GHz
the average incidence angle	$\theta = 30^\circ$
vertical wave number	$k_z = 0.2$

Table 2: Parameters of oriented volume.

Parameters	values
frequency	$f = 4.5$ GHz
the average incidence angle	$\theta = 30^\circ$
vertical wave number	$k_z = 1.16$

3.2. Parameter Inversion of Random Volume

The inversion parameters are h_v , Φ_0 , σ , then a three-stage method is used for parameter inversion.

i) Least-square line fit

By the Equation (1), we can see that complex interferometric coherence should be a straight line in the complex plane. If the polarization interferometric simulation datas γ_{HH} , γ_{VV} , γ_{HV} ,

γ_{HH+VV} , γ_{HH-VV} are known on the assumption that $z_0 = 0$, the straight line can be fitted by using the real and imaginary parts of the datas.

ii) Estimation of Φ_0 and γ_v

The best-fit straight line has two intersections (see Fig. 2) with the unit circle, we need to determine the point which represents the topographic phase. In general, γ_{HV} is far from $\exp(j\Phi_0)$, so the coordinate of true phase point is at (0.997, -0.02719), then $\Phi_0 = -0.0273$ rad.

Curves gained by fixing the parameter σ and changing the parameter h_v have different intersections with the best-fit straight line. Usually, cross-polarization interference coherence γ_{HV} can be used to instead of γ_V approximatively [2], so choose $\gamma_V = -0.26 + 0.6904i$.

iii) Estimation of h_v and σ

According to the known information, the relationship among coherence, h_v and σ can be drawn as Fig. 3. Refer this figure, we can determine that, $h_v = 14.6$ m, $\sigma = 0.056$ dB/m.

3.3. Parameter Inversion of Oriented Volume

In contrast with the random volume case, the expression in Equation (3) is not a line. Simulations reveal that when $\sigma = 0 \sim 2$ dB/m [6], possible coherences are enclosed in a feasible region that is basically a triangle, as shown in Fig. 4.

The borderline of the possible region is caused by the differences between the extinction coefficients associated with different polarizations, as demonstrated by the separation between the extreme cases (0 and 2 dB/m). The simulation data is employed to fit a straight line, as shown in Fig. 5. The line has two intersections with the unit circle, we also need to determine which point represents the topographic phase. In general, γ_{HH-VV} is closed to the topographic phase and

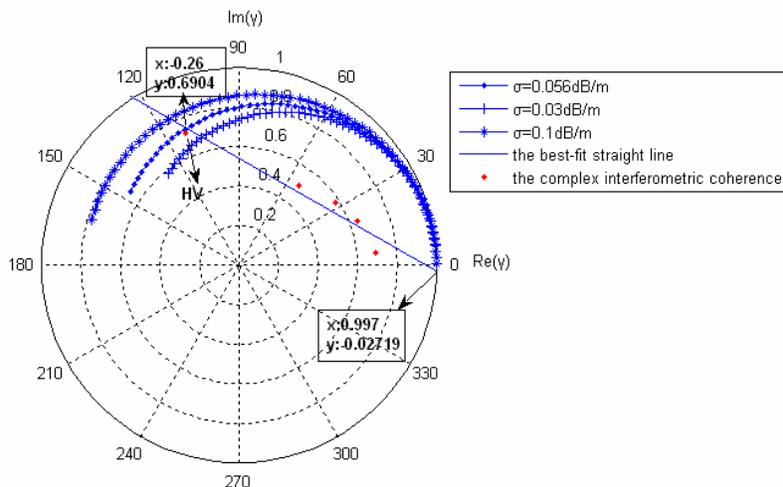


Figure 2: The reference geometry for estimation of γ_v and Φ_0 .

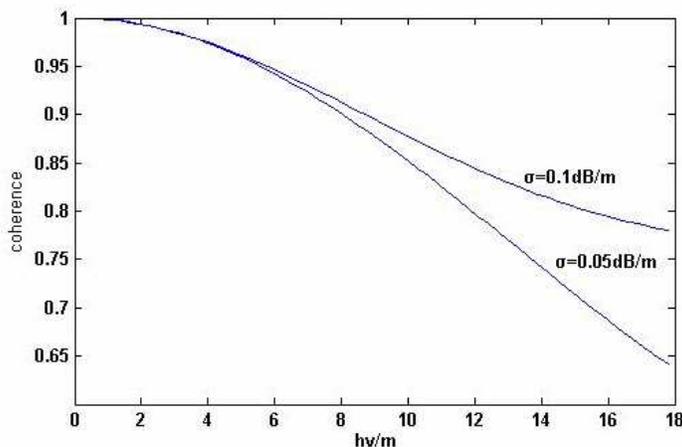


Figure 3: Coherence variation against h_v and σ .

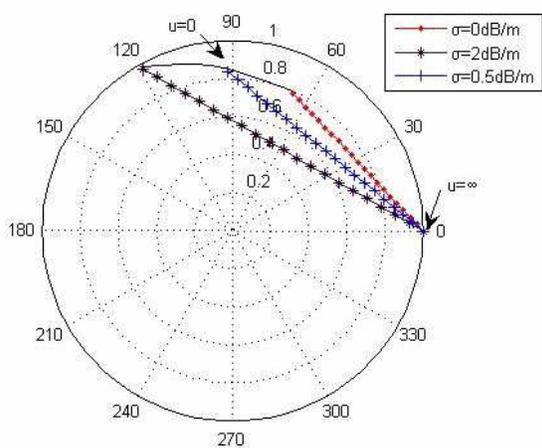


Figure 4: Possible region of the coherences in the complex.

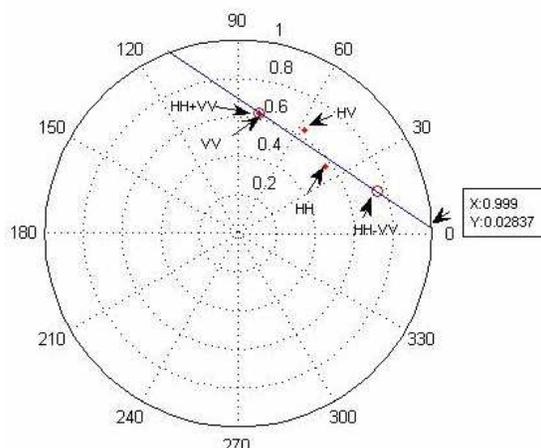


Figure 5: The best-fit line in the complex plane.

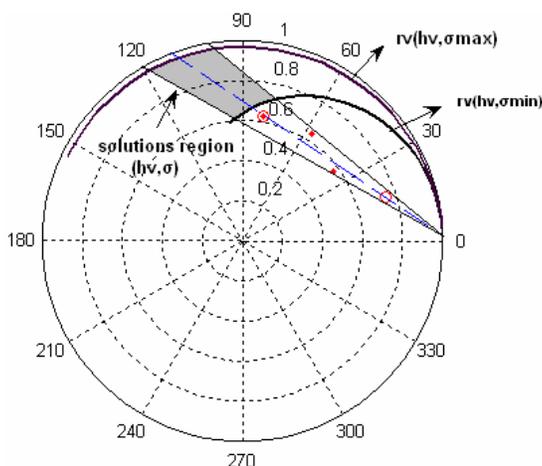


Figure 6: Parameters inversion for oriented volume.

γ_{HH+VV} is far from it [6], so the coordinate of the topographic phase point is at (0.999, 0.02837), then $\Phi_0 = 0.0284$ rad.

Two borderlines containing all the data can be drawn from the topographic phase point. From σ_{min} to σ_{max} , the curves of γ_V and the two borderlines form a shadow region (see Fig. 6) and all the possible values about γ_V corresponding to a unique parameter h_v are contained in it. The center of the shadow region is selected as the estimate of parameter γ_V , then $\gamma_V = -0.135 + 0.7911i$. According to the priori information, parameter h_v can be determined at 2.2 m, but parameter σ is difficult to accurately estimate [6].

3.4. Comparison of Inversion Parameters

Table 3: Random volume inversion parameters.

True values	Inversion values
$h_v = 15.1$ m	$h_v = 14.6$ m
$\sigma = 0.05$ dB/m	$\sigma = 0.056$ dB/m
$\Phi_0 = 0$ rad	$\Phi_0 = -0.0273$ rad

Table 4: Oriented volume inversion parameters.

True values	Inversion values
$h_v = 2$ m	$h_v = 2.2$ m
$\Phi_0 = 0$ rad	$\Phi_0 = 0.0284$ rad

4. CONCLUSION

Based on the relationship between polarization and extinction coefficient, two kinds of coherent scattering model are proposed and three-stage method is used for inversion of vegetation parameters

in the paper. The simulation results show that the error of inversion of vegetation parameters can be controlled in 10%.

REFERENCES

1. Treuhaft, R. N., "Vertical structure of vegetated land surfaces from interferometric and polarimetric radar," *Radio Sciences*, Vol. 35, No. 1, 141–177, 2000.
2. Cloude, S. R. and K. P. Papathanassiou, "Three-stage inversion process for polarimetric SAR interferometry," *IEE Proc. Radar Sonar Navig.*, Vol. 150, No. 3, 125–134, 2003.
3. Treuhaft, R. N. and S. R. Cloude, "The structure of oriented vegetation from polarimetric interferometry," *IEEE Trans. Geosci. Remote Sens.*, Vol. 37, No. 5, 2620–2624, 1999.
4. Cloude, S. R., K. P. Papathanassiou, and W. M. Boerner, "The remote sensing of oriented volume scattering using polarimetric radar interferometry," *Proceedings of International Symposium on Antennas and Propagation, (ISAP 2000)*, 549–552, Fukuoka, Japan, Aug. 2000.
5. Papathanassiou, K. P. and S. R. Cloude, "Single-baseline polarimetric SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, Vol. 39, No. 11, 2352–2363, 2001.
6. Ballester-Berman, J. D., J.-M. Lopez-Sanchez, and J. Fortuny-Guasch, "Retrieval of height and topography of corn fields by polarimetric SAR interferometry," *IEEE International Geoscience and Remote Sensing Symposium, IGARSS'04, Proceedings*, 1228–1231, 2004.
7. Ulaby, F. T., A. Tavakoil, and T. B. A. Senior, "Microwave propagation constant for a vegetation canopy with vertical stalks," *IEEE Trans. Geosci. Remote Sens.*, Vol. 25, No. 6, 714–725, 1987.

Detection of Interfaces between Frozen and Melted Sediment Using GPR: A Case Examination on Qinghai-Tibet Railway

Zhen-Wei Guo^{1,2}, Jian-Xin Liu^{1,2}, Jian-Ping Xiao^{1,2}, and Xiao-Zhong Tong^{1,2}

¹The Institute of Info-physics and Geometrics Engineer, Central South University
Changsha, Hunan 410083, China

²Key Laboratory of Non-ferrous Resources and Geological Hazard Detection
Changsha, Hunan 410083, China

Abstract— To survey the frozen earth roadbed of Qinghai-Tibet railway and to treat the hidden trouble of the roadbed are the important approach to steady the roadbed. And it is essential guaranteed to be open to traffic in security. According to the geological data of the typical testing section of Qinghai-Tibet railway, the field experiments are accomplished. Based on the processing results, a significant technology for practical application of ground penetrating radar in detecting frozen roadbed is useful and indispensable.

1. INTRODUCTION

Railway is nowadays considered as the most ecological and economical alternative to cars and planes; on the other hand, in order to be really competitive, trains have to go faster and faster on safe tracks that allows a high velocity. Permafrost is a unique problem with which Qinghai-Tibet railway is confronted. One fifth of the global land masses are underlain by permafrost [3]. Annual thawing and freezing of ground underlain by permafrost has major implications for the ground properties [3,6]. With the future scenarios of climate change, the dynamic response of permafrost to climate change is a major concern. Subgrade in permafrost region varies with the temperatureshift.

Serious difficulties of the railway construction, operation and maintenance are brought about by the special conditions of the Qinghai-Tibet railway. In order to guarantee safe operation of the Qinghai-Tibet railway, ballast sacks, contaminations, moist areas, packed layers and permafrost must be located and monitored continuously.

Thus it is important to understand how these ice masses and the frozen ground might respond to the predicted climatic forcing and feedback mechanisms such as: Changes in albedo, thermal and mechanical properties and melting of ice and changes in water path ways. Therefore, there is a need to be able to detect and map such ice and permafrost features, both in space and over time, to better understand their dynamics and behavior [5].

Ground Penetrating Radar (GPR) is a remote sensing method that has been used extensively in investigations for inspection of railway tracks [4] and has also been used as a geophysical method in permafrost research to map subsurface structures and composition (Moorman et al., 2003 and references therein). Frozen ground is, in general, GPR “friendly” with low energy dissipation and a good potential for large depth penetration and the possibility of mapping the subsurface geometries and internal structures [1, 2]. In recent years the potential of using GPR surveys on other planetary bodies, such as Mars, has increased awareness and research on the performance of GPR over frozen ground [1]. Nevertheless, the success of GPR studies on frozen grounds, such as sediments and tills, is sometimes limited due to difficulties in data interpretation and information extraction.

In this paper, several investigations are described and discussed, and a reliable and effective way is presented GPR equipment especially designed for inspecting and verifying railroad tracks and permafrost sediments.

This project is supported by National Key Project of Scientific and Technical Supporting Programs Funded by Ministry of Science & Technology of China (NO.: 2006BAC07B00).

2. GEOLOGICAL SETTING

These marlstones are of alluvial origin, their thickness ranges from 2.1 m to more than 5.5 m and they are covered, at a regional scale, by important sandstones. The sandstones appear as layers of finely sorted sandstones of colors ranging from yellow brown to brown. Petrographically speaking the sandstones is composed of quartz grains marginally associated with feldspars. Above the layers there are conglomerate and sandstone joints of sliced stones and that can reach 30 centimeters

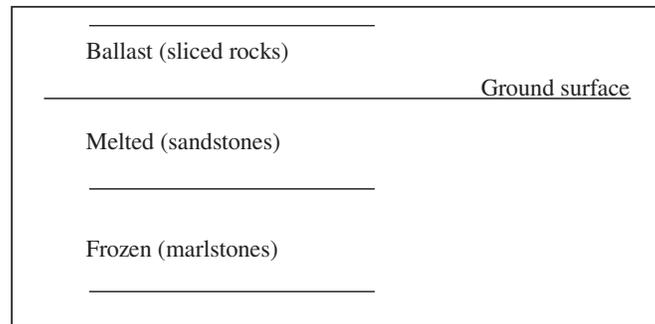


Figure 1: Geological model.

in thickness. Some marlstones layers show abundant frozen sedimentary traits of cross layered stratification in which the volume has a 40% ~ 90% ice content.

3. GPR DATA COLLECTION AND PROCESSING

We used a GR-III Ground Penetrating Radar with different antennas to map the buried frozen and melted sediment. Both systems are impulse radars. The GR-III unit was operated with the shielded 100, 200 and 400 MHz antennas.

The GPR profiles were run in a common offset mode. The center frequency of the detection used is 100 MHz, 200 MHz and 400 MHz, for the 1024 sampling points, sampling rate of 1024 ~ 1706 MHz, which is more 10 ~ 17 times than the antenna center frequency. Data were collected with a 5 ns range and 1024 samples per scan. When the depth is less than 15 m, the permittivity of the layer is between 6 and 13. Thus, one wave velocity in typical permafrost layer is limited between 0.08 and 0.15 m/ns. Therefore, this time window is chosen as 60 ~ 300 ns.

The antennas were towed by people (to keep external radio noise low), with antennas mounted in plastic boxes to avoid interference with nearby metal objects. The different antennas were driven along the same tracks, with common start and stop points. The start and end positions were determined with a GPS code receiver.

The GPR data were, after the survey, directly low- and high-pass filtered, and post-processing consisted only of 1) block average smoothing 2) down-sampling the traces to 1024 samples using linear interpolation (to avoid on-screen display aliasing) 3) two fold stacking and 4) applying a linear time-varying gain. Background signal removal to remove constant instrumental noise, antenna ringing and to suppress the direct wave, has been performed by subtracting a weighted part of the average trace (calculated using all traces) from the profile (hereafter referred to as “background removal”). The weight has been set lower in profiles containing long and flat reflections to avoid subtracting valuable information.

The GR-III data required more processing. It has been low- and high-pass filtered using a zero-phase filter, both in the forward and backward direction. The data also contained different DC (low frequency voltage offset) trends both in the individual traces and along the profiles. These DC shifts have been removed by calculating a trimmed mean within the data (the mean calculation to avoid weighting of the data by extremes), and subtracting it from the data. In addition to linear time-varying gain, an adaptive gain control has also been used for display purposes to additionally enhance weak reflections that are still above the noise floor. This gain control calculates the standard deviation of the data and the gain is a weighted function of the standard deviation. The GPR profiles have not been elevation corrected because the area is flat compared to the GPR profile lengths and penetration depth.

4. PERMAFROST-GPR ELUCIDATION

All geologic materials exhibit lower electrical conductivities frozen than thawed due to the absence of liquid water, implying a good potential use of GPR as a remote sensing instrument to map frozen ground properties. In highly (none frozen) conductive media the electromagnetic (EM) waves emitted by the GPR will be absorbed, leading to shallow penetrating depths. In the literature there are a number of different definitions of permafrost but a common one is ground with a temperature below 0°C for at least two consecutive years [3]. This definition is not always practical from a GPR perspective, due to the fact that water freezes in the ground at a range of temperatures at or below

0°C, depending on the soil properties (Hansson, 1988 and references therein). It should therefore be kept in mind that what in general terms is called permafrost can contain a significant degree of unfrozen soil with free liquid water. This can lead to substantially higher conductivity and thus larger loss of the EM waves and large differences in permittivity contrasts, despite the fact that the ground temperature is below 0°C and is defined as permafrost.

5. METHODS

Calculation of EM wave velocity and depth

The GPR reflections were verified where possible with shallow auger holes, sediment cores, ice cores and water depth measurements. From these data, both the calculated radar depth and the radar velocity were determined. To convert two-way travel time to depth, we developed a multi-layer model. The latter increases the accuracy of depth estimations when radar waves travel through multiple layers, each with a different velocity. The multi-layer depth calculation was derived from:

$$D_T = \sum_{i=1}^n \frac{t_i(v_i)}{2} \quad (1)$$

where D_T is the total depth (m), i is the interface number, v_i is average propagation velocity (m ns⁻¹) from interface ($i - 1$) to interface (i), and t_i is the two-way travel time (ns) from interface ($i - 1$) to interface (i). The difference in the GPR signal response was determined by the reflection coefficient. Using the dielectric constants, an estimation of the reflection coefficient for an interface under vertical incidence can be determined from:

$$R = \frac{\sqrt{\varepsilon_1} - \sqrt{\varepsilon_2}}{\sqrt{\varepsilon_1} + \sqrt{\varepsilon_2}} \quad (2)$$

where R is the reflection coefficient, ε_1 is the permittivity of the overlying material and ε_2 is the permittivity of the underlying material (A-CUBED, 1983). The dielectric constants for ice ($\varepsilon = 3$), water ($k = 80$) and frozen sediment ($k = 7.5$) are calculated using:

$$\varepsilon = (c/v)^2 \quad (3)$$

where c is the velocity of the EM wave in free space (m ns⁻¹) and v is the measured wave velocity through the low loss material (m ns⁻¹).

6. RESULTS

On the ground

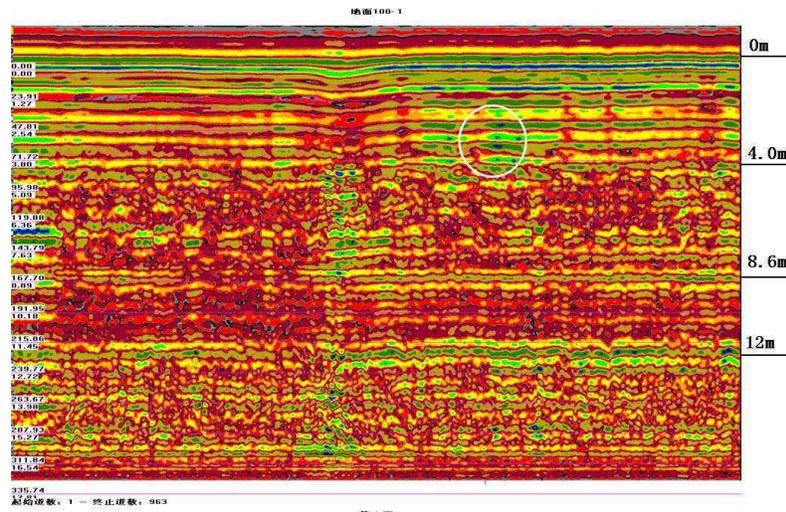


Figure 2: 100 MHz data set.

In the geological radar detection range down, the measurement regions are parts of the backfill areas, from the main formation include (Fig. 2): The first layer: 0 ~ 4 m, the second layer: 4 ~ 8.6 m, the third layer: 8.6 ~ 12 m.

For more information on the shallow strata is detected by 200 MHz and 400 MHz antenna, its initial results are as follows (Fig. 3, Fig. 4).

An interface is buried at the depth of 0.7 m, which is arranged between 0.66 m and 1.1 m. The sand covering is judged affirmatively.

The thick layer, with a preliminary judge for the relatively small amount of ice freeze-thaw zone, continuously varies between 0.7 m and 1.3 m in depth.

The layer from 1.3 m to 2.58 m in depth, with a thickness of ups and downs, initially judged to contain more ice than the freeze-thaw zone.

The permafrost sediments are believed in the depth of 2.58 meters to 4 meters.

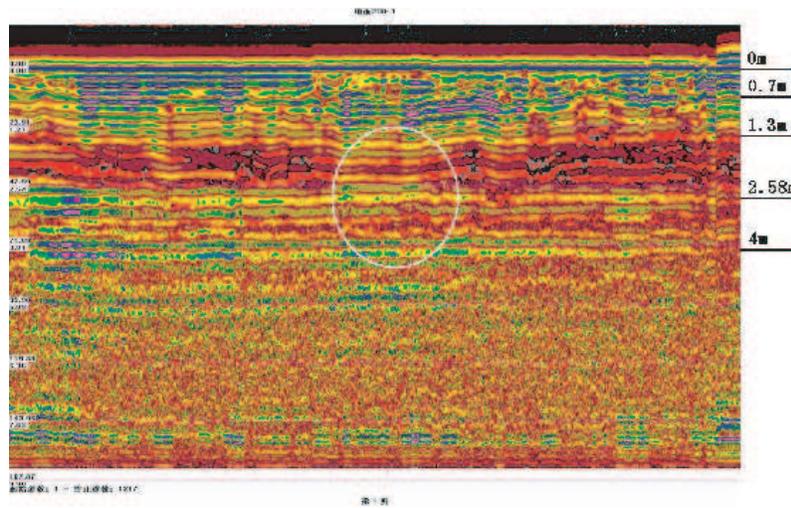


Figure 3: 200 MHz data set.

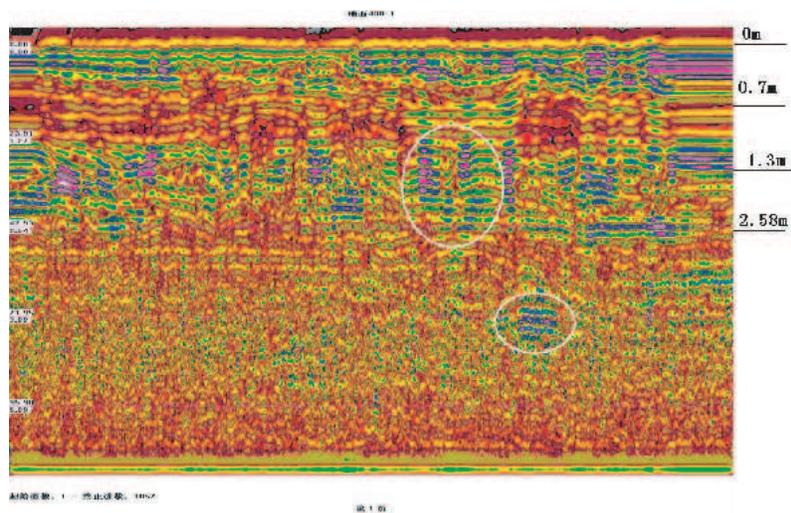


Figure 4: 400 MHz data set.

7. CONCLUSION

The results show that basic multi-frequency GPR techniques can be valuable for detection and mapping of buried ice as well as for revealing structures and the make-up of permafrost sediments. Information collected by the dedicated GPR system are valuable for preventing severe hazard to railway traffic and planning renovation; furthermore the possibility of optimizing the location of drill spots and surveying without closing the tracks, makes this method of investigation appreciable.

ACKNOWLEDGMENT

I feel indebted to my supervisor professor Jian-Xin Liu for funding and instructing. Special thanks to Jian-Ping Xiao and Xiao-Zhong Tong for discussing the results and correcting on the English.

REFERENCES

1. Arcone, S. A., M. L. Prentice, and A. J. Delaney, "Stratigraphic profiling with ground-penetrating radar in permafrost: A review of possible analogs for Mars," *Journal of Geophysical Research*, Vol. 107, No. E11, 18-1-18-4, 2002.
2. Daniels, D. J., *Ground Penetrating Radar*, 2nd Edition, IEE Radar, Sonar and Navigation Series 15 (Ed.), The Institution of Electrical Engineers, London, ISBN 0-86341-360-9, 2004.
3. French, H. M., *The Periglacial Environment*, 2nd Edition, Addison-Wesley-Longman, Reading, Mass., ISBN 0-582-30536-5, 1996.
4. Gobel, C., R. Hellmann, and H. Petzhold, "Georadar model and in-situ investigations for inspection of railway tracks," *Proceedings of Ground Penetrating Radar Conference*, Kitchener, Canada, June 12–16, 1994.
5. Kääh, A., "Remote sensing of mountain glaciers and permafrost creep," Geographisches Institut der Universität Zürich, Stiftung Zentralstelle der Studentenschaft der Universität Zürich, Vol. 48, ISBN 3-85543-244-9, 2005.
6. Williams, P. J. and M. W. Smith, *The Frozen Earth — Fundamentals of Geocryology, Studies in Polar Research*, Cambridge University Press, ISBN 0-521-36534-1, 1989.

GPR Data Processing for Permafrost Detection in Qinghai-Tibet Railway

Zhen-Wei Guo^{1,2}, Jian-Xin Liu^{1,2}, Jian-Ping Xiao^{1,2},
Xiao-Zhong Tong^{1,2}, Wei Zhang^{1,2}, and Jie Li¹

¹The Institute of Info-physics and Geometrics Engineering, Central South University
Changsha 410083, China

²Key Laboratory of Non-ferrous Resources and Geological Hazard Detection
Changsha, Hunan 410083, China

Abstract— As an advanced technology, Ground-Penetrating Radar (GPR) is a geophysical method based on electromagnetic wave propagation commonly used for nondestructive subsurface imaging, including thruway, glacier, permafrost and pipe. It has also been used as a geophysical method in permafrost research to map subsurface structures and composition. The Qinghai-Tibet Plateau is the world's highest plateau, averaging over 4000 meters above sea level and covering an area of 200,000 square kilometers. The Qinghai-Tibet railway has always been puzzled by permafrost. It is imperative to detect permafrost subgrade fast.

Permafrost recognition algorithm was proposed with applying the rolling power spectrum. In the Central South University, we did a test. As a result, it has shown that anomalies have been significantly strengthened. The most important points of this study, related to methodology and processing, are discussed, and finally, a method is proposed using data processing for permafrost detection.

1. INTRODUCTION

Ground-Penetrating Radar (GPR) is a geophysical method based on electromagnetic wave propagation commonly used for nondestructive subsurface imaging [1, 2]. It was originally used for ground geological investigation and the identification of major discontinuities within rock beds [4]. It has also been used as a geophysical method in permafrost research to map subsurface structures and composition [3]. However, most of these studies used low frequency antennae (100 to 400 MHz) in order to detect permafrost between one to twenty meters depth.

Despite the fact that computer processors are becoming more efficient and GPR software packages are becoming more user friendly, the processing and interpretation of the GPR data from roads, railways and airports is still the most time consuming phase and an interpreter's skills play a key role in the success of a GPR project. GPR data processing can be divided into four phases: (1) preprocessing, (2) data processing, (3) interpretation and visualization and (4) reporting. The processing was divided into several classic steps namely amplitude compensation, filtering, migration, 2D powering for coherent filtering, and display. Permafrost recognition is a significantly indispensable part of data processing for permafrost detection.

In this paper, we will describe the environment where the GPR detects. After presentation of the field-survey characteristics, we will describe the processing applied to the GPR data; explain how the permafrost is discriminated well. The most important points of this study, related to methodology and processing, are discussed, and finally, a method is proposed using data processing for permafrost detection.

2. GEOLOGICAL SETTING

In Qinghai-Tibet Plateau, the natural conditions are harsh, the air is oxygen-poor and the climate is bitterly cold. Most of the land consists of mountains, wilderness, and permafrost and snow zones. The centre of the study site (35°29'52"N, 93°41'59"E) is at the northeast of Hohxil. The study site is within a region of permanent permafrost. The mean annual air temperature in Qumar is -4°C . The study area, which is subjected to the test section in the Qinghai-Tibet railway, located above Qumar high plains. The height of the ballast is 6.3 m, and the width of pavement is 8.3 m, while the berm width around both sides is 3.0 m. Conglomerate and sandstone are filled about 30 cm under the base of ballast, at the same time, 4% of the cross-slope, along both sides of the filling-based centers, is necessary. Sliced stones and rocks, which have a thickness of 1.5 m, are filled above the conglomerate and sandstone. At the top of the sliced rock chip to build a layer of gravel in a round cushion and a thickness of 30 cm, which is a layer of impermeable layer on the laying of sand-gravel cushion, while the ballast is filling with coarse soil.

3. FIELD SURVEY, DATA PROCESSING AND INTERPRETATION

3.1. Field Survey in Qumar, 2008

A GR-III Ground Penetrating Radar detector is used in this study with the shielded 100, 200 and 400 MHz antennas. Data were collected with a 5 ns range and 1024 samples per scan. Measurements can be made as continuous sections or on a point, according to the permafrost of the studied area or the presence of thermometer hole. The antennae were slid on the train at a constant speed 15 km/h in order to obtain 15.386 traces per meter (about 0.10752 m between each trace).

Before any investigation of the stability of painted interface frozen sediment, tests were made to confirm that the permafrost recognition algorithm can be used on the data processing for this non-destructive geophysical method for the detection of permafrost sediment. The detection of permafrost depends on the difference between the dielectric impedance of the frozen and the melted domains. When the depth is less than 15 m, the permittivity of the layer is between 6 and 13. Thus, one wave velocity in typical permafrost layer is limited between 0.08 and 0.15 m/ns.

3.2. Data Processing

Permafrost recognition algorithm

3.2.1. Log Power Spectrum

For a given signal $x(t)$, whose spectral density $P(\omega)$ is non-negative and even symmetry, the power spectrum $|P(\omega)|^2$ highlight the main component of that $|P(\omega)| > 1$, which suppress those the minor of that $|P(\omega)| < 1$. Thus it is benefited for the power spectrum to highlight and identify the main features of the spectrum. When several features of the power spectrum needs to be identified, the log-power spectrum can be used, which is defined as:

$$L_x(\omega) = |\lg P_x(\omega)|^2 \quad (1)$$

The GPR signal received was a result of sub-waves and the formation reflection coefficient convolution result, so the power spectrum of the radar data is the stratigraphic sequence of the power spectrum and the reflection coefficient of the radar wavelet power spectrum multiplied by the results of operations. Using on the operations of logarithm, the information of spectrum can be translated into the superposition of stratum and the wavelet power spectrum. It is convenient for us to observe that the stratum make the transform of the frequency response in the radar wavelet.

3.2.2. Time Window Selection

To study the radar cross-section of road-based diseases, we hope the road only to be the spectral data of the basis, so the time window is necessary on the time profile of the radar in order to analyze the localized features of the spectrum which the time signal on the basis of partial section is corresponding to.

The function which has the window features satisfied the following qualification

$$\omega \widehat{G}(\omega), tg(t) \in L^2(R) \quad (2)$$

where $g(t)$ is the time window function, $\widehat{G}(\omega)$ is the result of Fourier transform of $g(t)$.

Many functions can be used as the time window function, such as Rectangular function, Yamagata functions and m -splines and so on. However, the selection of window function must be based on the “uncertainty principle” in order to achieve time and frequency domain localization of the highest resolution. “Uncertainty principle” is given: For the window function $g(t)$, its window size to meet the following formula.

$$4\Delta_{g(t)}\Delta_{\widehat{G}(\omega)} \geq 2 \quad (3)$$

Only when $g(t)$ is Gaussian function can the equation be defined. Gaussian window is the best window for partial analysis. Gaussian window function is defined as:

$$g_a(x) = \frac{1}{2\sqrt{\pi a}} e^{-\frac{x^2}{4a}}, \quad a > 0 \quad (4)$$

3.2.3. Construction of Rolling Spectral Profile

As the ground-penetrating radar to obtain the real signal, while the real value of the signal spectrum is symmetric, we take a positive semi axis part of the frequency analysis. The beginning of time window, as the radar reflectivity profile, there is a ground-penetrating radar as an arbitrary mapping relationship:

$$x(t_0) \Rightarrow W(t_0 + \Delta T) \Rightarrow \left| \lg \left\{ \frac{1}{N} P(n\Delta\omega) \right\}_{n=0}^{N-1} \right| \Rightarrow Q(t_0) \tag{5}$$

where, $x(t_0)$ is that the time profile corresponds to t_0 time information; $W(t_0 + \Delta T)$ in the selected time window between $t_0 \sim t_0 + \Delta T$; $\left| \lg \left\{ \frac{1}{N} P(n\Delta\omega) \right\}_{n=0}^{N-1} \right|$ is the discrete spectrum which the time profile is in the $t_0 \sim t_0 + \Delta T$ time window; $\Delta\omega$ is the frequency interval; $Q(t_0)$ is the mean spectrum of discrete spectral values which corresponds to:

$$Q(t_0) = \sum_{n=0}^N \left(\left| \lg \left\{ \frac{1}{N} P(n\Delta\omega) \right\} \right| * n\Delta\omega \right) \tag{6}$$

Be expressed on the time profile of the radar reflection time t_0 the energy conversion for a given time window of the rolling average spectra of $Q(t_0)$.

If ΔT is as the time interval from the starting point along the time-depth, scroll down to the depth at each time point corresponds to the mean spectrum of type (6), which form a rolling cross-section is as follows:

$$G(m) = \left(\sum_{t=0}^{T_m} Q(t) \right) (m) \quad m = 1 \text{ — Sample Points} \tag{7}$$

where: T_M is the sampling time window.

3.2.4. The Improvement of Scroll Spectrum Algorithm

There are some problems as follows, if we adopt the Equation (7) to calculate the rolling spectral profile:

- a) Computer is running slow. Radar detection is a fast method to detect. If when the data interpretation is processed, using complicated time algorithm to obtain results, the effectiveness of radar detection will be hugely impacted.
- b) It is vulnerable to external interference in the signal mutation. Ground Penetrating Radar detects by the use of emission and reflection of electromagnetic waves. It is susceptible to outside interference, so how to reduce the impact of interfering signals, it is significantly irreplaceable to the ultimately results in the data interpretation.

For these reasons, scroll spectrum algorithm is improved.

First, it is necessary to give the frequency of rolling scan multiple N , by controlling the size of the N values to achieve fast computation purposes. If the original sampling points were 512, 512 times the mean spectrum need to be computed for each of trace; if the N takes 8, only 64 times the mean spectrum need be computed, and computation time is increased eight-fold. Type (7) by rolling multiple becomes:

$$G(m) = \left(\sum_{t=i*N*\Delta T}^{t \leq T_m} Q(t + N * i) \right) (m) \tag{8}$$

Secondly, by the Equation (8) the number of secondary sample is resumed.

$$\tilde{G}(m) = \sum_{t=0}^{T_m} G(m) \frac{\sin\left[\left(t - m\right) \frac{\pi}{T_s}\right]}{\left(t - m\right) * \frac{\pi}{T_s}} \quad m = 1 \sim \text{Sample Points} \tag{9}$$

Finally, the results of the second sample will be filtered by the sliding windows open two-dimensional filtering.

$$\widehat{G}(m) = \sum_{n=-N}^N \sum_{m=-M}^M \tilde{G}_{mn} \tag{10}$$

where n and m are respectively the number of the traces and sample points.

3.2.5. Effect of Rolling Spectrum Algorithm

Figure 1 is the floor, Central South University, a school-run time after the detection of abnormal radar images of the lot due to perennial gardening, underground soil moisture larger, because the soil has a strong absorption of electromagnetic waves, SIR series and the Italian radar, are very difficult to find an exception exists. Figure 2 is the result of the adoption of recognition algorithms, unusual features to be strengthened.

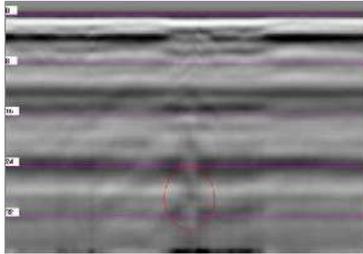


Figure 1: Time profile image.

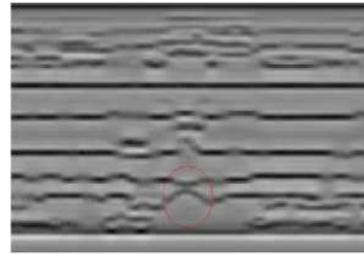


Figure 2: ARMA Log-power spectrum profile.

4. CONCLUSION

Tests and measurements made in detection show that the GPR can be used to identify the interface of water and soil. For very shallow both discontinuities and continuities, GPR reflections are masked by other waves of strong energy and often affected by coupling changes due to the roughness of the air–rock interface. The data processing proposed here, which improves the rolling spectrum algorithm enhances the speed of the interpretation. The method is effective and efficient to distinguish the interface between the frozen and the unfrozen ground. The permafrost can be distinguished easily.

REFERENCES

1. Conyers, L. B., *Ground-penetrating Radar for Archaeology*, Alta Mira Press, Walnut Creek, 224, 2004.
2. Daniels, D. J., *Ground Penetrating Radar*, 2nd Edition, Institution of Electrical Engineers, London, 726, 2004.
3. Moorman, B. J., S. D. Robinson, and M. M. Burgess, “Imaging Periglacial conditions with ground-penetrating radar,” *Permafrost and Periglacial Processes*, Vol. 14, 319–329, 2003.
4. Sass, O., “Bedrocks detection and talus thickness assessment in the European Alps using geophysical method,” *Journal of Applied Geophysics*, Vol. 62, 254–269, 2006.

GPR Polarization Simulation with 3D HO FDTD

Jing Li¹, Zhao-Fa Zeng^{1,2}, Ling Huang¹, and Fengshan Liu²

¹College of Geoexploration Science and Technology, Jilin University
Changchun 130026, China

²Applied Mathematics Research Center, Delaware State University, DE 19901, USA

Abstract— Polarization signal is important in designing GPR data measurement. Its scattering characteristics can be used to discriminate the targets. For the target with different medium properties or orientation, the GPR signal response has a significant difference in different polarization mode. The targets such as metallic, low impedance object and PVC, high impedance objects have different polarization signals while are difficult to discriminate in the conventional GPR acquiring methods. The polarimetric measurement has the potentials to improve the discrimination, reduce the clutter interference and get the better prospecting effect. Meanwhile, the simulation is one of key research issues to understand the radar polarization theoretical. In order to improve simulation accuracy, we use 3D high-order finite difference time domain (HO-FDTD) method to simulate the polarization. It obtains a good application effect. Through the simulation study of GPR polarization measurement, we can provide more accurate and richer theoretical for the practical work. provide more accurate and richer theoretical for the practical work.

1. INTRODUCTION

Ground Penetrating Radar (GPR) is a geophysical technique which use of high frequency electromagnetic wave ($10^6 \sim 10^9$ Hz) to detect electrical distribution for investigating the shallow subsurface medium [1]. The polarization characteristic signal is an important information such as amplitude, phase and Doppler frequency of radar for data interpretation. The signal from the receive antenna is a function of the polarization of the transmit antenna and properties of subsurface targets [2]. The dipole antenna of GPR radiate linearly polarized wave with the predominant polarization direction parallel to the long axis of the antenna. Receive antennas are very sensitive to the electric field component parallel to the long axis. In practical, the complexity geological environment, the distribution of underground target without regularity, and complex medium properties make the results shown obvious difference in the same condition. The polarization of electromagnetic wave is a fundamental property of propagation that provides the GPR with a unique opportunity for producing improved images of object in the subsurface. Many researches have done the studies and got some important improvement. Yong and Caldecot [3] designed the GPR system which used to locate underground pipe lines. Two orthogonal dipole antennas are used in the system. Roberts, Daniels and Radzevicius [4, 5] have done research on the polarization characteristics of GPR. They pointed out that there are significant difference between TE and TM polarization scattering. P. Capizzi and P. L. Cosentino [6] study the GPR multi-component data analysis, and so on.

In this paper, we use 3D high-order FDTD method [7] to simulate the GPR polarization measurement. Our study mainly focus on the polarization characteristic in different antenna geometric modes when the targets have different dielectric properties, shape and orientation. The results show that choosing appropriate polarization not only effectively improve the SNR and resolution but also can discriminate the medium properties, shape and orientation of targets. It could provide more effective geological information.

2. PRINCIPLE OF HIGH-ORDER FDTD

FDTD is a direct time-domain method to solve Maxwell equation which is simple, visual, flexible, and has a large dynamic range in calculation that apply in the GPR polarization simulation. However, one of the major limitations of the FDTD approach for the numerical solution of Maxwell's equations is its calculation accuracy. High-order FDTD method is one of ways to improve simulation accuracy. It uses Taylor expansion to expand the time-domain difference equation and to set a new difference equation. The time integration is implemented using second-order central finite difference and the space integration is implemented by using 2M-order central finite difference.

It can be obtained the discrete Maxwell's equations.

$$E_x^{n+1} = \frac{1}{\left(\frac{\epsilon_0 \epsilon_r}{\Delta t} + \frac{\sigma_e}{2}\right)} \left[\sum_{l=-M}^{M-1} a(l) \left(\frac{H_z^{n+\frac{1}{2}}(i, j+l+\frac{1}{2}, k)}{\Delta y} - \frac{H_y^{n+\frac{1}{2}}(i, j, k+l+\frac{1}{2})}{\Delta z} \right) + \left(\frac{\epsilon_0 \epsilon_r}{\Delta t} - \frac{\sigma_e}{2} \right) E_x^n \right] \quad (1)$$

where, $\Delta x, \Delta y, \Delta z$ as the along x, y, z direction's grid size, i, j, k is the grid number. the along x, y, z direction, $a(l)$ gives as:

$$a(l) = \frac{(-1)^l}{2 \left(l + \frac{1}{2}\right)^2} \frac{[(2M - 1)!!]^2}{(2M - 2 - 2l)!! (2M + 2l)!!} \quad (2)$$

Numerical stability condition:

$$\Delta t \leq \left(v_{\max} \sum_{l=0}^{M-1} |a(l)| \sqrt{\left(\frac{1}{\Delta x}\right)^2 + \left(\frac{1}{\Delta y}\right)^2 + \left(\frac{1}{\Delta z}\right)^2} \right)^{-1} \quad (3)$$

For the absorbing boundary condition, we use the uniaxial anisotropy perfectly matched boundary (UPML) which proposed by Sacks [9] and Gedey [8]. It does not need split the electromagnetic field. The algorithm is very concise and easy to understand. It is a non-physical absorbing medium and wave impedance which is not dependent on outward-wave angle and frequency.

3. ELECTROMAGNETIC WAVE POLARIZATION

The EM wave field strength changes with direction of EM wave. Such phenomenon can be observed: when the metal wire parallel to the electric field, electromotive force which electric field induced in the wire is strongest; when the metallic wire vertical to the electric field, electromotive force is zero. It shows that the direction of electromagnetic field in space is a very important issue. The polarization of plane wave can characterize the features that fixed electric field vector in space change with time. The conventional GPR antennas use two linear dipole antennae with identical properties which is vertical to the propagation direction of electromagnetic. The polarization modes are less to be used in which antenna parallel to the direction of electromagnetic wave propagation and cross-cutting along the direction. The GPR multi-polarization (Fig. 1) provides more information of underground target. The geometries shown as Fig. 1 is four components of acquired electromagnetic field: parallel broadside with y -directed antenna (YY), parallel broadside with x -directed antenna (XX), perpendicular antennas with y -directed source and x -directed receiver (YX) and perpendicular antennas with x -directed source and y -directed receiver orientation (XY).

The electric field traveling in the z direction can be described by two orthogonal components as given in Balanis 1989:

$$E_x(z, t) = E_{x0} e^{-\alpha z} \cos(\omega t - \beta z - \phi_x), \quad E_y(z, t) = E_{y0} e^{-\alpha z} \cos(\omega t - \beta z - \phi_y)$$

where α represents the attenuation coefficient, β represents the phase constant, ω the angular frequency, ϕ the phase, and E_{x0} and E_{y0} are the maximum amplitudes of the E_x and components respectively E_y .

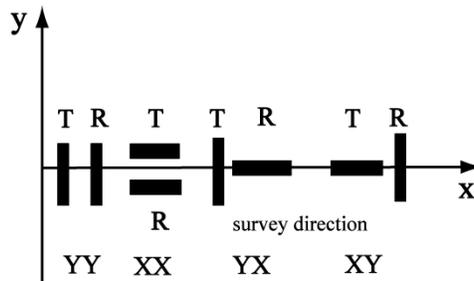


Figure 1: The polarization mode of GPR antenna.

4. GPR POLARIMETRIC SIMULATION

The steel and other metallic with low impedance, or PVC and other dielectric with high impedance are the common detecting target in the GPR detection. The result is often difficult to interpretation due to the dielectric properties of object as well as the clutter interference of the surrounding medium. The high-frequency electromagnetic waves which emissions by GPR have different polarization signal responses. 3D high-order FDTD method used to simulate the polarization characteristics of two object models: *the PVC* $\epsilon_r = 3, \sigma = 0.006 \text{ S/m}$.

We can clearly see that there are significantly differences for the two models (Fig. 3 and Fig. 4). For the PEC cylinder models, the signal from YY polarization mode could clearly discriminate the location of targets and has very small diffraction interference error. the steel cylinders in YY polarization mode and we can clearly discriminate the lower steel cylinders. The signal amplitude is stronger than other modes. But the other polarization can not achieve prospecting effect. For the PVC pipes model, the XX polarization signal response is stronger than other modes (Fig. 5). The clutter interference is also low. The signal is best imaged.

The position relation of survey direction and target orientation is an important factor to improve the GPR polarization effect. So, we design the model respectively vertical and parallel to the survey direction (Fig. 6). From the profile we can conclude that, for the metallic which parallel to the survey direction and the dielectric which vertical to the survey direction, the XX polarization has stronger signal response. For the metallic which vertical to survey direction and the dielectric which parallel to survey direction, the YY polarization signal response is stronger than the signal from other modes (Fig. 7).

The above analysis based on the target trend along a certain direction. It may be effect by the trend of target and the survey direction. Whether have the same conclusion for the equiaxed target such as sphere and cube. In order to verify this problem, we select the sphere and cube model with different dielectric properties. The map is the single-trace reflection signal in the same location of XX and YY polarization mode after removing the direct wave. From the single-trace signal, the metallic models have stronger signal amplitude in YY mode, the XX mode make the dielectric have stronger signal amplitude. It is well to prove that different dielectric properties target suitable for different polarization mode (Fig. 8 and Fig. 9).

Summing up all the simulation and analysis, we can get that for the equiaxed target, the metallic and low impedance dielectric objects, are best imaged with the long axis of the dipole antennas

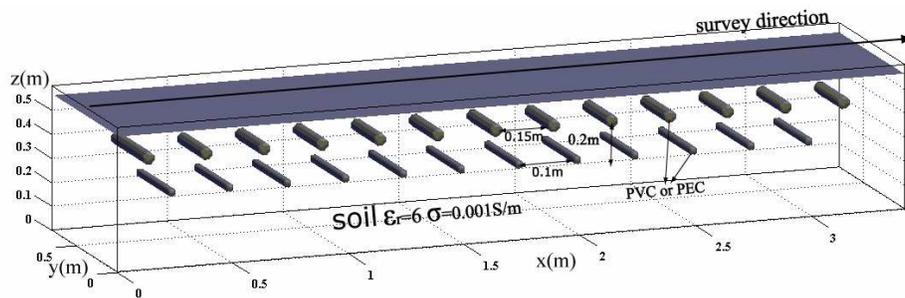


Figure 2: PEC cylinder or PVC pipe model.

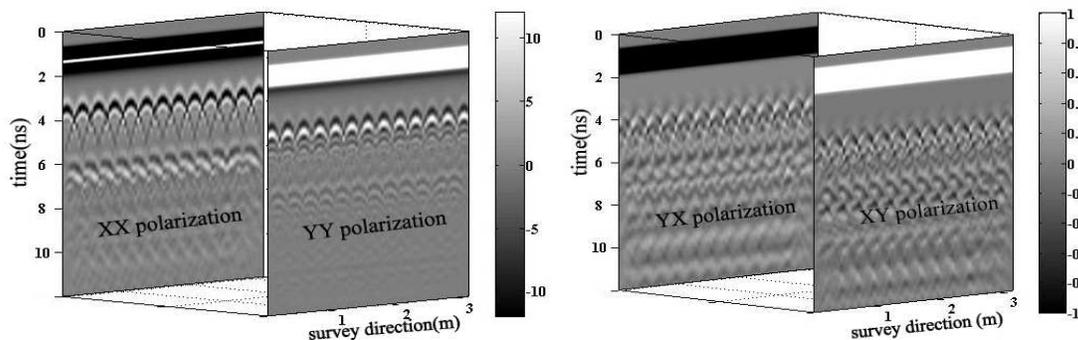


Figure 3: The GPR slice of double PEC cylinder model with different polarization.

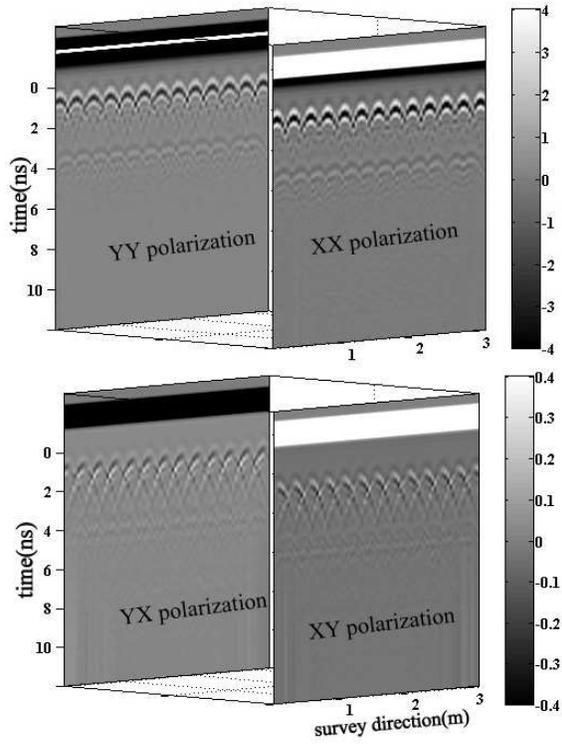


Figure 4: The GPR slice of double PVC model.

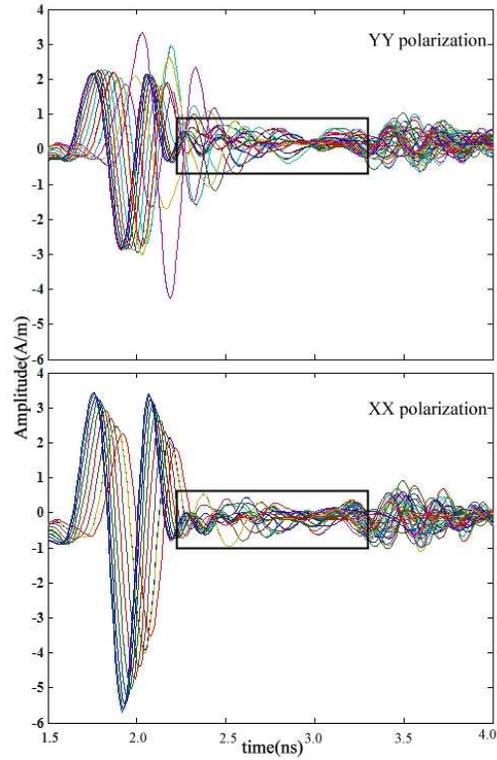


Figure 5: The GPR signal response of double PVC.

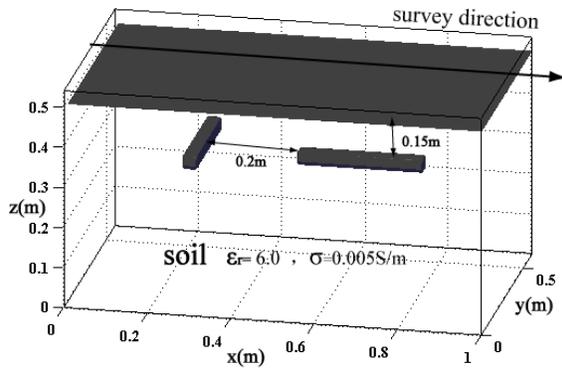


Figure 6: Cylinder model in different trend.

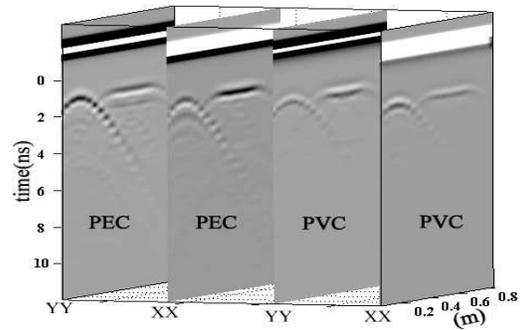


Figure 7: The GPR slices of different model in different polarization.

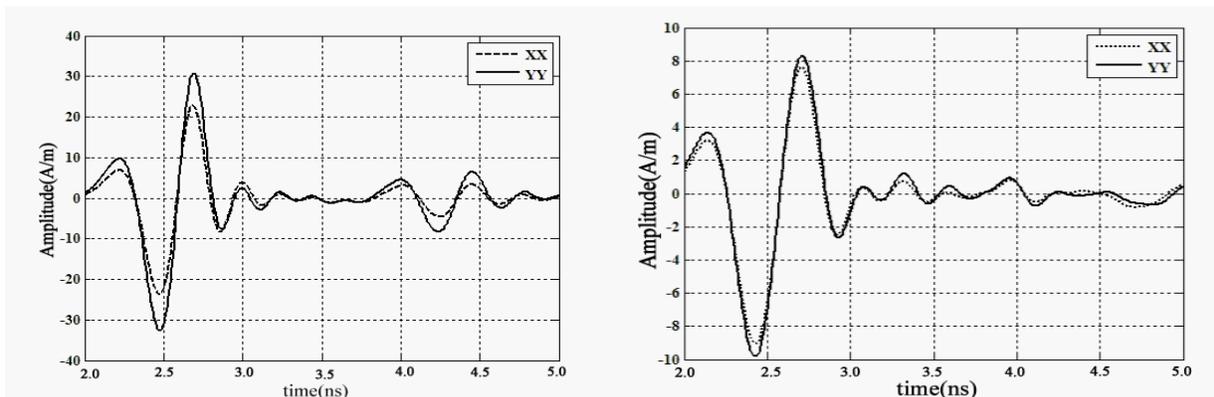


Figure 8: The single-trace signal of metallic cube and sphere model (left map is cube, right is sphere).

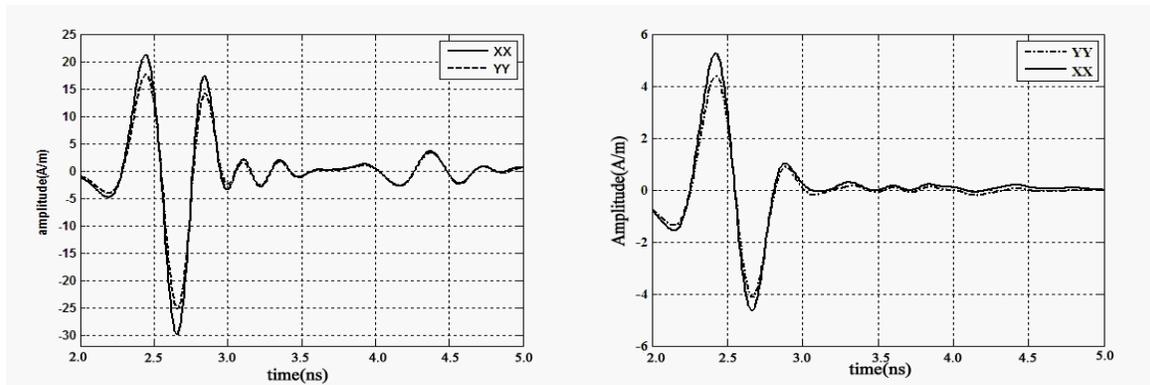


Figure 9: The single-trace signal of dielectric cube and sphere model (left map is cube, right is sphere).

oriented parallel to the long axis (YY mode). The dielectric, and high impedance dielectric objects, are best imaged with the long axis of the dipoles oriented orthogonal to the long axis (XX mode); If the object have a certain extension direction, the metallic object which is parallel to the survey direction and the dielectric which is vertical to the survey direction are best imaged with the XX polarization mode; the metallic which is vertical to the survey direction and the dielectric which parallel to the survey direction are best with the YY polarization mode. For the XY and YX polarization modes can not have a good prospecting effect in any case.

5. CONCLUSIONS

Through the study of GPR polarization measurement simulation with high-order FDTD method, we have an overall understanding to the GPR signal response for the target with different dielectric properties and shape in different polarization modes. By choosing appropriate GPR polarization mode, we can significantly improve the SNR, the recognition ability and the resolution. Moreover, the GPR polarization measurement also plays a prominent role in detecting 3D complex target. In the case of complex targets, the effect of polarization is even more critical to ultimately obtaining a good 3D view of individual objects within a group of complex features.

ACKNOWLEDGMENT

The research work was supported by NSFC (40774055) of China, and a DoD DEPSCoR W911NF-07-1-0422 of USA.

REFERENCES

1. Annan, A. P., "Radio interferometry depth sounding: Part 1. Theoretical discussion," *Geophysics*, Vol. 38, No. 3, 557–580, 1973.
2. Radzevicius, S. J. and J. J. Daniels, "Ground penetrating radar polarization and scattering from cylinders," *Journal of Applied Geophysics*, Vol. 45, 111–125, 2000.
3. Yong, J. and T. R. Caldecot, "A portable detector for plastic pipe and other underground objects, Final Report 404X-1," *Columbus*, The Ohio State University Electro Science Laboratory, 1973.
4. Daniels, D. J., D. J. Gunton, and T. H. F. Scot, "Introduction to subsurface radar," *IEE Proceedings*, Vol. 133, No. 4, 278–326, 1988.
5. Roberts, R. L. and J. J. Daniels, "Analysis of GPR polarization phenomena," *Journal of Environmental and Engineering Geophysics*, Vol. 1, No. 2, 139–157, 1996.
6. Capizzi, P. and P. L. Cosentino, "GPR multi-component data analysis," *Near Surface Geophysics*, 87–95, 2008.
7. Georgakopoulos, S. V., C. R. Birtcher, et al., "Higher-order finite difference schemes for electromagnetic radiation scattering and penetration," *IEEE Antenna's and Propagation Magazine*, Vol. 44, 134–142, 2002.
8. Gedney, S. D., "An anisotropic perfectly matched layer-absorbing medium for the truncation of FDTD lattices," *IEEE Trans. Antennas Propagat.*, Vol. 44, No. 12, 1630–1639, 1996.
9. Sacks, Z. S., D. M. Kingsland, and J. F. Lee, "A perfectly matched anisotropic absorber for use as an absorbing boundary condition," *IEEE Trans. Antennas Propagat.*, Vol. 43, No. 12, 1460–1463, 1995.

Fine Exploration Based on Dense Frequency Pseudo-random Harmonic Electromagnetic Method

Weibin Luo and Qingchun Li

Department of Geology Engineering and Geomatics, Chang'an University, China

Abstract— In controlled source electromagnetic (CSEM) sounding the exploration depth is determined by frequency of source signal and the earth resistivity. In current CSEM survey systems the mono frequency square or harmonic signal was adopted at each shot. It often took a long time to complete a deep EM sounding. And the vertical resolution of electromagnetic sounding is low. For improving the vertical resolution a new method of Pseudo-random harmonic electromagnetic sounding is introduced, based on the study of 2^n series Pseudo-random multi-frequency (PRMF). The energy of PRMF concentrates on several main frequencies. It is an ideal excitation sources for electrical exploration such as Induced Polarization method and CSEM sounding. The fine vertical resolution based on dense signal frequency was analyzed. The designs of dense frequency ratio pseudo-random signal using the complex programmable logic device (CPLD) and the tool of hardware description language (VHDL) are introduced. It can produce different dense frequency waveforms. The feasibility of this method is tested and its effectiveness is discussed. It is proved that this method will be an ideal source signal for design and development of fine CSEM sounding.

1. INTRODUCTION

Electromagnetic geophysical methods are based on the study of EM field propagation in the earth and on determining the electromagnetic parameters of the earth, containing unique information about petrophysical properties, lithological characteristics, and mineral content of the rocks in the earth's interior. In CSEM sounding, multiple frequencies are needed, because the depth to which the field is significantly affected by the earth is a function of the normalized electrical distance, δ/R (Zhdanov, 2009). There are two controllable parameters in the ratio δ/R : The frequency, and the transmitter-receiver separation. If the transmitter-receiver separation is varied to obtain a sounding curve, such a sounding is called a geometric sounding. If the frequency is varied, the sounding is a parametric sounding. If both are changed, the sounding is a hybrid geometric-parametric sounding. In frequency domain CSEM exploration, both in marine and on land environment, a source signal, mostly a square wave is frequently used as the transmitter waveform to induce the earth, and the electromagnetic response of subsurface is measured to evaluate the rock distribution.

Also it can transfer maximum energy to the subsurface because the transmitter current is running at its peak amplitude at all times, but a limitation of the square wave is that most of the transmitted energy is in the first harmonic. The transmitter sweep all of the setting frequencies of the square wave, the receiver measures the electromagnetic fields response of the earth, and the Cagniard resistivity, phase, or other parameters are derived. Clearly, the work efficiency is low for the measurements follows setting frequencies one by one; the receiver and transmitter need accurate communication; the noise ratio of measurements of EM response in various frequencies are different.

Square wave signal can be decomposed with the fundamental frequency and a series of odd harmonics. So some scholars have prompted that the odd harmonic can to be observed, it is called "odd harmonics method". The EM response of fundamental frequency and some odd harmonics are measured synchronously, but it has two fatal weaknesses: Firstly, the frequency difference of adjacent frequencies is fixed, and the distribution of harmonics is not meet the requirements of fine electromagnetic sounding; secondly, harmonic amplitude decrease fast, for example, the amplitude of 5th odd harmonics is only as strong as 1/5 of fundamental frequency. In other words, each harmonic component is observed at different current intensity.

In CSEM surveys both in marine and in land multiple frequency measurement can give us more details and help us achieving a good evaluation of economic abnormalities [6]. So multiple frequency waveforms were adopted in CSEM survey such as pseudorandom binary sequence (PRBS) used in multi-transient electromagnetic (MTEM) [3]. In this paper, we design a dense frequency pseudo-random signal for fine electromagnetic sounding based on the definition of vertical resolution.

2. THE VERTICAL RESOLUTION OF ELECTROMAGNETIC SOUNDING

According to skin depth, we define the vertical resolution as the ratio of minimum thickness of the geological body can be distinguished and its depth:

$$R_V = \frac{d}{H} \times 100\% \quad (1)$$

where d is the minimum thickness of a geological body can be distinguished, H is the depth. Even without considering the equivalent of the phenomenon of electromagnetic sounding, vertical resolution is also a very complex concept, it not only with the electromagnetic wave frequency, earth resistivity, but also with the resistivity difference of geological body and the rock, space between transmitter and receiver, instrument sensitivity, observation error and other factors. In this paper we only study the electromagnetic wave frequency on the impact of vertical resolution.

If we take Bostick depth as the exploration depth, then the vertical resolution can be defined as:

$$D(f) = 356\sqrt{\frac{\rho}{f}} \quad (2)$$

$$R_V = \frac{D(f_1) - D(f_2)}{D(f_1)} \times 100\% = \left(1 - \sqrt{\frac{f_1}{f_2}}\right) \times 100\% \quad (3)$$

where D is Bostick depth, ρ is the resistivity and f is the frequency with $f_2 > f_1$. Formula (3) shows that the vertical resolution is inversely proportional to the ratio of the frequency sampling interval.

3. 2^n SERIES OF PSEUDO-RANDOM WAVEFORM

2^n series of pseudo-random signal is random in a period but can be generated periodically [J. S. He, 1988], The pseudo-random n -frequency waves $p(2, n, t)$ with cycle $[0, T)$ can be defined as:

$$p(2, n, t) = \begin{cases} A & 0 \leq t < \frac{l_1 T}{2^n} \\ -A & \frac{l_1 T}{2^n} \leq t < \frac{l_2 T}{2^n} \\ \dots & \dots \\ A & \frac{l_{m-2} T}{2^n} \leq t < \frac{l_{m-1} T}{2^n} \\ -A & \frac{l_{m-1} T}{2^n} \leq t < \frac{l_m T}{2^n} \end{cases} \quad (4)$$

where, m is the number of switching times per period, l_k ($k = 1, 2, \dots, m$) is an integer, A is the amplitude.

As can be seen from the curve in Figure 1, it has seven frequencies with significant amplitude values at 1 Hz, 2 Hz, 4 Hz, 8 Hz, 16 Hz, 32 Hz and 64 Hz, these main spectrum in Figure 1(b) are called main frequency of pseudo-random signal. The ratio of high frequency and its next low frequency is two. The main frequency of the amplitude spectrum is basically the same. The amplitude of waveform in time is one, and the period is 0.1 s.

The 2^n series of pseudo-random signal have other main feature: Firstly, the initial phase between the main frequency is simple, it is $-\pi/2$; secondly, the pseudo-random signal energy are concentrated in the main frequency; thirdly, with n increased, the power of the dominant frequency decline gradually and the energy in non-primary frequency will increase; Fourthly, it can be generated periodic. So, in an actual application, we are only need to select an appropriate n and the multiple frequencies with enough power can be generated easily.

Clearly, the pseudo-random electromagnetic signal is a broadband signal and the frequency ratio between two adjacent main frequency is two, so the vertical resolution of 2^n series of pseudo-random electromagnetic sounding is also low, not less than 29.3%.

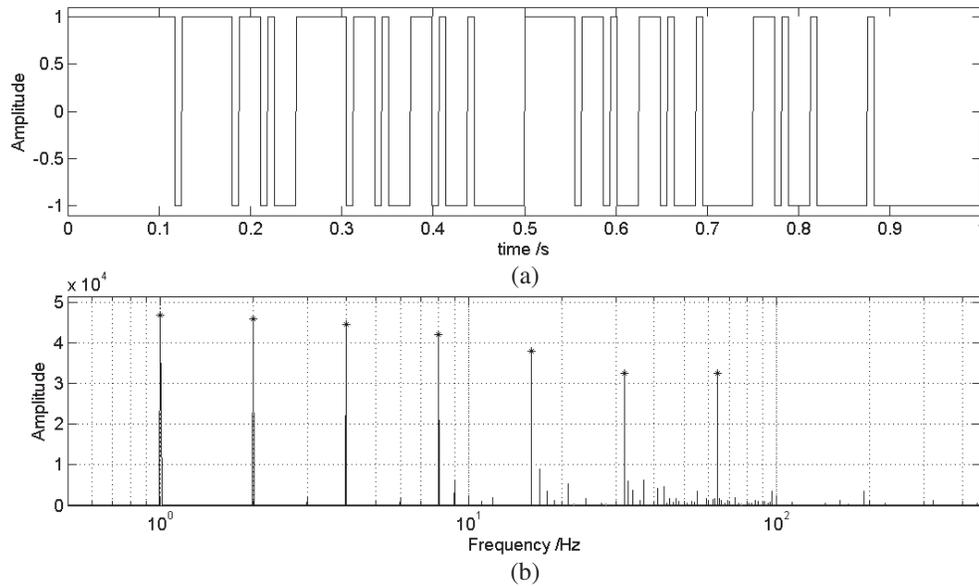


Figure 1: The base frequency of 1 Hz of 2^n series of pseudo-random waveforms (a) and its seven-frequency wave power spectrum (b) ($A = 1$, $T = 0.1$ s).

4. FINE EXPLORATION WITH DENSE FREQUENCY

From formula (3), we can see that reducing the difference of frequency is an effective method to improve the vertical resolution. In fact, since 2^n series of pseudo-random signal has fixed n -master frequency, for a given n pseudo-random signal the switching times per period is fixed, so it can be easily generated. Design a finite state machine (FSM) according to the switching times per period of pseudo-random signal, and stimulate the FSM with different clock frequency, we can get n pseudo-random signal with different n main frequency. This design idea can also be used to generate invert-repeated m -sequence and can be easily achieved using programmable logic device (CPLD).

According to a period of pseudo-random 5-frequency waves in time-domain expression, as shown in Figure 2, we build digital logic binary sequence of FSM code to generate the pseudo-random 5-frequency signal, using one stand for ON, and zero stand for OFF. We complete FSM code using hardware description language (VHDL) and implemented with CPLD. And we use a program counter to get 131072 Hz, 155344.59 Hz, 182361.04 Hz and 220752.84 Hz of 4 clock frequency division from crystal of 4194304 Hz. Using these four frequency clock to drive the FSM, we can get four pseudo-random 5-frequency waves with different 5 main frequencies. Figure 3 shows the fundamental frequency separately at 32 Hz, 38 Hz, 45 Hz, 54 Hz pseudo-random five main frequency distribution, together there are 20 main frequency by four times excite, and the ratio between the adjacent main frequency is less than two, thus it is equivalent to as add three frequencies 38 Hz, 45 Hz and 54 Hz between 32 Hz and 64 Hz. Similarly we can get more dense frequencies, together with higher power harmonic frequencies.

The four group Pseudo-random 5 frequency signal generated by four times clock stimulation lists in Table 1, together with harmonic frequency which power is larger more than 10% of base

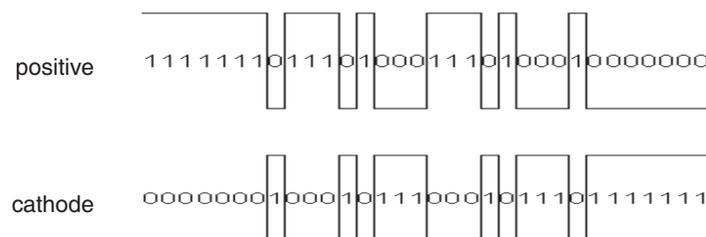


Figure 2: A pseudo-random 5-frequency waveform and state machine codes.

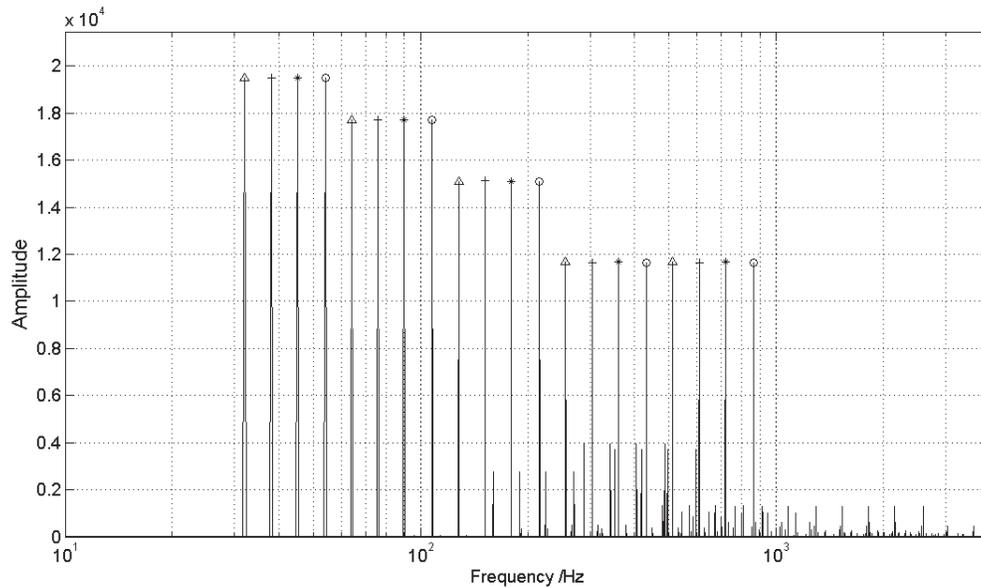


Figure 3: Four groups of five main frequency dense combination with fundamental frequency were 32 Hz, 38 Hz, 45 Hz, 54 Hz respectively.

Table 1: Frequency distribution of four groups of pseudo-random 5-frequency signal including the main frequency and harmonic with bigger power (Hz).

No. 1 group	No. 2 group	No. 3 group	No. 4 group
32	38	45	54
64	76	90	108
128	152	180	216
160	190	225	270
256	304	360	432
288	342	405	486
352	418	495	594
512	608	720	864

frequency. There are 32 frequencies between 32 and 864 Hz. With these dense frequencies the maximum vertical resolution in CSEM sounding can be up to 0.91% in theory.

Therefore, using dense frequency signals other than the pseudo-random program, stimulated just four times, we can get high-resolution electromagnetic sounding data. The base frequency and the power spectrum distribution can be controlled easily by simple select the FSM drive clock frequency. It is feasible using these dense pseudo-random main frequencies and harmonic in electromagnetic survey. Dense frequency EM response can help us obtaining high vertical resolution, and the increase in working hours and costs can be negligible.

5. CONCLUSION

For a given condition, the vertical resolution is largely depend on the signal frequency we used in electromagnetic sounding. In traditional CSEM sounding or CSAMT, the source signal is square wave, and frequency stepping is two times from low frequency to higher frequency. The vertical resolution is not less than 29.3%, it can not meet the requirements of fine exploration.

2^n series of pseudo-random signal is a complex combination of different frequency square-wave according to a certain relationship. It is wideband signal with several main frequencies and some main harmonic components. It is also a periodic signal and can be generated easily. Therefore, the 2^n series of pseudo-random signal is an ideal signal for fine CSEM and IP exploration. Dense frequency distribution can help us obtain more information of subsurface and higher vertical resolution.

For a given 2^n series of pseudo-random signal, the number switching times per period is fixed. So it is convenient using programmable logic devices and the state machine constructed in VHDL to code the waveform as a binary sequence. Just using four different frequency clock to stimulate the FSM we can get dense frequency 2^n series of pseudo-random signal. The source waveform we generated by this method including square wave, dual frequency wave, three-frequency wave, five-frequency wave, seven-frequency wave, nine-frequency waves, eleven-frequency wave and thirteen frequency wave and other more complex waveform signal.

REFERENCES

1. Kaufman, A. A. and G. V. Keller, *Frequency Domain and Time Domain Electromagnetic Sounding*, Geological Publishing House, Beijing, 1987.
2. He, J. S., *2^n Series Pseudo-random Signal and Its Application*, *Annual of the Chinese Geophysical Society*, Xi'an Map Press, Xi'an, 1998 (in Chinese).
3. Ziolkowski, A., B. A. Hobbs, and D. Wright, "Multitransient electromagnetic demonstration survey in France," *Geophysics*, Vol. 72, No. 4, F197–F209, 2007.
4. Srnka, L. J., J. J. Carazzone, M. S. Ephron, and E. A. Eriksen, "Remote reservoir resistivity mapping," *The Leading Edge*, Vol. 25, 972–975, 2006.
5. Mittet, R. and T. Schaug-Pettersen, "Shaping optimal transmitter waveforms for marine CSEM surveys," *Geophysics*, Vol. 73, No. 3, F97–F104, 2008.
6. Um, E. S. and D. L. Alumbaugh, "On the physics of the marine controlled-source electromagnetic method," *Geophysics*, Vol. 72, No. 2, WA13–WA26, 2007.

GPR Migration Imaging Algorithm Based on NUFFT

Hao Chen, Renbiao Wu, Jiaxue Liu, and Zhiyong Han

Tianjin Key Laboratory for Advanced Signal Processing, Civil Aviation University of China
Tianjin 300300, China

Abstract— Stolt migration is one of the useful GPR imaging algorithms, which has the advantage of fast implementation speed. The fast Fourier transform of integral formula in image reconstruction is realized by coordinate transformation and done through interpolation. Interpolation determines the effectiveness of imaging directly. In the practical applications, zero-fill interpolation in time-domain is used to improve accuracy, but which will increase the computational complexity. In this paper, non-uniform fast Fourier transform (NUFFT) algorithm is introduced to realize the fast computing of integral formula in image reconstruction directly. The simulation results show that this method achieve good results.

1. INTRODUCTION

Stolt migration is one of the useful GPR imaging algorithms which has the advantage of fast implementation speed. The date, which is uniformly distribute in frequency-wavenumber domain, is transformed from frequency-wavenumber domain to wavenumber-wavenumber domain. Then using interpolation method to make sure that the date is uniformly distribute in wavenumber-wavenumber domain. So it will be possibility to use the FFT, which lead efficiencies, to calculate the date. That is the key method of Stolt migration method. Usually, the Linear interpolation and sinc interpolation [2, 3] is used. Interpolation determines the effectiveness of imaging directly, in the practical applications, zero-fill interpolation in time domain is used to improve accuracy but which will increase the computational complexity.

This paper analyzed the root causes of why could not use FFT to calculate the formula for image reconstruction before stolt interpolation. Then introduced the algorithms of computing Fourier transform with unequally spaced data in any of the domain (time or frequency) have been proposed. These algorithms are known as Non-Uniform FFT (NUFFT), NUFFT was first proposed by Dutt and Rokhlin in 1993 [4], Beylkin; Nguyen [5] and Liu; Fessler [6] and Sutton [7] are made great contributions to its development.

In this paper, NUFFT is introduced to realize the fast computing of formula in image reconstruction directly. Both the stolt interpolation and the final range inverse FFT have been substituted by a single NUFFT. This paper describes the detail application of the NUFFT's to the stolt migration algorithm. Compared the new method with the old ones in calculation efficiency and image result. The performance is illustrated with numerical simulations. The simulation of B-scan data of a two-layered medium is generated by the Gprmax2.0 software. The simulation results show that this method reduce the complexity of the interpolation while ensure the operational speed. at the same time, it use least-square to approximate the function so the energy of the image result is more concentrated. This method achieves good results.

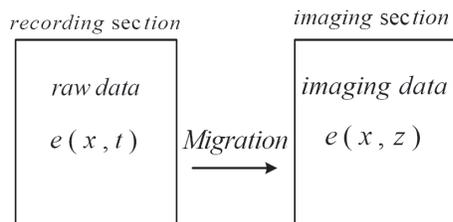


Figure 1: Migration schematic.

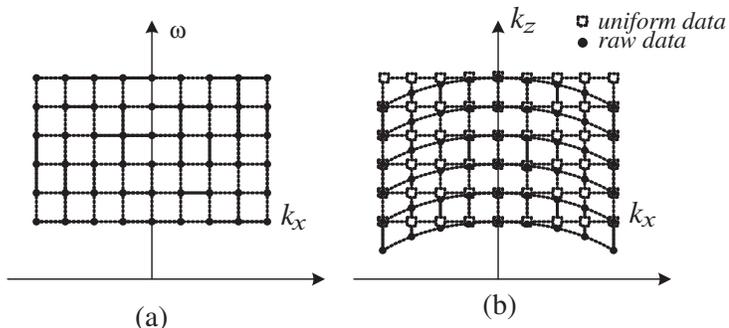


Figure 2: Coordinate transformation and stolt interpolation.

2. STOLT MIGRATION ALGORITHMS

Migration method back propagates the received wave front to the instant the target “explodes” and the image is then the scalar field at the instant before it begins to propagate. From the recording section extended to the imaging section is called migration that is show in Fig. 1.

Consider the 2-D case, the raw data $e(x, z = 0, t)$ is measured on the surface $z = 0$. The imaging data after migration is $e(x, z, t = 0)$, in particular, the 2-D Fourier transform can be written as:

$$E(k_x, z = 0, \omega) = \iint e(x, z = 0, t) e^{-i(k_x x + \omega t)} dx dt \quad (1)$$

The wave field extrapolation formula in the ω - k domain is:

$$E(k_x, z, \omega) = E(k_x, z = 0, \omega) e^{ik_z z} \quad (2)$$

The 2-D inverse Fourier transform about k_x, ω can be written as:

$$e(x, z, t) = \frac{1}{(2\pi)^2} \iint E(k_x, z, \omega) e^{i(k_x x + \omega t)} dk_x d\omega \quad (3)$$

According to Eq. (2) and make $t = 0$, the integral formula in image reconstruction is:

$$e(x, z, t = 0) = \frac{1}{(2\pi)^2} \iint E(k_x, z = 0, \omega) e^{i(k_x x + k_z z)} dk_x d\omega \quad (4)$$

We could Calculate Eq. (4) by direct summation and thus is an accurate evaluation of Eq. (4) but it takes $o(N^2)$ as many as operations. An interpolation approach is usually utilized to improve efficiency. consider the ω - k domain algorithm:

$$k_x^2 + k_z^2 = \left(\frac{\omega}{v/2} \right)^2 \quad (5)$$

It is easy to see form Eq. (5) that $k_z = \sqrt{[\omega/(v/2)]^2 - k_x^2}$, so k_z is unequally spaced, that is why Eq. (4) can not calculate though FFT. However, according to Eq. (5), Eq. (4) becomes

$$e(x, z, t = 0) = \frac{1}{(2\pi)^2} \iint E' \times G e^{ik_x x} e^{ik_z z} dk_x dz \quad (6)$$

That

$$E' = E \left(k_x, z = 0, \frac{v}{2} \sqrt{k_x^2 + k_z^2} \right)$$

$$G = \frac{v}{2\sqrt{1 + k_x^2/k_z^2}}$$

Then, though interpolation k_z becomes uniform, so that Eq. (6) can be Calculate by regular FFT, Coordinate transformation and interpolation as show in Figure 1.

3. MIGRATION BASED ON NUFFT

In the stolt migration a variable transform is made form ω to k_z and then interpolation is performed along k_z . There are many of the traditional interpolation methods, such as linear interpolation, sinc interpolation, and their improved methods [2, 14]. If the method selected inappropriate, the energy will not concentrate and will Cause the parasitic Component [2]. By zero-padding in the time domain can improve the accuracy of interpolation, thereby improving image quality, but this will increase computational complexity. This paper introduces the non-uniform fast Fourier transform (NUFFT) to achieve Eq. (4).

3.1. NUFFT Algorithms

The main idea of NUFFT is for a known point of non-uniform space and its close points in the uniform space, we can use the points' Fourier transform basis functions to approximate the point' which is in the non-uniform space.

Our aim is to develop a fast algorithm to find the following summation:

$$F(\alpha)_j = \sum_{k=-N/2}^{N/2} \alpha_k e^{i\frac{2\pi}{N}jt_k} \quad j = -\frac{N}{2}, \dots, \frac{N}{2} \quad (7)$$

in which t_k is un-uniform, for most NUFFT algorithms the procedures is:

1. Use $(q + 1)$ uniform points to interpolate one point's (t_k) Fourier transform basis function ($\exp(i2\pi jt_k/N)$)

$$s_j e^{i\frac{2\pi}{N}jt_k} = \sum_{l=-q/2}^{q/2} x_l(t_k) e^{i\frac{2\pi}{mN}j([mt_k]+l)} \quad (8)$$

where the “accuracy factors” $0 < s_j \leq 1$ are chosen to minimize the approximation error. Use least-square to approximate this sequence by a small number of uniform points.

2. Calculate Fourier coefficients

$$\eta_n = \sum_{l,k,[mt_k]+l=n} \alpha_k x_l(t_k) \quad (9)$$

3. Regular FFT: use uniform FFT to evaluate

$$F_j = s_j^{-1} \sum_{n=-mN/2}^{mN/2-1} \eta_n e^{i\frac{2\pi}{mN}nj} \quad (10)$$

The number of arithmetic operations is $O(N(q + 1) + mN \log mN)$.

3.2. Stolt Migration Using NUFFT

Is easy to see that the discrete form of Eq. (5) is

$$\left[\frac{2\pi n_\omega}{dtN_t} / (v/2) \right]^2 = \left[\frac{2\pi n_{k_x}}{dxN_x} \right]^2 + \left[\frac{2\pi n_{k_z}}{dzN_z} \right]^2 \quad (11)$$

So

$$n_{k_z} = N_z dz \sqrt{\left[\frac{n_\omega}{dtN_t} / (v/2) \right]^2 - \left[\frac{n_{k_x}}{dxN_x} \right]^2} \quad (12)$$

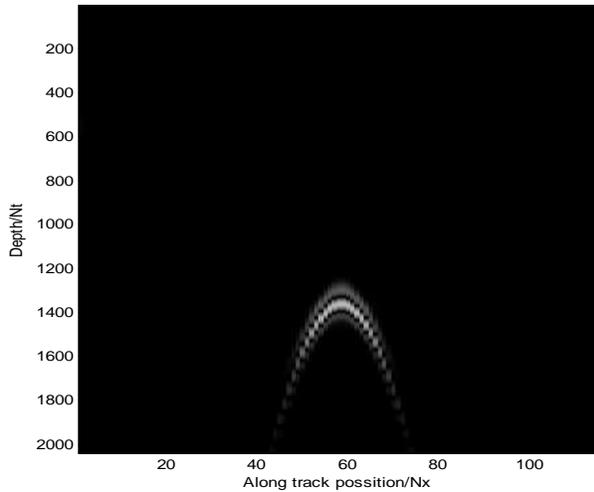


Figure 3: GPR B-scan of a two-layered medium.

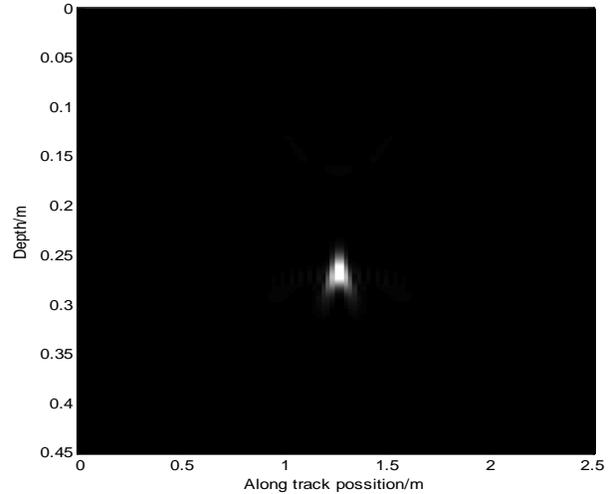


Figure 4: Result image of migration based on NUFFT.

the discrete form of Fourier transform basis function $\exp(ik_z z)$ is $\exp(i2\pi n k_z z/N)$, use the NUFFT method can realize the fast Computing.

4. EXPERIMENTAL RESULTS

In the simulation experiment. The case of a metal ball existing at a depth of 0.2525 m. The transmitter is located at 0.05 m Above the surface. The simulation of B-scan data of a two-layered medium is generated by the Gprmax2.0 software. An electromagnetic pulse of ricker form is assumed to be emitted. The simulated depth of medium air and concrete is 0.05 m and 0.45 m respectively. The relative dielectric constant of each layer is 1 and 20 respectively. The center frequency of transmit antenna is $f = 900$ MHz and sampling time is $dt = 5.89$ ps. After suppression of ground reflective wave and direct coupling wave interference, the B-Scan image is shown in Fig. 3. the migrate result with the NUFFT migration method as show in Fig. 4, where $q = 6$ and over sampling rate $m = 2$, accuracy factors $s_j = \cos \pi j / (mN)$.

For comparison, the direct-summation migration and interpolation migration methods are implemented. we use entropy to measure the effect of the migration image, the more greater the entropy, the more confusion the migration image is. For the matrix $A = (a_{ij})_{m \times n}$, define that the entropy is [10, 13]:

$$E(A) = \left[\frac{\sum_{j=1}^n \sum_{i=1}^m a_{ij}^4}{\left[\sum_{j=1}^n \sum_{i=1}^m a_{ij}^2 \right]^2} \right]^{-1} \quad (13)$$

The result is reported in Tab. 1.

Table 1: The entropy of the images.

	Direct-sum	interpolation	NUFFT
entropy	5.8645	10.3621	6.9298

5. CONCLUSIONS

Through the analysis above and simulation results we can see that, the Stolt migraton based on NUFFT is efficient while reducing the complexity of the interpolation, and has made good imaging results. NUFFT algorithm and stolt interpolation both are adopted to achieve the fast computing of Eq. (4) by approximate, but NUFFT algorithm is to approximate the Fourier transform-based by least-squares method in the frequency domain which is more normative, and the energy of the result image is more concentrated. In addition, in practical applications, due to obstructions, restricted areas and underground media, the raw data is also non-uniformity, NUFFT algorithms will be more widely used in this field.

REFERENCES

1. Gu, K., G. Wang, and J. Li, "Migration based SAR imaging for ground penetrating radar systems," *IEEE Radar Sonar and Navigation*, Vol. 151, No. 5, 317–325, Oct. 2004.
2. Zhang, C. and Z. Zhou, "A modified ground penetrating radar 3-D synthetic aperture imaging method based on stolt migration," *Signal Processing*, Vol. 20, No. 6, Dec. 2004.
3. Zeng, H., T. Zeng, C. Hu, and Y. Zhu, "Comparison of several interpolation methods in spotlight SAR imaging algorithms," *Signal Processing*, Vol. 21, No. 4, Aug. 2005.
4. Dutt, A. and V. Rokhlin, "Fast Fourier transforms for nonequispaced data," *SLAM-J. Sci. Comp.*, Vol. 14, No. 6, 1993.
5. Beylkin, G., "On the fast fourier transform of functions with singularities," *Applied and Computational Harmonic Analysis*, Vol. 2, No. 4, 363–381, 1995.
6. Liu, Q. H. and N. Nguyen, "An accurate algorithm for nonuniform fast fourier transforms (NUFFT's)," *IEEE Microwave and Gueded Wave Lett.*, Vol. 8, No. 1, 18–20, 1998.
7. Fessler, J. A. and B. P. Sutton, "Nonuniform fast Fourier transforms using min-max interpolation," *IEEE Trans. Signal Process.*, Vol. 51, No. 2, 560–574, Feb. 2003.

8. Liu, Q. H. and N. Nguyen, “The regular fourier matrices and nonuniform fast fourier transforms,” *SLAM J. Sci. Comp.*, Vol. 21, No. 1, 283–293, 1999.
9. GprMax2.0. <http://www.gprmax.org>.
10. Wang, G., *Study of Imaging Methods in Ground Penetrating Radar*, Civil Aviation University of China, Tianjin, 2004.
11. Stolt, R., “Migration by Fourier transform,” *Geophysics*, Vol. 43, No. 1, 23–48, 1978.
12. Subiza, B., E. Gimeno-Nieves, and J. M. Lopez-Sanchez, “An approach to SAR image by means of non-uniform FFT’s,” *2003 IEEE International Geoscience and Remote Sensing Symposium Proceedings, IGARSS’03*, Vol. 6, 4089–4091, Jul. 2003.
13. Xiu, Z., J. Chen, G. Fang, and F. Li, “Ground penetrating radar imaging based on F-K migration and minimum entropy method,” *Journal of Electronics & Information Technology*, Vol. 29, No. 4, Apr. 2007.
14. Chang, W., Q. Song, and D. Liang, “Implementation and application of resampling interpolation,” *Systems Engineering and Electronics*, Vol. 22, No. 5, 2000.

Analysis of MMW Imaging System with Scanning Mirrors and Extended Hemispherical Lens

Zucun Zhang and Wenbin Dou

State Key Laboratory of Millimeter Waves, Southeast University, Nanjing 210096, China

Abstract— In this paper, Small planar focal plane array (FPA) for millimeter-wave (MMW) imaging system is realized by adding an extended hemispherical lens to a scanning and focusing system, which is composed of a scanning planar mirror and a focusing parabolic mirror. Firstly, optical path and spot-patterns for different incident angles are calculated and compared without or with an extended hemispherical lens by using ray-tracing method. Numerical results demonstrate that for the small planar FPA the focusing performance can be improved greatly by adding an extended hemispherical lens when the incident wave is unparallel to the z -axis. Secondly, for different incident angles, the contour diagram of electric field in the focal plane of imaging system with extended hemispherical lens is calculated by ray-tracing method and Stratton-Chu vector diffraction integral formula. The numerical results show that the system with extended hemispherical lens has good imaging performance.

1. INTRODUCTION

Millimeter-wave (MMW) imaging has proven to be a useful technique for penetrating bad weather, seeing through some certain building materials, or detecting concealed articles under clothing. In the last several decades, many MMW imaging systems have been proposed [1–4]. However, long acquisition times pose serious limitations to the use of MMW imaging. Recently, MMW imaging systems that utilize focal array systems are of particular interest because they can solve the problems of long acquisition times. However, the number of detectors on focal plane, which is very important for the performance of imaging, is limited due to the cost and techniques. Fortunately, the scanning system can minimize the receiver numbers while keep high performance, thus using small receiving arrays combining the mechanical scanning is a feasible method [5].

Offset parabolic mirror is usually used as energy focusing antenna in multi-beam system. However, due to the characteristics of offset parabolic mirror, the receiving arrays must be arranged on the curved surface, which is a hard work and high-cost [6, 7]. The extended hemispherical lens has properties that if an optical system is designed such that all the rays are being focused to a point, the extended hemispherical lens can be added to the system and all the rays will still focus to a point [8]. Taking use of the properties, we introduce an extended hemispherical lens in our system to construct a small plane array on the extended surface of hemispherical lens. Numerical results show that the imaging system with extended hemispherical lens has good performance.

2. ANALYSIS AND CALCULATION

2.1. Geometry of the Imaging System

Figure 1 shows the optical path of the imaging system with extended hemispherical lens. Incident plane waves enter from the left and are brought to the parabolic mirror by the rotatable planar mirror, and then the waves are focused on to the focal plane of the imaging system, see Fig. 1. For incident wave with different incident angles, the planar mirror will scan so that to keep the wave reflected from the planar mirror be focused on to the center point of focal plane. The parabolic mirror is 90° offset and the equation is $y^2 + z^2 = 4f(f - x)$, its focus point is on the origin point and the focal length f is 160 mm.

The diameter of incident diaphragm is 160 mm, and the frequency is 100 GHz. The size of the system is about $400 \text{ mm} \times 600 \text{ mm} \times 160 \text{ mm}$. In this paper, θ_x , θ_y and θ_z are the angles between the incident plane wave and the x -axis, y -axis and z -axis respectively. be greatly improved by using extended hemispherical lens.

2.2. Optical Path and Spot-patterns

In Fig. 2 the focusing characteristics of the imaging system for different incident angles without or with extended hemispherical lens are shown respectively. As shown in Fig. 1, the position of focal plane for the system without or with extended hemispherical lens is different. From Fig. 2, it can be seen clearly that if the incident wave is unparallel to z -axis, the focusing performance can

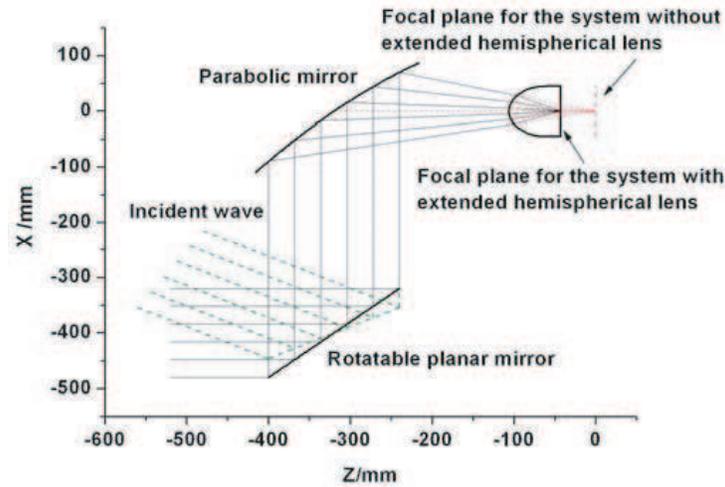


Figure 1: The optical path of the imaging system.

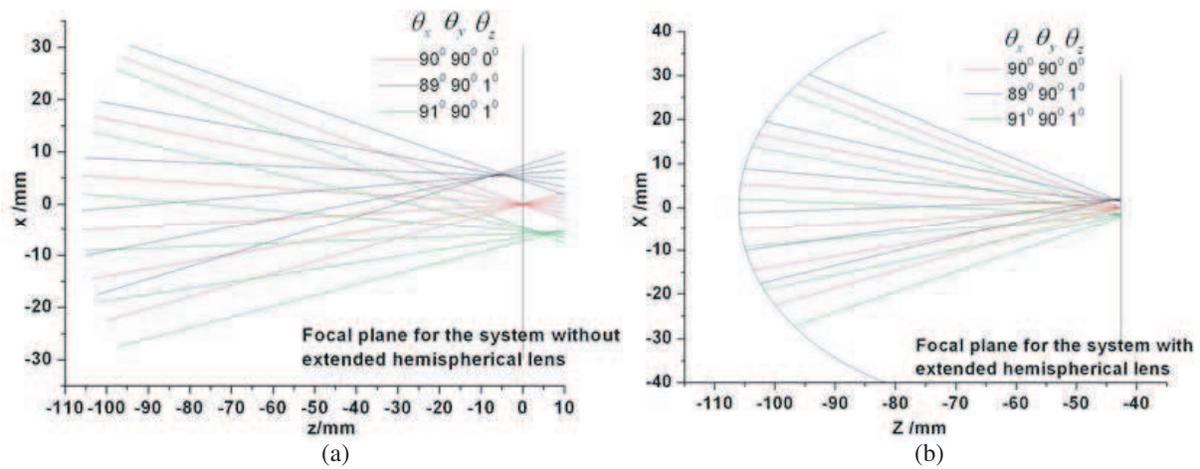


Figure 2: For different incident angles, the focusing characteristics of the imaging system without or with extended hemispherical lens. (a) Without extended hemispherical lens, (b) with extended hemispherical lens.

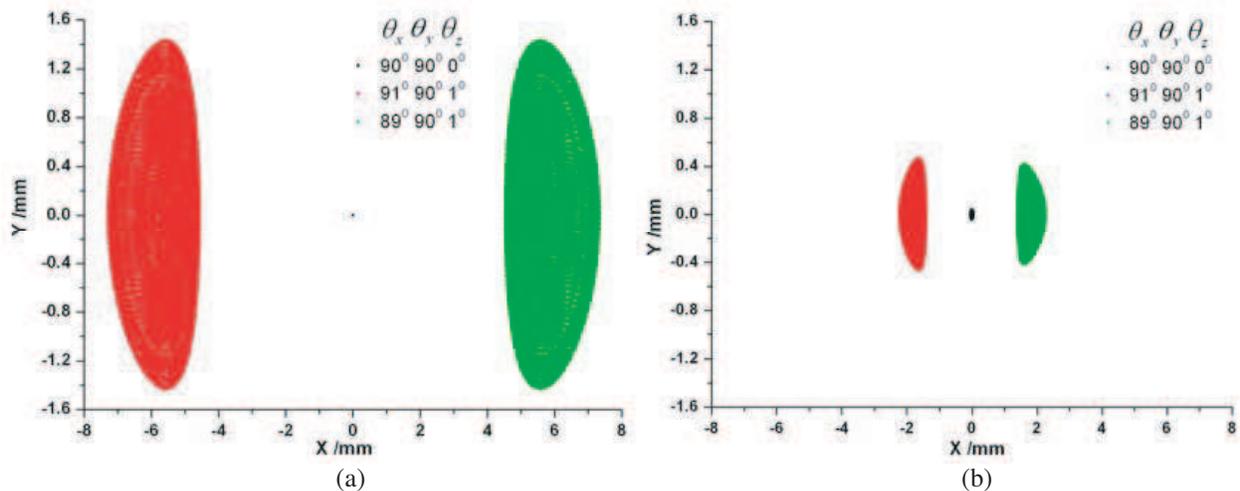


Figure 3: For different incident angles, the spot-patterns on focal plane of the imaging system without or with extended hemispherical lens. (a) Without extended hemispherical lens, (b) with extended hemispherical lens.

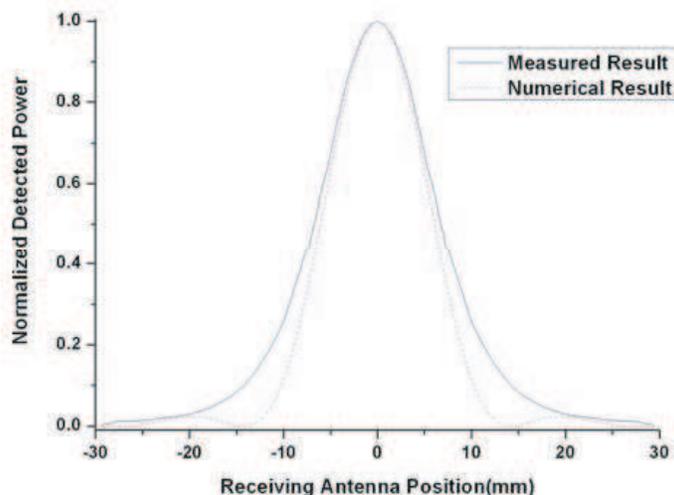


Figure 4: Calculation result compared with experiments given in reference [10].

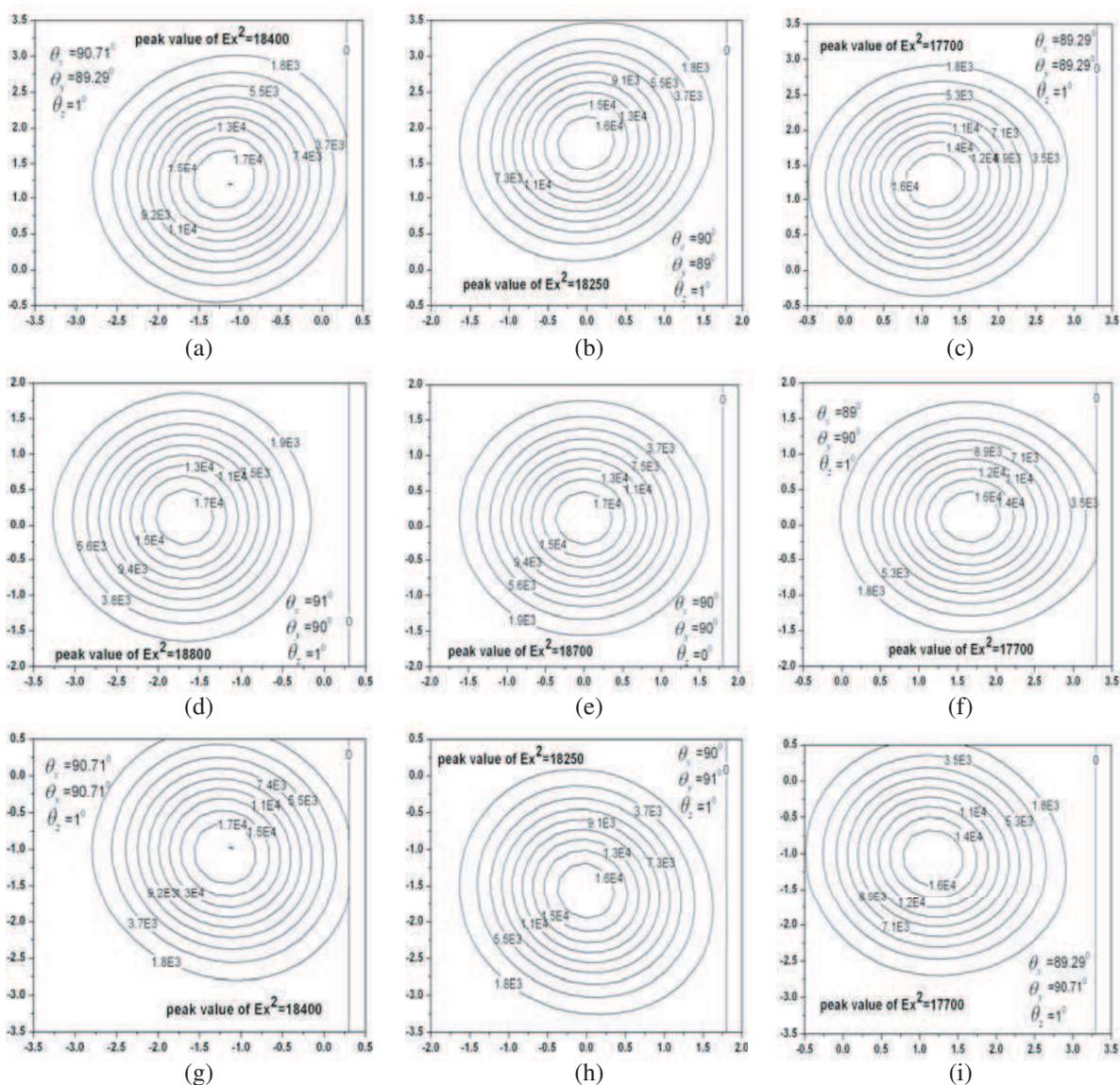


Figure 5: For different incident angles, the diffraction pattern of $|E_x|^2$ for different incident angles.

The spot-patterns of the imaging system without or with extend hemispherical lens are illustrated in Fig. 3 respectively. The incident wave with different incident angles will focus in the different position on the focal plane. Comparing the two figures, we can find that when the incident wave is unparallel to the z -axis, extended hemispherical lens can decrease the spot-patterns largely, which means the imaging performance will be improved greatly. In the following text, we will calculate and analysis the field on the focal plane of the imaging system with extended hemispherical lens.

2.3. Contour Diagram of Electric Field

Using the ray-tracing method and Stratton-Chu vector diffraction integral formula [9], we have calculated the field distribution on the focal plane of the reverse-microscope system in [10] at X waveband, the parameters are the same as the reference [10]. Our numerical results are compared with the measured result, better agreement has been found as shown in Fig. 4.

In Fig. 5, the diffraction pattern for different incident angles are shown when the planar mirror are kept fixedly. It can be seen clearly that the plane wave for different incident angles has its diffraction pattern in the different position of focal plane. The aberration and the difference of peak value of $|E_x|^2$ are very small and can be tolerated. A linearly polarized plane wave with different incident angle corresponds to the different point of the scene. According to these, for the imaging system, we can arrange the detectors in different position of the focal plane to receive the wave emitting from the different point of the scene. Moreover, the diffraction pattern can provide enough information for evaluating the imaging performance and selecting the detector elements.

3. CONCLUSION

In the present paper, small planar FPA for MMW imaging application is constructed by bringing an extended hemispherical lens into an imaging system, which is composed of a scanning planar mirror and a focusing parabolic mirror. The numerical results indicate that the improvement of the imaging performance can be reached by using the system with an extended hemispherical lens, when compared with the system without an extended hemispherical lens. Our work is very meaningful to design a low-cost and practical system for millimeter-wave imaging.

ACKNOWLEDGMENT

This work is supported by the Astronomy United Foundation of NSFC and CAS (No. 10778602).

REFERENCES

1. Goldsmith, P. F., C. T. Hsieh, et al., "Focal plane imaging systems for millimeter wavelengths," *IEEE Trans. Microwave Theor. Tech.*, Vol. 41, No. 10, 1164–1675, 1993.
2. Appleby, R., R. N. Anderton, S. Price, N. A. Salmon, et al., "Compact real-time passive millimetre-wave imager," *Proc. SPIE*, Vol. 3703, 13–19, 1999.
3. Martin, C., J. Lovberg, S. Clark, and J. Galliano, "Real time passive millimeter-wave imaging from a helicopter platform," *Proc. SPIE*, Vol. 4302, 22–28, 2000.
4. Anderton, R. N., R. Appleby, J. E. Beale, P. R. Coward, and S. Price, "Security scanning at 94 GHz," *Proc. SPIE*, Vol. 6244, 62110C-1-7, 2006.
5. Doyle, R., B. Lyons, A. Lettington, et al., "Stand-off detection of hidden threat objects on personnel at checkpoints and in public areas using active millimetre-wave imaging," *Proc. SPIE*, Vol. 5619, 90–97, 2004.
6. Rudge, A. W., "Multiple-beam antennas: Offset reflectors with offset feeds," *IEEE Trans. Antennas Propag.*, Vol. 23, No. 10, 317–322, 1975.
7. Jasik, H, *Antenna Engineering Handbook*, Mcgraw-Hill Book Company Inc., 1982
8. Filipovic, D. F., S. Steven, and G. M. Rebeiz, "Double-slot antennas on extended hemispherical and elliptical silicon dielectric lenses," *IEEE Trans. Microwave Theor. Tech.*, Vol. 41, No. 10, 1738–1749, 1993.
9. Dou, W. B., J. G. Wang, and Z. L. Sun, "Diffracted field in the focal plane of a reverse microscope imaging system at millimeter wavelengths," *J. of Southeast University*, Vol. 32, 340–345, 2002.
10. Rutledge, D. B. and M. S. Muha, "Imaging antenna arrays," *IEEE Trans. Antennas Propag.* Vol. 30, No. 4, 535–540, 1982.

Simulation for GPR Echoes Based on Non-constant-Q Attenuation Model

Weikun He^{1,2}, Zhigang Su¹, Renbiao Wu^{1,2}, Zhiyong Han¹, and Jiaxue Liu¹

¹Tianjin Key Lab for Advanced Signal Processing
Civil Aviation University of China, 300300, China

²Electronic and Information Engineering Institute
Tianjin University, 300300, China

Abstract— In this paper, we discuss the simulation for GPR echoes propagated in a lossy media and propose the echoes simulation method based on non-constant-Q attenuation model which conforms to the actual situation of the underground media. The signal model of GPR echoes and the simulation wave based on non-constant-Q attenuation model are established. Simulation results demonstrate the new method has better performance than the method based on the constant-Q attenuation model.

1. INTRODUCTION

Ground Penetrating Radar (GPR), as a high-resolution, high-precision and non-destructive detection tool, is widely applied in the quality surveillance for the highways, bridges and runways ect [1]. The underground media belongs to lossy media whose dielectric response is frequency-dependent, so its quality factor, Q , is frequency-dependent which makes it more complicate for the electromagnetic wave to propagate in the layer media [1]. It is important to study electromagnetic wave propagation characteristic in non-constant-Q layer media and echoes simulation technique for the inversion of the underground structure.

Much attention has been attracted to the simulation for GPR echoes propagated in a lossy layer media. The outstanding products include the methods based on the echo model [2, 3] and on the Finite Difference Time Domain (FDTD) [4]. For the echo-model-based method, the propagation attenuation of electromagnetic wave is based on the constant-Q model in which dielectric permittivity of the media obeys a complex power function of frequency [2, 3]. However, for the lossy media whose dielectric permittivity doesn't comply to a complex power function of frequency, the quality factor isn't constant, but frequency-dependent. This mismatch will affect the simulation precision and the constant-Q model above is invalid for such situations. For the FDTD-based method, more complexity and operation burden are required concerned with the solution to the Maxwell equations iteratively and the setting of the absorbing boundary conditions (ABC).

In this paper, we discuss echoes simulation based on non-constant-Q attenuation model. The simulation results are compared and analyzed based on the data obtained by the GPR data simulation software (GprMax) and demonstrate the new method is valid.

2. PROBLEM FORMULATION

GPR transmits high frequency and ultra-wideband pulse which can be reflected due to variations in the electrical properties of the investigated medium. The returned echoes of GPR signal for a multilayered medium are superimposed signals reflected from the boundaries of different media, which can be described by:

$$y(t) = \sum_{n=1}^N s(t) * h_n(t) + \varepsilon(t) \quad (1)$$

where N denotes the number of independent echoes, $s(t)$ is the transmitted pulse, $h_n(t)$ is the impulse response of the channel through which the n -th echo propagates and $\varepsilon(t)$ denotes unmodeled clutter and noise. The impulse response $h_n(t)$ can be written as:

$$h_n(t) = \prod_{l=1}^{L_n} a_{n,l} h_{n,l}(t) \quad (2)$$

where L_n is the number of the media layers which the n -th echo propagates through, $h_{n,l}(t)$ is the impulse response of the n -th echo in the l -th media and $a_{n,l}$ is the change of the n -th incidence

wave amplitude which enters the l -th media and it is frequency-dependent. For example, if the EM wave is reflected on the boundary, $a_{n,l}$ is reflection parameter and $h_{n,l}(t)$ is the same as $h_{n,(l-1)}(t)$. However, if the EM wave is refracted on the boundary, $a_{n,l}$ is the refraction parameter and $h_{n,l}(t)$ is different from $h_{n,(l-1)}(t)$.

The impulse response of any media can be described as

$$h_{n,l}(t) = \exp(-\alpha_{n,l}d_{n,l})\delta(t - \tau_{n,l}) \quad (3)$$

where $\alpha_{n,l}$ is the attenuation coefficient of the corresponding media, $d_{n,l}$ and $\tau_{n,l}$ are the depth and time delay of the n -th echo in the l -th media respectively. They are both frequency-dependent.

From Equations (1)–(3), the spectrum of the received signal $y(t)$ can be written as

$$Y(\omega) = S(\omega) \sum_{n=1}^N \prod_{l=1}^{L_n} a_{n,l} \exp\{-\alpha_{n,l}d_{n,l} - j\omega\tau_{n,l}\} \quad (4)$$

By Equation (4), we can see that the simulation problem for GPR echoes is actually to solve the model parameter $a_{n,l}$, $\alpha_{n,l}$ and $\tau_{n,l}$ on the condition that the transmitted pulse $s(t)$ and the media property are given. The received signal $y(t)$ or $Y(\omega)$ then can be obtained by the above model parameter.

3. QUALITY FACTOR

To describe attenuation characteristic of a non-magnetic lossy media, we use complex dielectric permittivity $\bar{\epsilon}$:

$$\bar{\epsilon} = \epsilon' - j\epsilon'' \quad (5)$$

where, ϵ' and ϵ'' are the non-negative real values.

The Q factor is useful for characterizing wave attenuation. The Q factor is defined as the ratio of total energy restored and the energy loss in one cycle and can be written as [3]:

$$Q = \epsilon' / \epsilon'' \quad (6)$$

For the constant-Q model, the dielectric permittivity obeys a complex power function of frequency [2, 3]:

$$\bar{\epsilon} = \epsilon^0 (j\omega/\omega_0)^{-2\gamma} \quad (7)$$

where ω_0 is the reference frequency, ϵ^0 is equal to the dielectric permittivity value for $\gamma = 0$ and γ is constant, then the Q factor can be described as [2, 3]

$$Q = 1/\tan(\pi\gamma) \quad (8)$$

here, $Q > 0$ and $0 < \gamma < 1/2$. We can see that the Q factor is constant and frequency-independent.

For a lossy media, Debye model can be used to describe the dispersion characteristic of the media and the frequency-dependent permittivity can be described as:

$$\bar{\epsilon} = \epsilon_\infty + (\epsilon_s - \epsilon_\infty)/(1 + j\omega\zeta) - j(\sigma/\omega) \quad (9)$$

where, ϵ_s is the static (DC) permittivity, ϵ_∞ is the permittivity at theoretically infinite frequency, ζ is relaxation time of the media and σ is conductivity. From Equations (6) and (9), we can see that the Q factor is frequency dependent for the non-constant-Q model. The simulation for GPR echoes based on non-constant-Q model is mainly studied in the following section.

4. INTERFACE EFFECT ON ECHOES

The amplitude of the received signal is related with the loss caused by the reflection and refraction in the interface and that caused by the electromagnetic propagation within the media. The time delay lies on the depth of the media and the phase velocity propagated in the media. This section will discuss the effect of the interface between the different media.

When the n -th echoes is reflected in the l -th interface,

$$a_{n,l} = \left(\sqrt{\epsilon'_l} - \sqrt{\epsilon'_{l+1}} \right) / \left(\sqrt{\epsilon'_l} + \sqrt{\epsilon'_{l+1}} \right) \quad (10)$$

When the n th echoes is refracted in the l th interface,

$$a_{n,l} = 2\sqrt{\varepsilon'_l} / \left(\sqrt{\varepsilon'_l} + \sqrt{\varepsilon'_{l+1}} \right) \quad (11)$$

where, ε'_l and ε'_{l+1} are the real part of the dielectric permittivity of the l -th and $(l+1)$ -th layer media.

5. LAYER MEDIA EFFECT ON ECHOES

To analyze the loss caused by electromagnetic wave propagated within the layer media, the complex wave number k is used and it can be defined as:

$$k = \beta - j\alpha \quad (12)$$

where, β is phase coefficient and α is attenuation coefficient. Then β and α can be described again as [3]:

$$\beta = \omega \left[\mu\varepsilon' \left(\sqrt{1 + (1/Q)^2} + 1 \right) / 2 \right]^{\frac{1}{2}} \quad (13)$$

$$\alpha = \omega \left[\mu\varepsilon' \left(\sqrt{1 + (1/Q)^2} - 1 \right) / 2 \right]^{\frac{1}{2}} \quad (14)$$

The effective factor of the echoes amplitude $e^{-\alpha_{n,l}d_{n,l}}$, caused by the loss of the n -th echoes propagated in the l th layer media, is related with attenuation coefficient $\alpha_{n,l}$ and the depth of the media $d_{n,l}$. Let

$$\tan \delta = 1/Q \quad (15)$$

then α is related to β as follows:

$$\alpha/\beta = \tan(\delta/2) \quad (16)$$

The effective factor $e^{-\alpha_{n,l}d_{n,l}}$ can be written as:

$$\exp(-\alpha_{n,l}d_{n,l}) = \exp[-\beta_{n,l}d_{n,l} \tan(\delta_{n,l}/2)] = \exp[-(\omega\tau_{n,l}) \tan(\delta_{n,l}/2)] \quad (17)$$

where, $\tau_{n,l}$ and $\delta_{n,l}$ are the corresponding time delay and δ value for non-constant-Q media. From Equation (17), it can be seen that the loss of electromagnetic wave within the layer media is frequency-dependent.

The phase velocity of the n -th echoes propagated in the l -th layer media can be written as:

$$v_{n,l} = \omega/\beta_{n,l} \quad (18)$$

Substituting Equations (13) and (16) to Equation (18), the corresponding phase velocity can be obtained,

$$v_{n,l} = \sqrt{\cos \delta_{n,l}} / \left[\sqrt{\mu\varepsilon'_l} \cos(\delta_{n,l}/2) \right] \quad (19)$$

From Equation (19), the time delay can be obtained,

$$\tau_{n,l} = d_{n,l}/v_{n,l} = \sqrt{\mu\varepsilon'_l} \cos(\delta_{n,l}/2) / \sqrt{\cos \delta_{n,l}} \quad (20)$$

6. EXPERIMENT RESULTS

To verify the efficiency of the proposed method, the simulation results were compared and analyzed based on the data obtained by GPR data simulation software (GprMax) which is widely used and is representational for the GPR data simulation. In GprMax software, GPR echoes can be obtained based on the FDTD method [5].

We present some numerical examples of runways to illustrate the performance of the proposed algorithms. In the following examples, the runways consist of surface media and base media. The characteristic of surface media asphalt, is Debye dispersion characteristic. The corresponding parameters for the media are $\varepsilon_s = 6\varepsilon_0$, $\varepsilon_\infty = 2\varepsilon_0$ (ε_0 is the air dielectric permittivity), $\zeta = 0.1$ ns, $\sigma = 0.001$ S/m. The base media is assumed to be non-dispersion media because we needn't consider

the bottom reflection of the base media. The corresponding parameters of the base media are $\varepsilon_\infty = \zeta = \sigma = 0$, $\varepsilon_s = 22\varepsilon_0$. The transmitted signal is a Ricker pulse [5] with the centre frequency $f_c = 900$ MHz, the time window $T_w = 10$ ns, the sampling interval $T_s = 0.07$ ns and the sampling frequency $f_s = 1.41 \times 10^4$ MHz which can be decreased by the sequencing sampling [1].

Firstly, suppose the depth of the surface media $d_1 = 0.1$ m. The simulation echoes are shown in Figure 1, in which the dashed line is the result with GprMax software (FDTD method), dashdotted line is the result with constant-Q model, solid line is the result with non-constant-Q model and lineation is the result without the consideration of dispersion. To improve the visualization effect, we amplified locally Figure 1(a) to obtain Figure 1(b). Compared with the echoes obtained by the method without the consideration of dispersion, the echoes obtained by the methods with the consideration of dispersion have changed a lot in both the waveform and the amplitude. So the estimation performance of the echoes parameters (amplitude, time delay ect) is effected if we don't consider the dispersion. Compared with the echoes obtained by constant-Q model, the echoes obtained by non-constant-Q model is closer to that obtained by GprMax software because of the consideration of the actual characteristic of the quality factor. However, because of the effect of the following factors, including the operation precise, absorbing boundary conditions ect, the echoes obtained by the proposed method and that obtained by GprMax software have small discrepancy.

To verify the relation of attenuation and dispersion with the depth of the media, we increase the depth of the surface media and suppose $d_1 = 0.35$ m. Figure 2 shows the received echoes from which we can see that the attenuation and dispersion increase with the depth of the media. In addition, the discrepancy of the echoes obtained by constant-Q model and non-constant-Q model increase with the depth of the media.

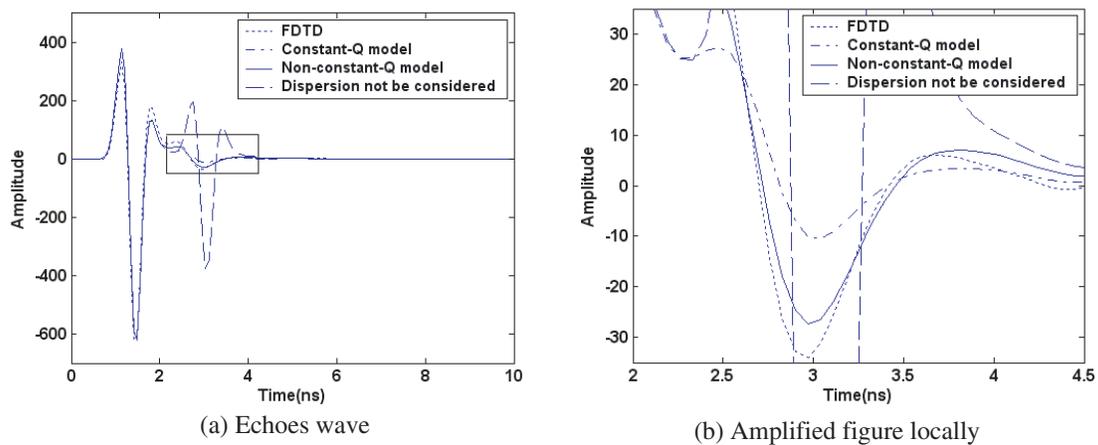


Figure 1: Echoes simulation for $d_1 = 0.1$ m.

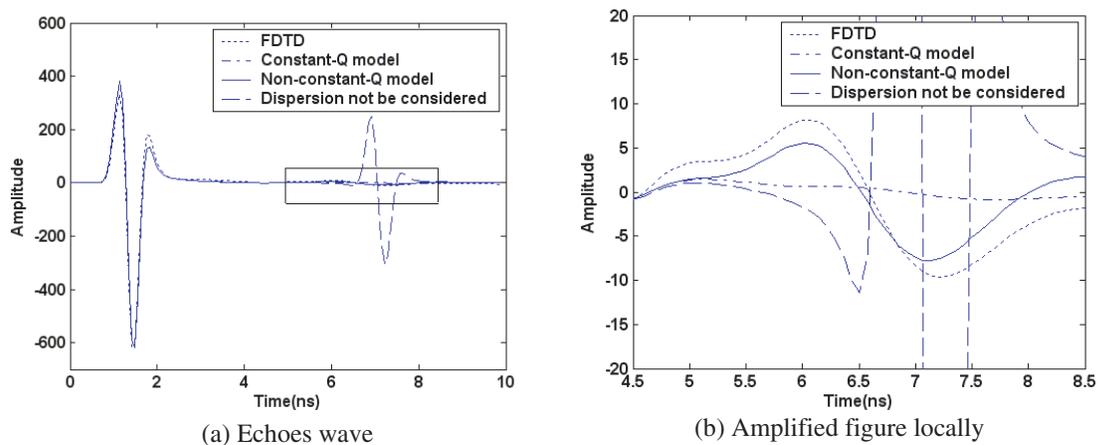


Figure 2: Echoes simulation for $d_1 = 0.35$ m.

7. CONCLUSION

Based on the EM characteristic of a lossy media, we proposed non-constant-Q model in this paper. The signal model of GPR echoes and simulation wave were established based on non-constant-Q model. By comparing with the widely-used GprMax software, the simulation results demonstrate the proposed method has better performance than that based on the constant-Q model.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China under grant 60879019, and by the Civil Aviation Technology Project of China under grant MHRD0701.

REFERENCES

1. Daniels, D. J., *Surface Penetrating Radar*, IEE Press, London, United Kingdom, 1996.
2. Carcione, J. M., "Theory and modeling of constant-Q P- and S-waves using fractional time derivatives," *Geophysics*, Vol. 74, T1–T11, 2009.
3. Bano, M., "Modeling of GPR waves for lossy media obeying a complex power law of frequency for dielectric permittivity," *Geophysical Prospecting*, Vol. 52, 11–26, 2004.
4. Sullivan, M., *Electromagnetic Simulation Using FDTD Method*, IEEE Press, 2000.
5. Antonis, G., "Modeling ground penetrating radar using GprMax," *Construction and Building Material*, Vol. 19, 755–762, 2005.

Adaptation in Front of Ground Penetrating Radar (GPR) Antenna by Layered Dielectric Slab and Resistive Loading

Y. Wahyu¹, R. S. Sianipar², A. Kurniawan³, Sugihartono³, and A. A. Lestari⁴

¹Research Center for Electronics and Telecommunication-LIPI, Indonesia

²Radar and Communication Systems (RCS)-Solusi247, Indonesia

³School of Electrical Engineering and Informatics, Bandung Institute of Technology, Indonesia

⁴IRCTR, Delft University of Technology-Indonesia Branch, Indonesia

Abstract— Reflections from the field boundary is a medium that can not be avoided in the GPR antenna. These reflections lead to a successful energy transmitted into the medium becomes smaller. This is indicated by a decline amplitude transmitted waves. To minimize this reflection is used as a medium of dielectric layer transition from the antenna to the medium. By using this dielectric layer from amplitude expected to signal that the medium will be higher. Apart from amplitude, should also noted ringing level. A good antenna has a small ringing level. Therefore, in this simulation is used resistive loading. With resistive loading is expected to wave reflection from the end of the antenna can be reduced. So with the imposition of resistive and dielectric layers are expected to signal that is transmitted to the medium has a low ringing level and also has a high amplitude.

1. INTRODUCTION

GPR is a useful device for the detection of objects buried beneath the soil surface to a certain depth without the need to be done soil excavation. With GPR, the various activities or research to find out information about the situation subsurface ground can be done quickly and easily. In the GPR antenna systems play an important role, because the general performance of the GPR using impulse radar is determined by the ability of impulse radiating antenna to the ground with the loss and minimum distortion. This means that the impulse GPR antennas must be able to minimize late-time ringing. Resistive loading is used to overcome these internal reflection [1].

Between the antenna and the medium used as a transition dielectric layer electromagnetic field to the medium. To analyze the amplituda form in the medium to be used with the method of numerical modeling Finite-difference time-domain (FDTD) using FDTD3D software. The selection of this method with the consideration that the outcome is to get the waveform in time domain. Other benefits include the use of FDTD: FDTD to work effectively on systems that use a monocycle pulse as the excitation source, then FDTD allows the user to define the material properties paa all points in computational domain so that the antenna is designed more realistic [2].

2. ANTENNA SYSTEM DESIGN

The proposed GPR antenna is dipole antennas with resistive loading and the dielectric layer. Use no other dipole antenna because the antenna is a dipole antenna is often used for GPR applications mainly because of its simplicity [3]. The main problem dipole antenna for GPR applications is its narrowband, but for the application required GPR antenna with ultra-wideband characteristics. To overcome this done at arm antenna resistive loading (let's call this arm load) with the Wu-King profile to reduce the late-time ringing due to multiple reflection between the tip of the antenna and the feed point. Antenna geometry can be seen in the Figure 1. Dashed lines represent the load arm and the gap that separates the line is a resistive load by inserting lumped element resistor in accordance with the Wu-King profile. The number of resistors used 65 pieces with initial resistance of 200 Ohms. From [5] note that the distance between the feed point with the first resistor is selected as far as $c/(f\sqrt{\epsilon_r})$ where c is the speed of light, f is the frequency of the pulse and a relatively permittivity substrate ($\epsilon_r = 4 : 34$) to radiation from the first resistor strengthen the radiation from the feed point to the antenna broadside.

Furthermore antenna mounted on a dielectric layer 6 as a transition from the antenna to the medium as shown in the Figure 2. Dielectric layer profile used is linear where ϵ_r difference between the successive dielectric layer equal. The reason this election is a linear profile because of its simplicity. ϵ_r value of each layer of dielectric and thickness can be seen in Table 1. This simulation is used for the observation point 6 as shown in the Figure 3. Point 1 and 2 are used to generate

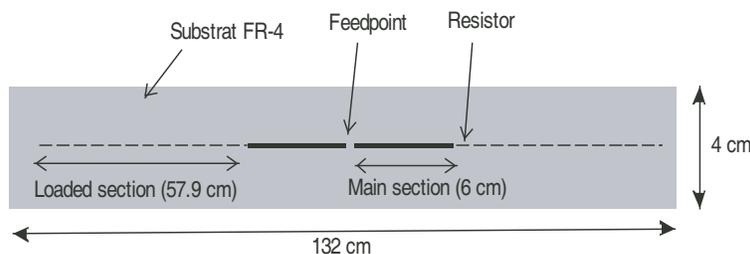


Figure 1: Geometry dipole antennas with resistive loading.

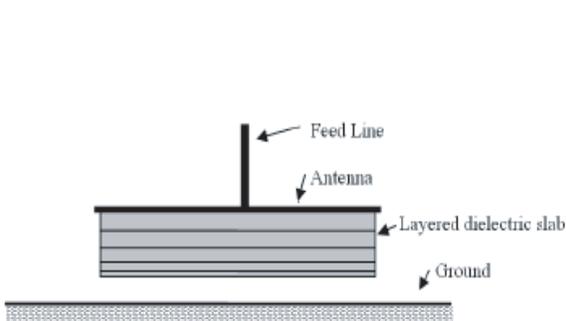


Figure 2: Antenna with dielectric layers.

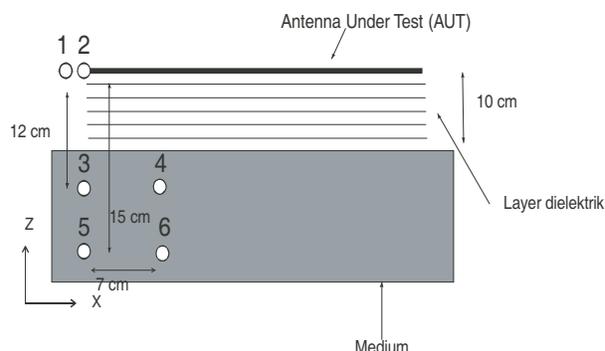


Figure 3: Observation points simulation.

graphs and VSWR input impedance because the program does not generate direct FDTD3D the input impedance and VSWR. The data generated from point 1 and 2 must be further processed by using Matlab to generate graphics input impedance and VSWR. Point 3, 4, 5 and 6 are used to seeing amplituda peak to peak signal and the resulting ringing. Medium used have $\epsilon_r = 10$ and $\sigma = 0.05$).

3. SIMULATION RESULTS

3.1. The Influence of Resistive Loading

3.2. The Influence of the Addition of a Dielectric Layer and Resistive Loading

4. RESULTS

The influence of resistive loading can be seen from Table 2. Obtained from the table that with the addition of the resistive loading level of ringing can be reduced to about 1% (-40 dB). In the high-resolution GPR applications necessary condition ringing maximum level 1% (-40 dB). Simulation results indicate a resistive load is able to meet high resolution requirements. However, this resistive loading will give dimension larger antennas. In terms of input impedance, when there is no resistive load is the input impedance fluctuations in the range 0–2.5 GHz. This shows the nature of the narrowband dipole antenna. As for the application of GPR antenna ultrawideband. Dengan required the addition of resistive loading the resulting graph is more flat. With $Z_0 = 200$ Ohms VSWR can be seen in the Figure 7. Resonannya frequency around 830 MHz. Frequency range of VSWR = 2

Table 1: Profile of a dielectric layer is used.

Layers	ϵ_r	Tickness
1	2.28	1.67 cm
2	3.57	1.67 cm
3	4.85	1.67 cm
4	6.14	1.67 cm
5	7.42	1.67 cm
6	8.71	1.67 cm

Table 2: Level ringing with resistive loading and without resistive loading.

Points	Level ringing without resistive loading (%)	Level ringing without resistive loading (%)
Point 3	160.54	1.14
Point 4	108.33	1.06
Point 5	168.67	1.11
Point 6	110.23	1.09

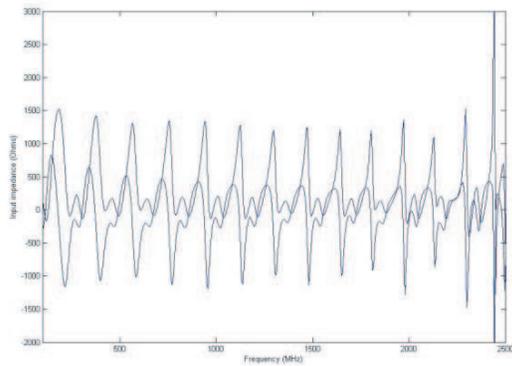


Figure 4: Input impedance without resistive loading.

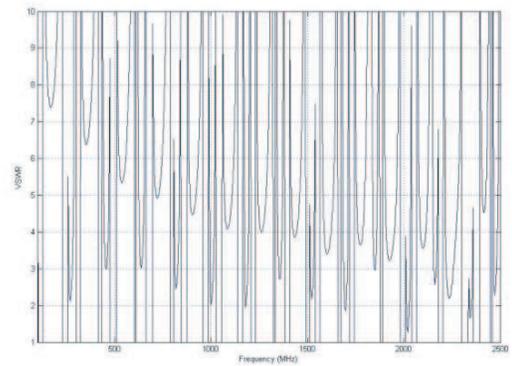


Figure 5: VSWR without resistive loading ($Z_0 = 200$ Ohm).

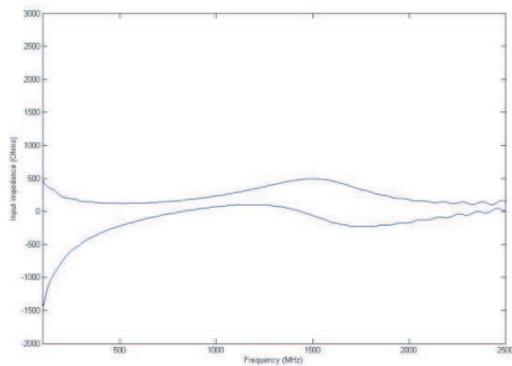


Figure 6: Input Impedance with resistive loading (initial resistance 200 Ohm).

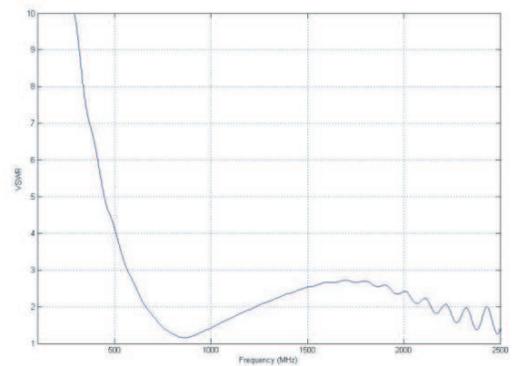


Figure 7: VSWR with resistive loading (first resistance 200 Ohm).

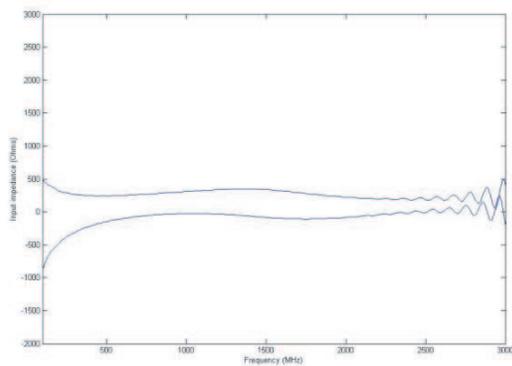


Figure 8: Input Impedance with resistive loading (initial resistance 40 Ohm).

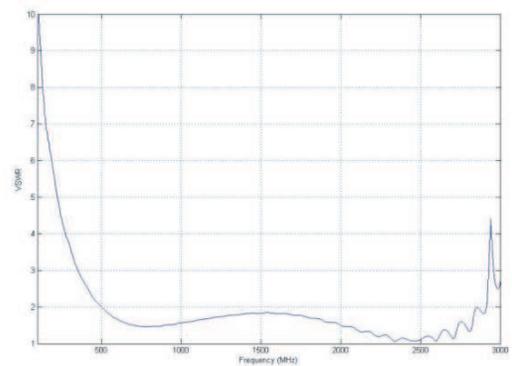


Figure 9: VSWR with resistive loading (initial resistance 40 Ohm).

is 666–1233 MHz. Bandwidth is obtained around 567 MHz (Fractional bandwidth = 0 : 59). Terms of ultrawideband Fractional bandwidth is greater than 0.25. Therefore, the proposed antenna are eligible ultrawideband. In Figures 6 and 7 used the early resistance of 200 Ohm and follow the Wu-King profile. If the initial resistance was changed to 40 Ohms and still follow the Wu-King profile of the results obtained as shown in picture 8 and 9. Obtained input impedance is more flat than using the initial resistance of 200 Ohms. With the initial resistance of 40 Ohms to get the frequency range of VSWR = 2 is 508 to 2910 MHz. Antenna bandwidth around 2.4 GHz (Fractional bandwidth 1.40) and the resonant frequency of 800 MHz. Input impedance produced by the Wu-

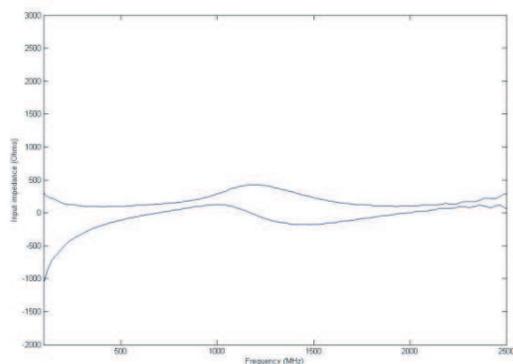


Figure 10: Input impedance with resistive loading (initial resistance 200 Ohm).

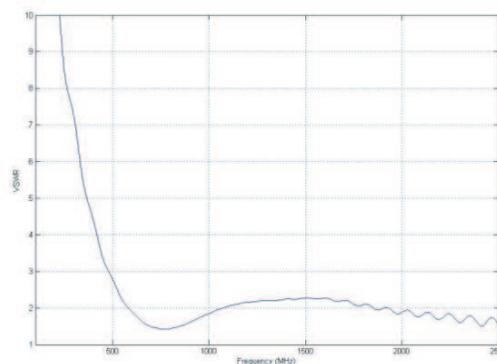


Figure 11: VSWR dengan layer dielektrik dan bebanan resistif (resistansi awal 200 Ohm).

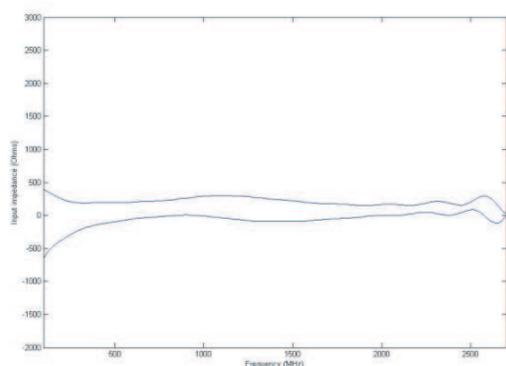


Figure 12: Input impedance with resistive loading (initial resistance 40 Ohm).

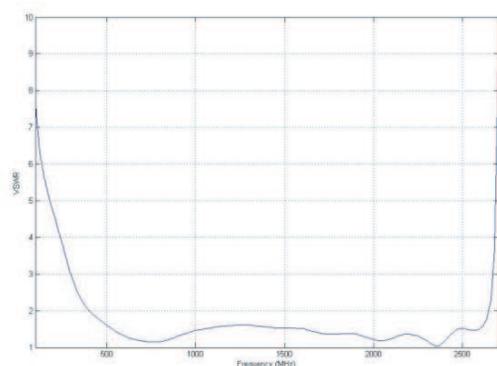


Figure 13: VSWR with resistive loading (initial resistancel 40 Ohm).

King profile with initial resistance of 40 Ohms is more flat than the initial resistance of 200 Ohms for a 40 Ohm resistansi early reflections occur from the lower end of the antenna. But in a simulation to see the effect of adding a dielectric layer using a resistive loading with initial resistance of 40 Ohms is only used to analyze the input impedance and VSWR alone. This is because when using FDTD to simulate a small resistance so it will take bigger. Especially for the ringing analyze the resulting data have to be greater than 2 times the pulse duration of 3.2 ns pulse begins early. Observation point inside the medium to make the necessary time interval for the signal travels up to the medium also increases thereby increasing the simulation time. So with the above considerations the effect of the addition of a dielectric layer is not done for the early resistance of 40 Ohms. Then after seeing the influence of the addition of a resistive load, see also the influence of the addition of a dielectric layer and the resistive loading at the same time. From Table 3 can be seen rising levels of ringing and the peak to peak primary pulse when compared with no addition of a dielectric layer. Ringing level on average increased to 4.56%–5.23% while the increase in the main pulse at each point varies from 304% at 3 points to 250% at point 6. The increase in the level of the main pulse shows a successful energy transmitted into the growing medium. The increase in the level of ringing is due to many reflections on the limits of 2 different medium. Because rapid vines in each medium is different then there is some signal to precede the other pulses. Signal to more than 3.2 ns pulse starts from the beginning would be considered a ringing. Input impedance and VSWR with initial resistance of 200 Ohms can be seen in the picture 3 : 10 and 3 : 11. From the picture when compared with Figures 6 and 7 can be seen form the graph is similar. The difference is that the dielectric layer shift of 130 MHz in the direction of smaller frequency. Frequency range of VSWR = 2 is 584 to 1050 MHz. Bandwidth is obtained around 466 MHz (Fractional bandwidth 0 : 57). Then if used resistive loading with initial resistance of 40 Ohms is obtained resonant frequency shift of 110 MHz to 690 MHz. Frequency range of VSWR = 2 is 397 to 2637 MHz. Bandwidth is obtained to be

Table 3: Level ringing and increased peak to peak primary pulse with the addition of a dielectric layer and resistive loading.

Points	Ringing Level (%)	Increased peak to peak primary pulse (%)
Point 3	5.23	304.62
Point 4	4.56	278.66
Point 5	5.22	270.46
Point 6	4.60	250.93

2240 MHz (Fractional bandwidth 1 : 47).

5. CONCLUSION

Of the overall simulation above can be seen that with the addition of resistive loading can suppress ringing the desired level. The addition will increase the dielectric layer peak to peak amplituda main pulse but will also raise the level of ringing. In terms of input impedance and VSWR, resistive loading can make the resulting input impedance is more flat. The addition of the proposed dielectric layer does not change the form input impedance and VSWR as a whole produced but will shift the graph is approximately 100 MHz to the lower frequencies. Ringing level generated by the linear-profile dielectric layer is not able to meet the high-resolution GPR applications that require ringing level is less than 1% (−40 dB). Therefore necessary to develop longer-layer dielectric profile and the optimal resistive load so that it can provide peak to peak amplituda high but can still keep ringing levels below 1% (−40 dB).

REFERENCES

1. Lestari, A. A., A. G. Yarovoy, and L. P. Ligthart, “Adaptive antenna for ground penetrating radar,” Delft University of Technology, The Netherlands.
2. Daniels, D. J., *Ground Penetrating Radar*, 2nd Edition, The Institution of Electrical Engineers, London, United Kingdom.
3. Montoya, T. P. and G. S. Smith, “A study of pulse radiation from several broad-band loaded monopoles,” *IEE Trans. Antennas Propagat.*, Vol. 44, No. 8, 1172–1182, Aug. 1996-a.
4. Lestari, A. A., D. Yulian, A. B. Sukmono, E. Bharata, A. G. Yarovoy, and L. P. Ligthart, “Rolled Dipole Antenna for Low-resolution GPR,” *PIERS Proceedings*, 2217–2220, Beijing, China, March 26–30, 2007.
5. Lestari, A. A., A. G. Yarovoy, and L. P. Ligthart, “RC loaded bow-tie antenna for improved pulse radiation,” *IEEE Trans. Antennas Propagat.*, Vol. 52, No. 10, 2555–2563, Oct. 2004.

Consideration of Antenna Pattern Design for FY3 Precipitation Measurement Satellite Dual-frequency Precipitation Radar

Honggang Yin¹ and Xiaolong Dong²

¹National Satellite Meteorological Center, China

²Center for Space Science and Applied Research CAS, China

Abstract— A dual-frequency precipitation radar (DPR), which will be the major payload of the FengYun-3 (FY3) Precipitation Measurement Satellite, is being developed in China. This radar will be used to observe Typhoon rain and other precipitation. In the case of spaceborne precipitation radar, the backscattered echo of the sea surface is much stronger than the rain echo. In this paper, the model of sea clutter interference with precipitation measurement of the FY3 DPR is built, and the effects of sea clutter on the design of the antenna radiation pattern are analyzed in detail. Based on the analysis of surface clutter, the requirements for antenna beam width and radiation sidelobe level are proposed.

1. INTRODUCTION

China is one of the countries which suffer most serious damage caused by typhoon. There are about 6.9 landing tropic cyclones per year on average in China that is more than other countries of the northwest Pacific [1]. Typhoon Morak claimed hundreds of lives in Taiwan 2009, and the direct economic losses from this typhoon exceeded RMB 10 billion. Hard rain accompanying with Typhoon is one of the primary factors which lead to disaster. Since Typhoon and other tropical cyclones are over sea during much of their life cycle, there is only the spaceborne radar that can obtain their three dimensional structure and energy transfer, and all these information will be helpful to estimate the surface rain rate and to forecast the track. China is planning to launch a FengYun-3 meteorological satellite, i.e., the Precipitation Measurement Satellite (FY3 PMS), to measure precipitation in the future. Its major payload is a dual-frequency precipitation radar (DPR).

For spaceborne precipitation radar, the rain volume is proximate to the Earth's surface, and the backscattered echo of land surface is much stronger than that from the rainfall. Usually, sea surface echo is stronger than the land surface backscattering. When rain rate is 0.5 mm/h, typical value of the power ratio of the returned signal from precipitation to that from sea surface for a precipitation radar with frequency of 13.8 GHz is $-35\text{ dB}\sim-55\text{ dB}$ [2]. Therefore the received rain echo may be masked or contaminated by the serious sea clutter through antenna beam while observing precipitation adjacent to the sea surface.

It is obvious that the sea clutter must be considered for the design of performance parameters for the FY3 DPR. The development of the FY3 DPR prototype is introduced briefly at first in this paper. Then the effects of sea clutter on the design of the antenna radiation pattern including antenna beam width and sidelobe level are analyzed in detail.

2. FY3 DPR

The FY3 PMS will operate in a non-sun-synchronous circular orbit, and the flight altitude is about 400 km. The satellite will adopt a low orbit inclination angle in order to make frequent measurement of typhoon. The FY3 DPR consists of two active phased array radars, which are Ku-band precipitation radar (KuPR) and Ka-band precipitation radar (KaPR). The main objective of the FY3 DPR is achieving accurate measurement of Typhoon rain and other types of precipitation, such as stratiform rain, snow, etc. Key technologies for the radar, such as the slotted wave guide antenna with low sidelobe level, the real-time digital signal processor and so on, are being developed by Beijing Research Institute of Telemetry and Beijing Institute of Remote Sensing and Equipment.

Table 1 gives the main expected performance parameters for the FY3 DPR. The KuPR and KaPR both execute $\pm 20^\circ$ across track scan, and the corresponding swath width is about 300 km. The range resolution is 250 m, which is implemented by transmitting 1.6 μs pulse. The observation range is 18 km in height, and mirror image observation will be done near the nadir. The minimum detectable rain rate of the KuPR is 0.5 mm/h. For the sake of measuring snow and light rain, the KaPR has higher detecting sensitivity (0.2 mm/h). No less than 64 independent samples in each resolution volume are needed for maximizing the signal-to-noise ratio (SNR).

Table 1: Key expected performance parameters for the FY3 DPR.

	KuPR	KaPR
Frequency	13.6 GHz	35.5 GHz
Swath width	300 km	
Horizontal resolution	5 km (at nadir)	
Range resolution	250 m	
Observation range	18 km ~ -5 km ASL	
Minimum detectable rain rate	0.5 mm/h	0.2 mm/h
Measurement accuracy	$\leq \pm 1$ dB	
Number of independent samples	≥ 64	

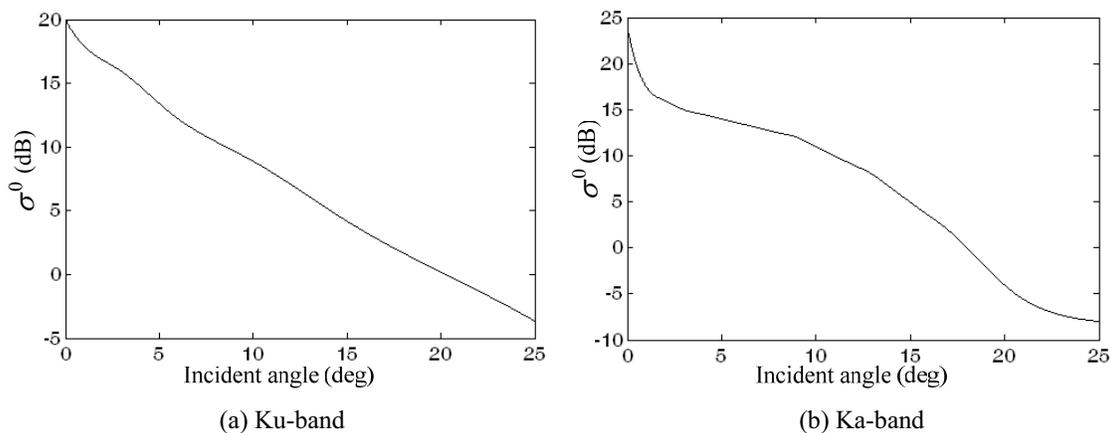


Figure 1: Sea surface backscattering model.

3. SEA CLUTTER MODLING AND SIMULATION

It is well known that microwave backscattering of the ocean is determined by both sea surface conditions and the observation geometry. Since it is hard to use a simple model to define the sea surface scattering, here the backscattering coefficients from practical measuring data are used, and only the strongest backscattering at different incidence angles are picked out to study the case in which the interference of sea clutter is most severe for determining the requirement on the antenna radiation pattern of the FY3 DPR. We use the sea scattering data measured by the precipitation radar (PR) onboard the Tropical Rainfall Measuring Mission (TRMM) satellite as the sea surface backscattering coefficients of Ku-band. The data set acquired at AAFE RADSCAT flight experiment is used while the scan angle is bigger than the maximum scan angle (17°) of the PR [3]. The backscattering coefficients of Ka-band are taken from the experiment results of Grant and Yaplee [4]. The cubic spline interpolation is applied to all these data in order to achieve the sea surface normalized radar backscattering sections σ^0 within the whole scan range, shown in Fig. 1.

A simple precipitation model is utilized to simplify the analysis, i.e., a uniform rain field extends from the sea surface to the 5 km altitude and there is a 0.5 km thickness bright band at the rain top, attenuation of which is twice as that of other rain region. In order to study the effect of the most severe sea clutter interference, the intensity of the rain field is assumed to be the minimum detectable rain rate of the FY3 DPR. The relationship between precipitation and the radar measuring parameters used here is expressed as [5, 6]

$$\eta = 4.7357 \times 10^{-4} I^{1.54}, \quad k = 0.032 I^{1.124} \quad \text{KuPR} \quad (1)$$

$$\eta = 1.3276 \times 10^{-2} I^{1.33}, \quad k = 0.215 I^{1.07} \quad \text{KaPR} \quad (2)$$

where η is the radar reflectivity (1/km), I is the rain rate (mm/h), and k is the attenuation coefficient (dB/km).

The received power from precipitation P_r can be calculated by integrating the radar equation

over the effective radar beam resolution volume V

$$P_r = \frac{P_t G_0^2 \lambda^2}{(4\pi)^3} \eta \int_{-L/2}^{L/2} \frac{A_r G_r(l)}{(R_0 + l)^2} dl \int_{\Omega} G_a^2(\theta, \phi) d\Omega \quad (3)$$

where P_t is the transmitted power, G_0 is the antenna gain at boresight direction, λ is the wavelength, A_r is the path attenuation of rain echo, R_0 is the distance from radar to the rain cell V , l is the distance between calculation point to the midpoint of V , $G_a(\theta, \phi)$ is the normalized antenna pattern, Ω is the corresponding solid angle of V , and $G_r(l)$ is the normalized range weighting function. In the case of non-pulse-compression radar like the FY3 DPR, the value of G_r is a constant.

The received signal power from sea clutter P_s can be obtained by integrating the radar equation over the sea surface area S at the same range gate as that of V ,

$$P_s = \frac{P_t \lambda^2 G_0^2}{(4\pi)^3} \iint_s \frac{G_a^2 - a(\theta, \phi) G_r}{h_s^2 \sec \theta} A_s \sigma^0 \sin \theta d\theta d\phi \quad (4)$$

where h_s is the satellite altitude, A_s is the path attenuation of sea clutter, θ is the incidence angle, and ϕ is the azimuth angle.

The effect of sea clutter on the antenna half power beamwidth of the FY3 DPR will be analyzed firstly. In the following numerical simulation, the signal to clutter power ratio (SCR, i.e., P_r/P_s) is calculated at every 10 meter along the main beam direction. Here the antenna peak sidelobe level of KuPR is assumed to equal to -35 dB. The altitudes which the sea clutter can interfere severely (i.e., $P_r < P_s$) with different antenna beamwidth (or horizontal resolution) are shown in Table 2. Normally the thickness of the thinnest stratiform precipitation is only about 3 km, so the horizontal resolution of the FY3 DPR must be not less than 6 km. But higher resolution means bigger antenna. Finally, 5 km horizontal resolution at nadir is chosen for the FY3 DPR because of the dimension limitation of the satellite platform, the corresponding antenna half power beamwidth is about 0.71° .

Next we will examine the interference of sea clutter for the FY3 KuPR and KaPR with different peak radiation sidelobe levels (SLL) in the case of which SLL is changed by adopting different weighting coefficients for exciting current. These results are presented in Fig. 2 and Fig. 3, where the abscissa represents antenna scan angle of the FY3 DPR and the ordinate represents the distance from calculating point to the sea surface along the antenna boresight. The colorbars at the right side of Fig. 2(a) and Fig. 3(a) indicate the mapping of the SCR values in dB to color for the KuPR and KaPR respectively. Here we consider that if the SCR is greater than 0 dB then the precipitation can be measured effectively by the FY3 DPR. It can be seen from Fig. 2(b) and Fig. 3(b) that those regions above the 0 dB contour almost do not suffer from the interference of sea clutter.

Table 2: Antenna beamwidth dependence of altitudes which sea clutter can interfere severely.

Horizontal resolution at nadir	5 km	6 km	7 km
Corresponding antenna beamwidth	0.71°	0.86°	1.00°
Altitude of sea clutter interference	2.38 km	2.82 km	3.18 km

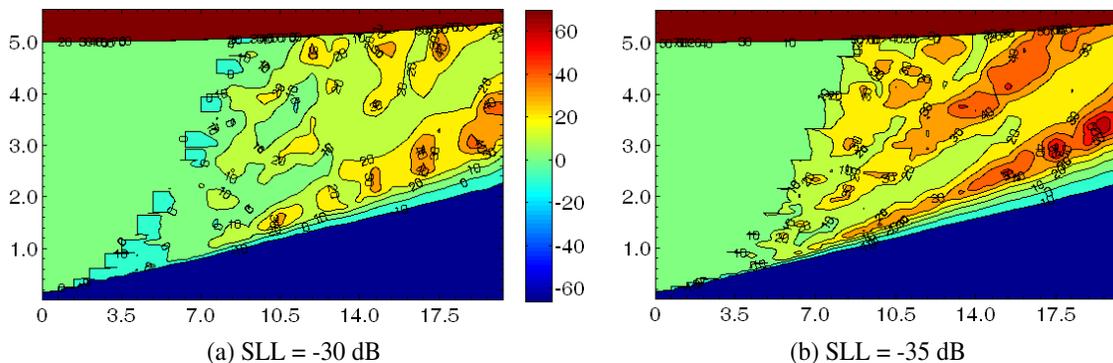


Figure 2: SCR with different antenna sidelobe levels for KuPR.

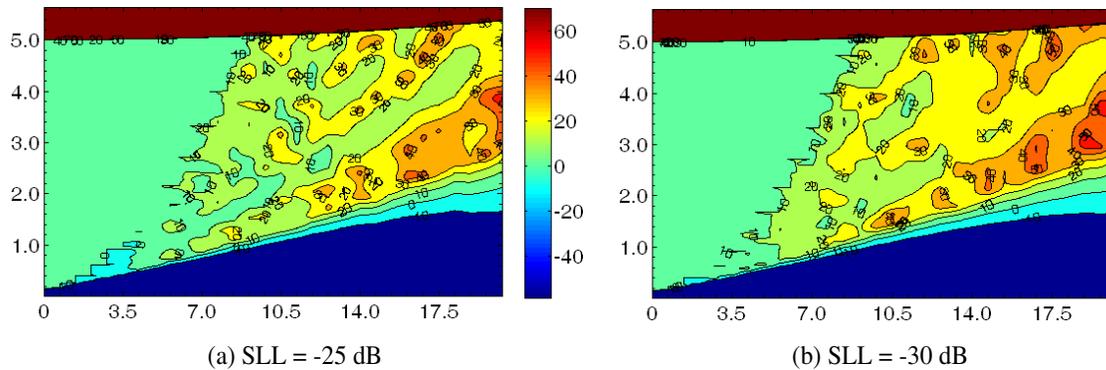


Figure 3: SCR with different antenna sidelobe levels for KaPR.

4. CONCLUSION

The DPR that will be installed on the FY3 PMS is being developed. The main objective of the FY3 DPR is to measure the three dimensional structure of precipitation, especially the Typhoon rain. The effects of sea clutter on rain measurement from the FY3 DPR have been analyzed in order to determine the requirements for the antenna radiation pattern. The power ratios of rain echo to sea clutter have been calculated. From the results of the numerical simulation, it is concluded that the performance parameters requirements of 0.71° antenna beamwidth, -35 dB radiation sidelobe level for the KuPR and -30 dB for the KaPR are needed to ensure to discriminate little rain from sea clutter with the FY3 DPR.

REFERENCES

1. Wang, X. L., Y. M. Wang, F. M. Ren, and W. J. Li, "Interdecadal variations in frequencies of typhoons affecting China during 1951–2004," *Advances in Climate Change Research*, Vol. 2, No. 3, 135–139, 2006.
2. Sato, K., H. Horie, H. Hanado, and H. Kumagai, "A digital-analog hybrid technique for low range sidelobe pulse compression," *IEEE Trans. Geoscience and Remote Sensing*, Vol. 39, No. 7, 1612–1615, 2001.
3. Schroeder, L. C., P. R. Schaffner, J. L. Mitchell, and W. L. Jones, "AAFE RADSCAT 13.9-GHz measurements and analysis: Wind-speed signature of the ocean," *IEEE Journal of Oceanic Engineering*, Vol. 10, No. 4, 346–357, 1985.
4. Grant, C. R., and B. S. Yapple, "Back scattering from water and land at centimeter and millimeter," *Proceedings of the IRE*, Vol. 45, 976–982, 1957.
5. Manabe, T., K. Okamoto, and T. Ihara, "A feasibility study of rain radar for the tropical rainfall measuring mission, 5: Effects of surface clutter on rain measurements from satellite," *Journal of Communication Research Laboratory*, Vol. 35, No. 145, 163–181, 1988.
6. Mega, T., H. Hanado, K. Okamoto, and R. Ushio, "Rain parameters calculated from raindrop size distribution for the design of future spaceborne precipitation radar," *URSI XXVII General Assembly*, [s. n.], 2002.

A Millimeter-wave Interferometric Radiometer for Atmosphere Observation from Geostationary Orbit

Ailan Lan, Shengwei Zhang, Hao Liu, Jingye Yan, and Ji Wu

Center for Space Science and Applied Research, Chinese Academy of Sciences, China

Abstract— In this paper, the advantages of a microwave interferometric sounder for atmospheric observation from geostationary orbit are introduced firstly. Then the major system parameters and application indices of the system are listed and the system design scheme is outlined. Furthermore, the four parts of the system, antenna, mm-wave receiver, wide band quadrature IF receiver and high speed digital unit, are stated in detail. At the same time, the difficulties needed to be solved further are pointed out.

1. INTRODUCTION

Meteorological satellite is primarily used to monitor and forecast strong convective weather, such as typhoon and storm. Comparing with geostationary meteorological satellite, polar-orbiting one has longer revisit time. The revisit time is not less than 6 hours, approximating to the typical life cycle of strong convective weather, even if two pole-orbiting meteorological satellites are used at the same time. Whereas geostationary satellite has characteristics of wider spatial coverage and timeliness, which make it possible to monitor the dynamic process of convective weather.

On the other hand, geostationary satellite has higher orbit. In order to achieving needed spatial resolution, a large antenna aperture and a complicated mechanical scanner must be used for a real aperture microwave sounder. It is comparatively difficult to realize on a satellite platform in terms of technique at present. However, this difficulty can be avoided by a interferometric sounder through using array antenna.

In addition, clouds are almost completely opaque at infrared and visible light wavelengths. Sounders working at these wavelengths can only obtain the information in cloud free areas and in the less important upper atmosphere above the cloud tops [1]. Consequently, the sounders can hardly get the vertical distribution of temperature and humidity in the troposphere under cloudy condition. In comparison with infrared and visible light, microwave has the advantage of all-weather observation. The microwave sounder is possible to penetrate cloud and obtain the vertical profiles of temperature and humidity in cloud.

Therefore, using a microwave interferometric sounder to observe atmospheric temperature from geostationary orbit can improve the detecting ability of meteorological satellite.

2. SYSTEM DESCRIPTION

The millimeter-wave interferometric radiometer for atmospheric observation from geostationary orbit, that is geostationary interferometric microwave sounder (GIMS), under development by Center for Space Science and Applied Research CAS, will be used to obtain the temperature and humidity profiles of cloud to monitor and forecast strong convective weather. The major system parameters of hardware and performance indices are listed in Table 1. According to the channels used in AMSU-A and the analysis from CMA, operating frequencies of GIMS are set as Table 2. Channel 1, at 50.3 GHz, is particularly sensitive to surface features variation, while other channels are sensitive to the atmospheric constituents at higher altitudes. Therefore, channel 1 and other channels will be used respectively to obtain the surface emissivity and to monitor air temperature at different heights in the atmosphere [2].

3. SYSTEM DESIGN

GIMS is a 2-D passive remote sensor. Its hardware can be divided into four parts according to different functions: antenna array, millimeter-wave receiver, wide band quadrature IF receiver and high speed digital unit. The input of the system is millimeter-wave noise signal and the output is visible function.

Table 1: System parameters and performance indices.

System Parameters of hardware	Frequency	50–60 GHz
	Channel number	8
	Bandwidth	≤ 320 MHz
	Synthetic aperture	≥ 2.5 m
	Minimum length of baseline	≤ 6 wavelengths
	Imaging time	≤ 5 min
	Averaged data rate	≤ 1 Mbps
Performance indices	Radiometric sensitivity	< 1 K@25 ms
	Spatial resolution	50 km (at the height of 36000 km)
	Imaging area	3000 * 3000 km
	Calibration accuracy	≤ 1.5 K

Table 2: Operating frequencies of 8 channels.

Channel	Centre Frequency (GHz)	Bandwidth (MHz)
1	50.30	320
2	51.76	320
3	52.80	320
4	53.29	320
5	53.84	320
6	54.40	320
7	54.94	320
8	55.50	320

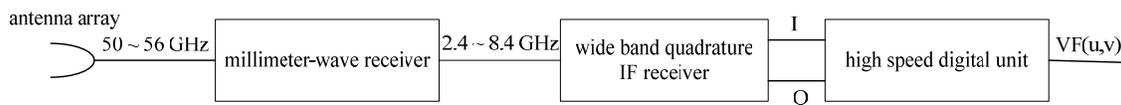


Figure 1: The diagram of GIMS.

3.1. Array Structure

In GIMS, a rotating circular thinned array is used, instead of a stationary Y-shape array, to reduce the required number of the antenna elements [3]. The array gets more spatial frequency sampling by moving the whole array or each element antenna respectively, and spatial frequency domain can be covered completely in one period. Comparing with stationary thinned array, this operation mode can improve spatial resolution with less elements of antenna and receiver. At the same time, it makes system a little more complex. In one word, time division sampling mode is a compromise between the complexity and cost of system.

In the system, 22 corrugated horn antenna elements are used. The structure of the array and the relative positions of each antenna are shown in Fig. 2. By the rotation of the disk, more baselines are obtained. The normalized length distribution of the baselines and spatial frequencies coverage are shown in Fig. 3 and Fig. 4 respectively. Fig. 3 illustrates that, for array structure shown as Fig. 2, baseline length distribution is approximately uniform. It can be seen from Fig. 4 that the spatial frequencies are like polar coordinate distribution, which makes conventional inversion algorithm, such as FFT or CLEAN algorithm, not suitable any longer. In addition, how to obtain accurate calibration data for this system is another difficulty need to be solved.

3.2. Millimeter-wave Receiver

In the system, 22 antennae will be used. Correspondingly, the same number of millimeter-wave receivers must be used. In order to keep the consistency of 22 millimeter-wave receivers, the design and assembly flow of each receiver must be identical.

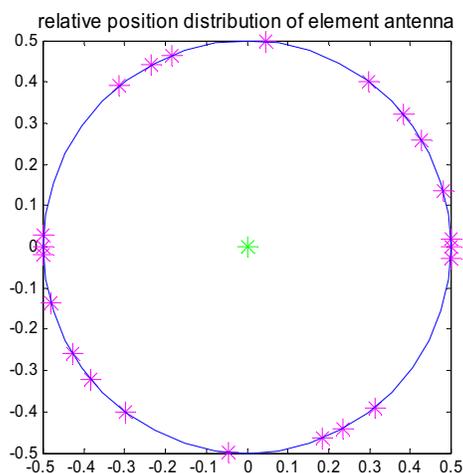


Figure 2: Array structure of GIMS.

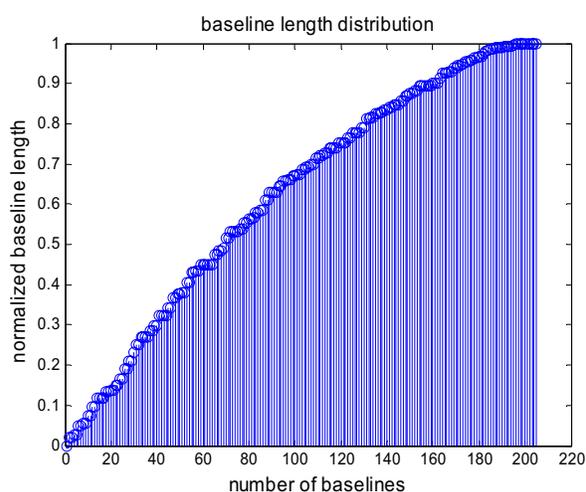


Figure 3: Baseline length distribution for the array shown in Fig. 2.

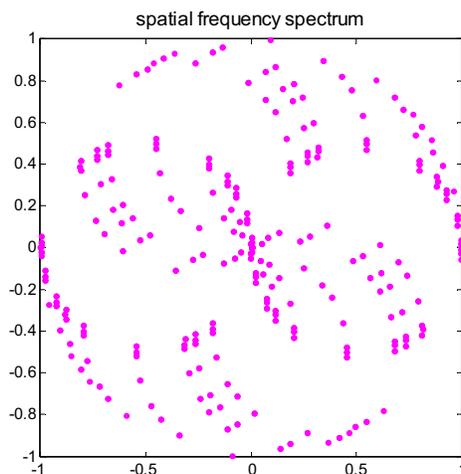


Figure 4: Spatial frequencies coverage for the array shown in Fig. 2.

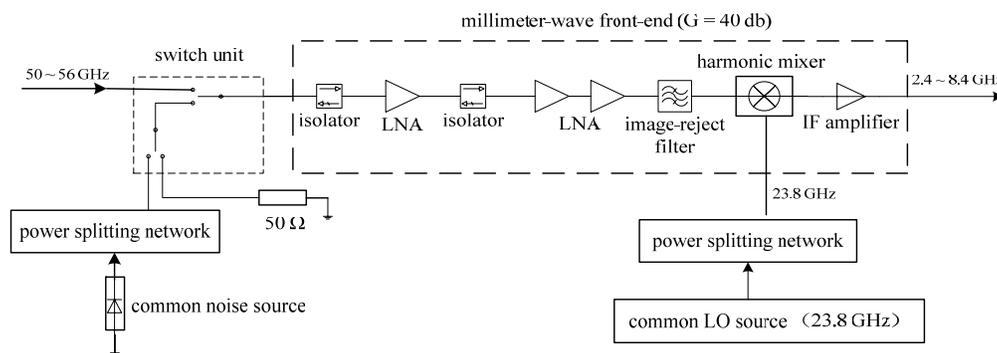


Figure 5: The diagram of a millimeter-wave receiver.

Figure 5 gives the diagram of one millimeter-wave receiver. It is composed mainly of a millimeter-wave front-end, a LO source, switches and noise sources. The millimeter-wave front end is the core of this unit. It receives and amplifies millimeter-wave noise signal (from 50 GHz to 56 GHz) and down-converts the signal into IF signal (from 2.4 GHz to 8.4 GHz). In the system, the gain of the front-end is required to be no less than 40 dB. Note that the mixer used in the front-end is a secondary harmonic mixer and the centre frequency of LO is 23.8 GHz.

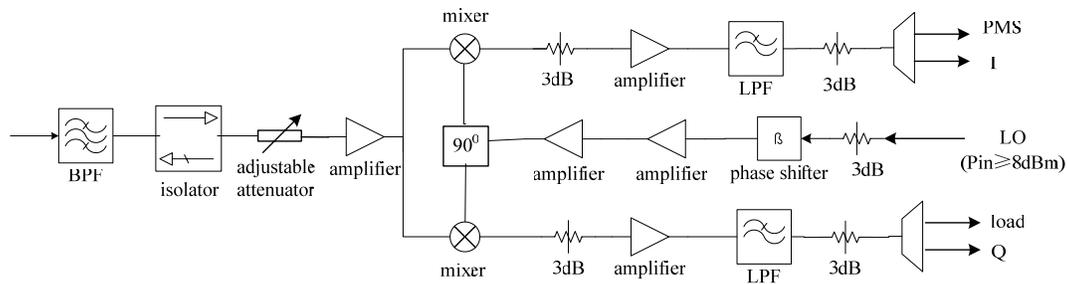


Figure 6: The diagram of coherent demodulator.

3.3. Wide Band Quadrature IF Receiver

As millimeter-wave receivers, 22 wide band quadrature IF receivers are also needed in the system. The IF receiver is used to pick up useful signals with required bandwidths from the input IF signal and convert the signals into quadrature LF signals by coherent demodulation. Consequently, coherent demodulator is the chief module. Although all of the 8 channels have the same bandwidth, the function of channel 1 is different from other channels and some of its performance indices are required differently. Therefore, two demodulators with same structure are needed. One has stationary LO frequency to get signal from channel 1, and the other must change its LO frequencies to obtain signals from other 7 channels. The diagram of coherent demodulator is shown as Fig. 6.

For this unit, the difficulty is to keep the consistency of amplitude frequency responses among different receivers and in pass band for a receiver. To solve this problem, high stability and consistency of common LO splitting network are required.

3.4. High Speed Digital Unit

High speed digital unit consists of data sampling modules whose sampling rate is up to 1 Gbps, a data processing module and a synchronous timing module. Each receiver needs a data sampling module to sample signals from all channels synchronously and the synchronous timing module is used to assure sampling synchronism. The sampled data, including I/Q data, consistency calibration data for amplitude and phase between channels, status data, temperature data and angle data, are transmitted to the processing module through optical fiber. In the processing module, in-phase data and quadrature data are correlated to obtain visible functions.

GIMS is a multichannel system with 22 receivers and 8 channels. It needs considerable amount of digital correlators to implement correlated process at a very high speed, which causes the scale of the digital unit very large.

4. CONCLUSIONS

At present, interferometric sounder is not used as widespread as real aperture sounder. However, the advantages of interferometric sounder are obvious: small aperture antenna and high resolution. The two characteristics are very significant for a space-borne sounder observing from geostationary orbit.

The prototype of GIMS is being developed. Most of the critical techniques are solved, but there are still several unsolved difficulties which will impact the resolution and accuracy of the system. Therefore, we are focusing our efforts on solving these difficulties in order to build a practical spaceborne system in the future.

REFERENCES

1. Lambrigtsen, B., W. Wilson, A. Tanner, et al., "GeoSTAR-a microwave sounder for geostationary satellite," *2004 IEEE International Geoscience and Remote Sensing Symposium Proceedings, IGARSS'04*, Vol. 2, 777–780, 2004.
2. L.M.D (France), "AMSU-A channel interpretation," http://web.lmd.jussieu.fr/~falmd/TP/results_interpret_AMSU/ans_interp_AMSU.pdf.
3. Liu, H., J. Wu, S. Zhang, J. Yan, C. Zhang, W. Sun, and L. Niu, "Conceptual design and breadboarding activities of geostationary interferometric microwave sounder (GIMS)," *The Proceeding of IGARSS*, 2009.

Forward Modeling of Direct Current Method Based on ANSYS

Dong-Feng Zhang

Institute of Information Engineering Physics, Central South University, Changsha, Hunan 410083, China

Abstract— In this paper, firstly we have reviewed the developing of numerical methods used in geophysical electric prospecting, and have listed the current state and developing trend of that, then simply have introduced the main characteristics of ANSYS and also mainly have focused on the steps and key problems of using ANSYS to model numerically DC forward problems; Secondly, based on the differential control equations of voltage potential of point current field, we have introduced the boundary value problem and the variational problem (involved in 2 dimension and 3 dimension); based on secondary development of APDL language raised in ANSYS, we have modeled lots of geophysical models, and we have got same perfect results by analyzing the results of numerical modeling; Lastly, we have achieved the proof that is the ANSYS can be used in geophysical electric modeling with high accuracies.

At the end, we have presented the suggestion that ANSYS can be widely used in geophysical forward modeling with valuable applications.

1. THIS PAPER USED IN THE GEOELECTRIC MODEL

The Earth's radius is about 6400 kilometers, the general electrical exploration of the emission source and observation points, spacing is usually not more than ten kilometers away, in this case that the ground is flat, do not take into account the curvature of the earth, in the discussion of geoscience in most of the basic electromagnetic field problems, do the following assumptions [2]:

- earth partition medium ν is uniform, isotropic, linear conductive medium;
- complex earth model can ν be composed of such a local area;
- electrical properties of the medium with ν time, temperature and pressure has nothing to do;
- medium permeability $\mu\nu$ and the permeability of free space μ_0 the same, namely, $\mu = \mu_0$.

Line current source of ANSYS simulation.

The potential for the line source should meet the differential equation [4] as follows:

$$\frac{\partial}{\partial x} \left(\sigma \frac{\partial u}{\partial x} \right) + \frac{\partial}{\partial z} \left(\sigma \frac{\partial u}{\partial z} \right) = -I \delta(x_A) \delta(z_A)$$

Assumptions on the type of geological body y -axis parallel to the direction of the source and the line direction and the direction of geological body, and therefore there is no y variable equation. I represent the unit line source for the length of the current intensity.

Homogeneous half-space potential distribution in winding current formula is:

$$U = -\frac{\rho I}{\pi} \ln R_A + C$$

2. POINT CURRENT SOURCE OF THE ANSYS THREE-DIMENSIONAL FORWARD MODEL

2.1. Model Calculation

Model shown in the figure, even a semi-infinite space, there is a radius of a sphere, the sphere resistivity, depth to the background resistivity, double-electrode power supply (current 10 A), power supply electrode A at -30 m, for the electrode B at 30 m. Following the results of graphics for different depths of apparent resistivity curves.

On the below model (Fig. 1), by calculating the following formula described in the application of theoretical value and the use of ANSYS simulation results for comparison, revealing at the formula near the surface of the irrational, because it is for the air — the interface of the earth, it's just a approximation to potential use will be extremely simple way of doubling the calculations, the following calculation was based on the sphere at different depths were cases of apparent resistivity anomaly curve.

$$U_0 = \frac{\rho_0 I}{2\pi} \left[\frac{1}{R} + 2 \sum_{n=0}^{\infty} \frac{n(\rho_1 - \rho_0)}{n\rho_0 + (n+1)\rho_1} \cdot \frac{r_0^{2n+1}}{D^{n+1}} \cdot \frac{1}{r^{n+1}} P_n(\cos \theta) \right]$$

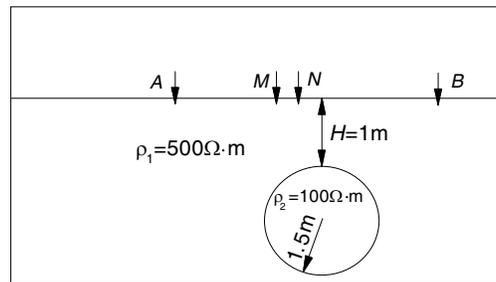


Figure 1: Sketch Models.

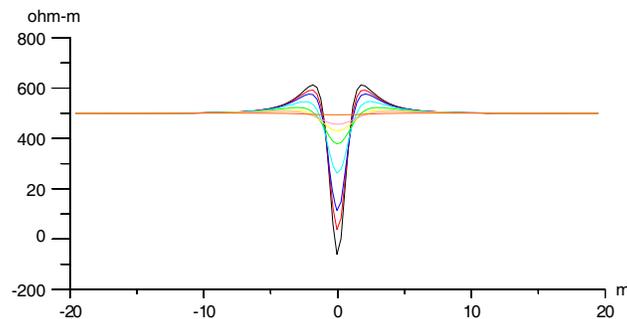


Figure 2: Balls at different depths of apparent resistivity curve of the theoretical value ($h = 0, 0.1, 0.2, 0.5, 1, 1.5, 2, 5 \text{ m}$)

From the above figure (Fig. 2) we can see that when h is less than 0.5 m, the formula (4-2) the calculated result is unreasonable application of ANSYS calculated results in line with the actual situation, when h is greater than equal to 0.5 m When the two curves in good agreement, which shows the formula (4-2) only if an exception when the ball farther from the ground is reasonable, and thus illustrates the ANSYS three-dimensional numerical simulation accuracy. Of course, on the other hand reflects the gradient in the middle of the ball guide device found good potential to detect the depth of ore bodies, under normal circumstances for.

3. MODEL 2

Model 2, respectively, some upright for underground low resistance and high resistance plate body, top of the depth of 2 m, rock resistivity $100 \text{ } \Omega\text{m}$, low resistivity body resistivity $30 \text{ } \Omega\text{m}$, high resistance body resistivity $300 \text{ } \Omega\text{m}$, with the dipole — even pole-pole measurements, $AB = MN = 2 \text{ m}$, numerical simulation results shown in Figures 3 and 4.

As can be seen from the figure, vertical plate-like body morphology abnormalities, low resistance anomaly board is closed, while the high resistance anomaly board was “eight” shape, unusual for the semi-enclosed, and the location of abnormal high resistance semi-enclosed This is because they do not coincide with the polar distance increases, high resistance value of outward displacement of the great peaks arising addition, abnormal body vertical plate and horizontal plate-like body shape abnormalities that just the opposite, this is an exception The “orthogonal characteristics”, the substance of this feature is for the underground by the current flow from the anode cathode path with low resistance to bypass the high resistance, and as far as possible by taking the shortest

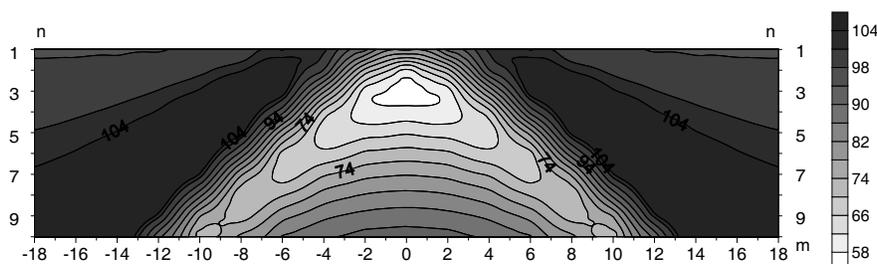


Figure 3: Vertical low resistivity forward modeling results of plate-like body plan.

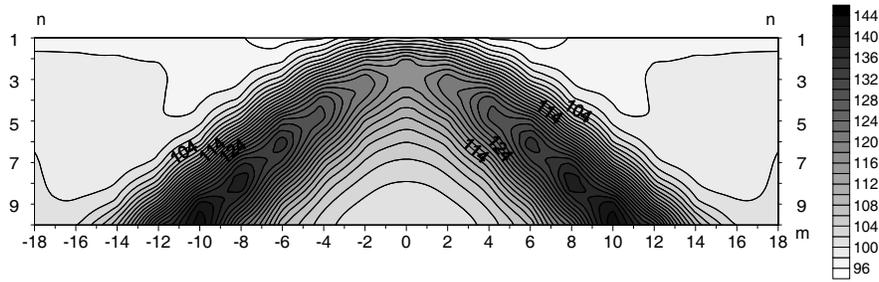


Figure 4: Vertical high resistance plate body forward modeling results of Figure.

distance characteristics of the decision.

4. CONCLUSION

Carried out through the application of a large number of DC ANSYS finite element method numerical calculation, we can draw the following conclusions: the finite element method is a more powerful numerical simulation method, which is suitable for electric model of a complex geophysical problems, finite element software ANSYS powerful, it can be applied to many aspects of geophysics: bit-field continuation, gravity and magnetic forward, DC law forward and magnetotelluric forward modeling and so on.

REFERENCES

1. Coggon, J. H., "Electromagnetic and electrical modeling by the finite element method," *Geophysics*, Vol. 36, 132–155, 1971.
2. Silvester, P. and C. R. S. Haslam, "Magnetotelluric modeling by the finite element method," *Prosp.*, Vol. 20, 872–891, 1972.
3. Rijo, L., "Modeling of electric and electromagnetic data," Ph.D. Thesis. Univ. of Utah., 1977.
4. Kaikkonen, P., "Numerical VLF modeling," *Geophysical Prospecting*, Vol. 27, 815–834, 1979.

Micro-motion Simulation and Micro-Doppler Extraction

Ning Chao and Huang Jing

National Key Lab. of Target and Environmental Electromagnetic Scattering and Radiation
Beijing 100854, China

Abstract— A model which simulates micro-motion is measured in the anechoic chamber. And the measurement comprises Two types of data: Data measured by narrow-band radar and data measured by wide-band radar. Some improvement on a micro-motion periodicity estimation algorithm is presented. The general joint-time-frequency (T-F) transforms are introduced and compared in the paper. A detailed description is given of a new method for confirming the dwelling time of short time Fourier transform (STFT). Using the T-F transforms and Hough transform, some micro-Doppler characteristics are extracted. The results demonstrate that a radar has the capability to measure micro-motion and obtain the micro-motion characteristics in some specific condition.

1. INTRODUCTION

If a target or any structure on the target has mechanical vibration or rotation in addition to its major translation, such as rotors of a helicopter, or the engine compressor, it may result in a frequency modulation on the returned signal which generates sidebands around the target's Doppler frequency shift. This is called the micro-Doppler effect. The micro-Doppler reflects some dynamic behavior. It can be used in target recognition [1, 2]. The micro-motion measurements in radar are still rare, in despite of many simulations with electromagnetic calculation theory. In this research, a metal cone model is used to simulate micro-motion and measured in anechoic chamber. Some micro-motion characteristics are obtained from the experiment data. The digital signal processing algorithms are also discussed in this paper.

2. THE MICRO-MOTION MODEL

The model is a metal cone with a spherical cap which is used to simulate micro-motion, shown as Fig. 1. The model is mounted on a metal bracket. There are two motors inside, controlling the spinning rotation and coning rotation respectively. The model rotates like a peg-top, as well as the rotations speed and the coning angle are adjustable. The bracket is covered by microwave absorb material in order to get less additional back waves. shown as Fig. 1.

3. MEASUREMENT OF THE MICRO-MOTION MODEL

The measurement of micro-motion model is completed in the anechoic chamber of National Key Laboratory of Target and Environmental Electromagnetic Scattering and Radiation. Two kinds of radar are in application: A narrow-band radar and a wide-band radar. We just set the model in coning rotation because the spinning rotation did less contribution to the radar back waves. These movements of rotation are illustrated in Fig. 1.

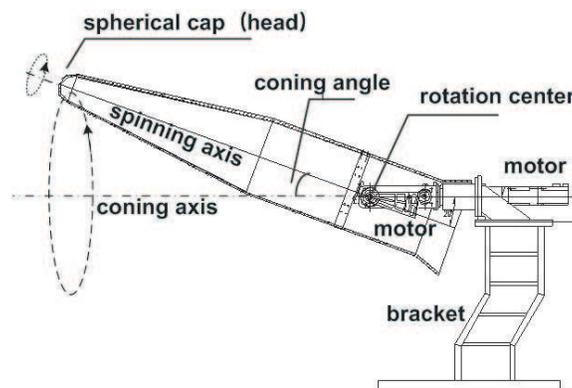


Figure 1: The micro-motion model.

4. MICRO-MOTION PERIODICITY ESTIMATION IN TIME DOMAIN

The testing radar is a narrow-band coherent radar. The magnitude and phase information of the back waves are contained in I channel and Q channels in time domain

The distance from the radar to the target, the azimuth and elevation angle are changed when the micro-motion model in coning rotation. This will induce the changing of reflective waves energy which are represented by RCS sequence in time domain. Then the coning periodicity could be estimated from its back waves. There are several methods of calculating periodicity, including autocorrelation algorithm, spectrum algorithm. The method on the combination of circle autocorrelation (CAUTO) and circle average magnitude difference function (CAMDF) is a simple, accurate and high practical algorithm [3]. The main formulas are:

$$R(k) = \sum_{n=1}^N x(n) \cdot x(\text{mod}(n+k, N)) \quad (1)$$

$$D(k) = \sum_{n=1}^N |x(\text{mod}(n+k, N)) - x(n)| \quad (2)$$

$$\phi(k) = R(k)/D(k) \quad (3)$$

where $k = 1, 2 \dots N$, $x(n)$ is the analyzed data, $\text{mod}(x, y)$ denotes Modulus after division (x/y), $R(n)$ is the result of CAUTO, $D(n)$ is the result of CAMDF and $\phi(n)$ is the final result.

The position of the maximal extremum in $\phi(n)$ must be the first index and the position of the second great extremum corresponds with the coning periodicity shown in Fig. 2.

Large calculations are required in the algorithm CAUTO + CAMDF. In order to calculate faster, We modify the algorithm as follow:

1. Since $R(n)$, $x(n)$, $\phi(n)$ are all eudipleural, then the calculation times can be $N/2$.
2. The coning periodicity is usually long because the A/D rate is much higher than the frequency of rotation, then moving step length of k , denoted as Δk , is unnecessary to be one data, it can be longer and the calculation times become shorter.

With the signal spectrum, the proper moving step length of k will be confirmed. Firstly, draw the spectrogram. Secondly, calculate the bandwidth B_W with more than 90% energy in spectrum. Then Δk is confirmed using the B_W . If $\Delta k = (1/10)B_W^{-1}$, the bias of periodicity estimation is 10% and Δk may be shorter in order to guarantee smaller bias.

In our test, the estimated periodicity is 2s with the original CAUTO + CAMDF method, the calculating time consumption is about 380 ms. The estimated periodicity is 1.94 s (bias is less than 10%) with the modified method, while the time consumption of calculating is only about 50 ms. The bias is acceptable. Though the bias is a little larger, the time consumption of calculating is much shorter. The result comparison of the two methods is shown in Fig. 2.

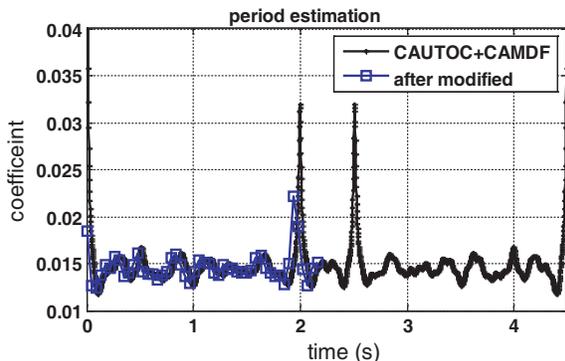


Figure 2: Periodicity estimation with back waves.

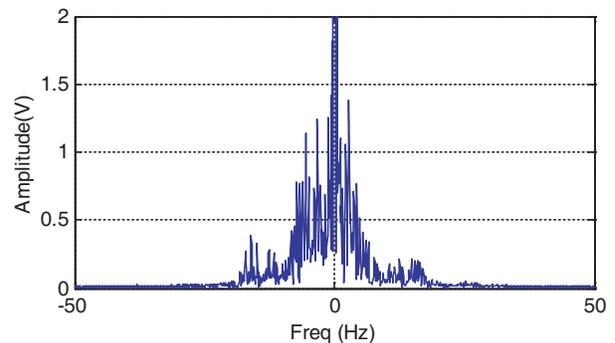


Figure 3: Spectrum of back waves.

5. TIME-FREQUENCY TRANSFORM

Time-Frequency (T-F) transforms are strong tools to analyze the time-varying frequency characteristics of the micro-Doppler modulation. STFT is defined as

$$F_x(t, \omega) = \int_{-\infty}^{+\infty} x(\tau)h(t)e^{-j\omega\tau} d\tau \tag{4}$$

where $h(t)$ is time window. The resolution of the STFT is determined by the window size. There is a trade-off between the time resolution and the frequency resolution. Denote the window size as T_W , the time resolution $\delta_t = T_W$, the frequency resolution $\delta_f = 1/T_W$ (in a rectangle window).

There are an effective method to confirm proper size of time window with the result of coning periodicity T and the maximum micro-Doppler frequency $f_{d\max}$, generally, $f_{d\max}$ is a half of B_W .

With Nyquist Sampling theorem, $(\frac{1}{2}f_{d\max})^{-1} < T_W < \frac{1}{2}T$, In order to have enough joint-time-frequency resolution and obtain more detail characteristics of cone rotation, the T_W will be $(\frac{1}{5}f_{d\max})^{-1} < T_W < \frac{1}{5}T$. In our experiment, the $f_{d\max}$ is about 35 Hz, the periodicity is about 2 s. Then, the window size $T_W = 300$ ms and complete STFT diagram which is shown as Fig. 4.

Wigner-Ville distribution (WVD) [4] is another T-F transform algorithm. The WVD of a signal $s(t)$ is defined as the Fourier transform of the time-dependent autocorrelation function:

$$\text{WVD}(t, \omega) = \int s\left(t + \frac{\tau}{2}\right) s^*\left(t - \frac{\tau}{2}\right) e^{-j\omega\tau} d\tau \tag{5}$$

WVD suffers from a problem of cross-term interference. To reduce the cross-term interference, the filtered WVD has been used to preserve the useful properties of the T-F transform with a slightly reduced joint-time-frequency resolution and largely reduced cross-term interference. Smooth Pseudo Wigner-Ville distribution (SPWVD) is one of the filtered WVD methods [4, 5], The diagram by SPWVD is shown as Fig. 5:

$$\text{SPWVD}(t, \omega) = \int h(\tau) \int g(\xi - t) R_s(\tau) e^{-j\omega\tau} d\xi d\tau \tag{6}$$

Another shortage of WVD is the time consumption of calculating. In a case, calculating the same data, SPWVD spends about 53.5 seconds., while STFT only spends about 0.4 second.

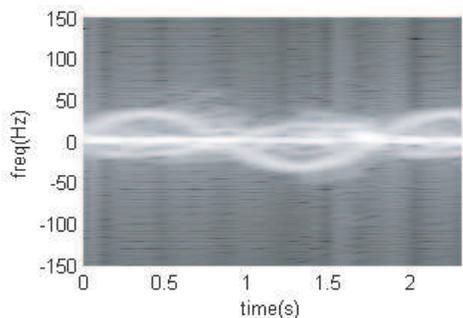


Figure 4: Results of STFT.

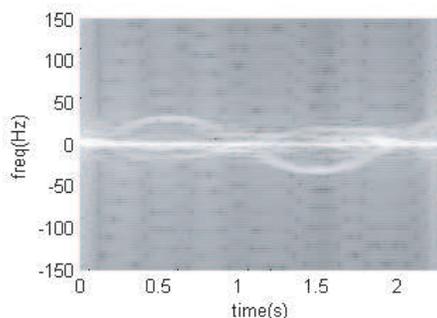


Figure 5: Results of SPWVD.

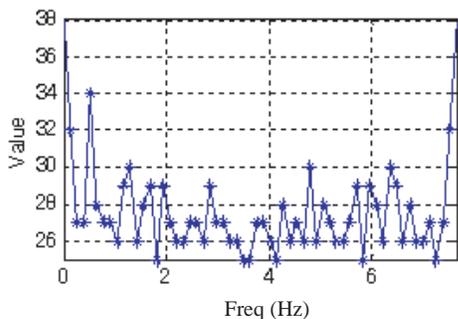


Figure 6: The accumulator array of f .

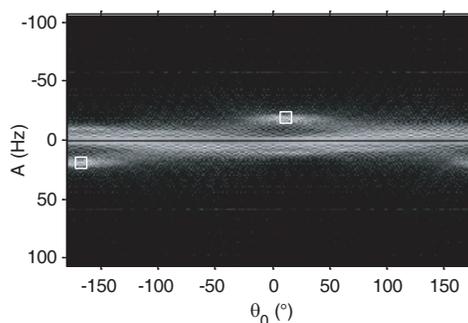


Figure 7: The accumulator array of (A, θ_0) .

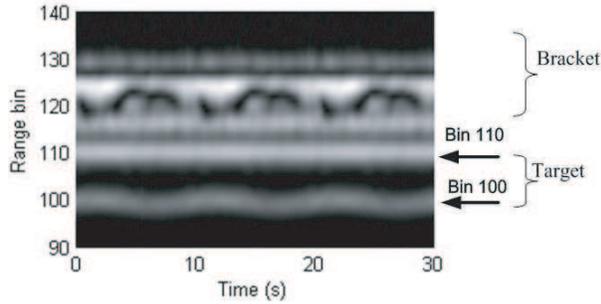


Figure 8: HRRP of micro-motion model in cone rotation.

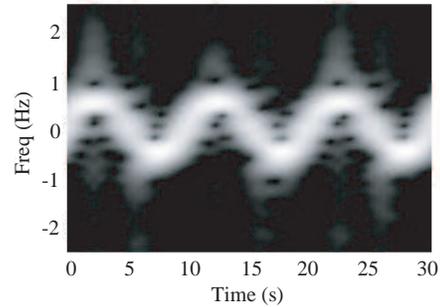


Figure 9: T-F diagram of range bin 100.

Although joint-time-frequency resolution of STFT is not exact enough, STFT has great advantage of speed and be used in the case of real time data processing, especially in Embedded System, such as in DSP, FPGA and Single-Chip Microcomputer.

6. MICRO-MOTION PERIODICITY ESTIMATION BY HOUGH TRANSFORM

Hough transform is a feature extraction technique used in image analysis. Hough transform is generally used to detect straight lines [6].

From Fig. 5 and Fig. 6, the micro-Doppler characteristic of spherical cap in the model mainly showed out as a sinusoid curve and the parameter of the curve could be estimated by Hough transform. A sinusoid curve is confirmed with four parameter and can be described as:

$$F = A \sin(2\pi ft + \theta_0) + F_0 \quad (7)$$

where A is amplitude, f is frequency; θ_0 is the original phase; F_0 is offset (here $F_0 = 0$).

The accumulator array are three dimensions. Firstly, we estimate f and $f = 0.5$ Hz as shown in Fig. 6. Then, (A, θ_0) could be found $A = 20$, $\theta_0 = 10^\circ$ or $A = -20$, $\theta_0 = 170^\circ$ (They are according with the same sinusoid curve) as shown in Fig. 7.

The physical meaning of f is the coning frequency and A is the maximum micro-Doppler frequency. The results are in agreement with the parameters we set before.

7. DATA ANALYSIS IN WIDE-BAND RADAR

After processing the wide-band data, obtain a group of one dimension high resolution range profiles (HRRP) with time shown as Fig. 8.

In HRRP, there are two greater scattering centers, one is the spherical cap of model which is at range bin 100, the other is the bottom edge which is at range bin 110. The bracket is the greatest center, it is farther than range bin 110.

With the I/Q data of the spherical cap, A T-F diagram are processed by STFT shown as Fig. 9. There are a obviously sinusoid curve and by Hough transform, we estimate its periodicity is about 10s. The result is accurate which is just coning periodicity.

8. CONCLUSION

This paper introduces the micro-Doppler experiment implemented by radar. Through analyzing the data, some characteristics of micro-motion are extracted from time domain and joint T-F domain. Our simulation also demonstrates the radar effectiveness in measuring micro-motion if enough dwelling time are given, including both narrow-band radar and wide-band radar. In addition, the periodicity estimation algorithm with back waves in time domain is improved and some applications of T-F transform and Hough transform is described.

ACKNOWLEDGMENT

The authors thank the Anechoic Chamber of Institute 207, CASIC for providing micro-Doppler experiment data.

REFERENCES

1. Ingwersen, P. A. and W. Z. Lemnios. "Radars for ballistic missile defense research," *The Lincoln Lab. Journal*, Vol. 12, No. 2, 245–266, 2000.

2. Chen, H., and Y. Liu, "Micro-motion recognize on radar," *Systems Engineering and Electronics*, Vol. 29, No. 3, 361–364, March 2007.
3. Feng, D. and J. Liu, "RCS periodicity of target and its estimation algorithm," *Journal of Astronautics*, Vol. 29, No. 1, 363–366, 2008.
4. Chen, V. C., F. Li, and S.-S. Ho, "Micro-Doppler effect in radar: phenomenon, model and simulation," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 42, No. 1, January 2006.
5. Hu, C. and S. Lou, "*System Design and Analysis Based on Matlab*, XiDian Press, 2002.
6. Barbarossa, S., "Analysis of multicomponent LFM signals by a combine wigner-hough transform," *IEEE Transactions on Signal Processing*, Vol. 43, No. 6, 1511–1515, June 1995.

An Integration of Electronic System and Some Solutions to Its Key Point

Yanhong Hao and Jiali Wang

School of Electromechanical Engineering, Xidian University, Xi'an 710071, China

Abstract— With the development of electronic technology, it becomes more and more complex and functional. It tends to an development of PHM system. The paper gives an integration of PHM (Prognostic and Health Management/Monitoring) system as a whole, and an elaborate description of it as well. Because every module of the PHM system is designed for special purpose, it becomes necessary to implement every part concretely, and proper algorithm may be in use. As the uncertainty of data source is one of its key point to PHM technology, the paper presents a BP network data fusion algorithm to reduce this uncertainty. It has voltage test pattern, temperature and energy as its sensors' inputs.

1. INTRODUCTION

As electronic equipments become more and more complicated, PHM (Prognostic and Health Management/Monitoring) technology has developed gradually. Because PHM is a newborn and interdisciplinary technology, especially in our country, it remains many problems to solve, such as uncertainty of data acquisition and differentiation of PHM system. America has PHM technology as its military and civil aviation maintenance technology since the year of 2000. The Joint Strike Fighter (JSF) Program has first developed a vision for and has specified a very stringent set of requirements for a comprehensive Prognostics and Health Management (PHM) system. These PHM capabilities are then integrated into the JSF Air System design. This vision and the associated specified requirements have resulted in the development of PHM integration. Along with it, HUMS (Health and Usage Management System), FUMS (Fleet and Usage Management System in UK) ACMS (Aircraft Condition Monitoring System), EMS (engine monitoring system) and IDPS (integrated diagnostics and prognostics system) come into being. Many foreign countries have already evolved their own PHM systems.

PHM technology is in the ascendant in our country. Thought it has begun in the 1980s, PHM is rather more uncertain deductions than certain calculation. This uncertainty comes from the uncertainty of data acquisition and the simple function of modules, lack of precise algorithms. With the development of electronic equipment, the PHM technology should keep its pace. COMPASS, ECM, ADEPT are some early PHM systems in civil aviation of our country. In 2006, QAR (Quick Access Recorder, a fault diagnostic and prognostic database) is imported to our country for PHM, but it still at its beginning. Compared with other countries in military and civilian, our country has a great urge to research efficient PHM system in deep.

2. INTEGRATION OF PHM SYSTEM

An integration of PHM system is an organic integration of functional modules, which exchange real-time data with each other under certain communication protocol. By processing the input data, it predicts the future condition of a system under test using prognostic algorithms and related methods. It outputs the fault points and fault models of a system. It also gives the estimation of a system's Remaining Useful Life (RUL) and some advice of its maintenance. These functions are practically necessary and efficient to condition-based maintenance (CBM). By this means, money and manpower can be reduced evidently. So it is the main way to equipment maintenance.

PHM system functional architecture can be shown as follows:

7 modules are integrated in PHM system. Every module has its own technique.

- (1) Data acquisition. It has active or passive sensors to collect different types of raw data from system under test, and sends them to data manipulation module. The big problem in this module is the uncertainty of data sources. The solution to this problem is to use data fusion algorithm following it to reduce this uncertainty.
- (2) Data manipulation. It converts the raw data from data acquisition module into the needed type of the following modules, such as A/D conversion or feature extraction.

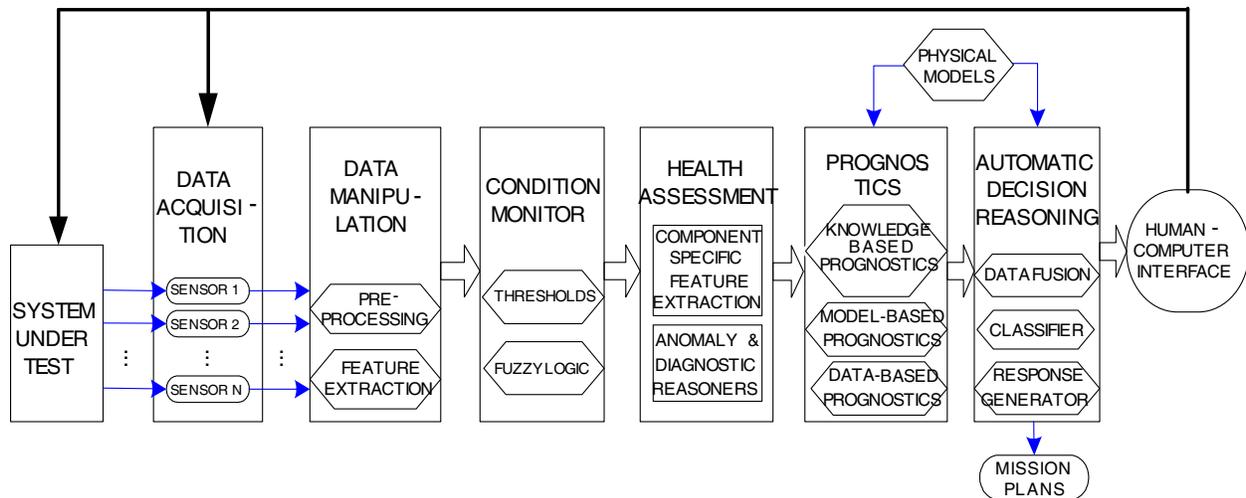


Figure 1: PHM System functional architecture.

- (3) Condition monitor. It set up some conditions as thresholds or fuzzy logic. These conditions are criterions for PHM system to give instructions whether the system is health or not.
- (4) Health assessment. This module includes component specific feature extraction, anomaly and diagnostic reasoners.
- (5) Prognostics. It is the core of PHM system. Some crucial algorithms are in it. The prognostic effect of a PHM system is mainly due to them. They can be divided into three categories: knowledge-based prognostics, model-bases prognostics and data-based prognostics.
- (6) Automatic decision reasoning. It is another important part of PHM system. It has some algorithms to draw a prognostic conclusion. Bayes decision rule is commonly used to data fusion. There are other algorithms such as neural network fusion algorithm. Fusion algorithms take an important role in prognostic result, so the selection of fusion algorithm is vital.
- (7) Human-computer interface. It integrates the output of every module and instructs system under test to adjust its operating state properly.

It should be noted that modules of PHM system not entirely distinguish one from another, and they may share same data or use feedback.

3. USING FUSION ALGORITHM TO REDUCE THE UNCERTAINTY OF DATA SOURCES

The prognostic accuracy of a PHM system depends dramatically on the accuracy of data source. Fusion algorithm is the key to solve this problem. Nowadays Bayes decision rule, Dempster-Shafer theory and neural network are widely in use. A multi-sensor data fusion method for PHM is as follows:

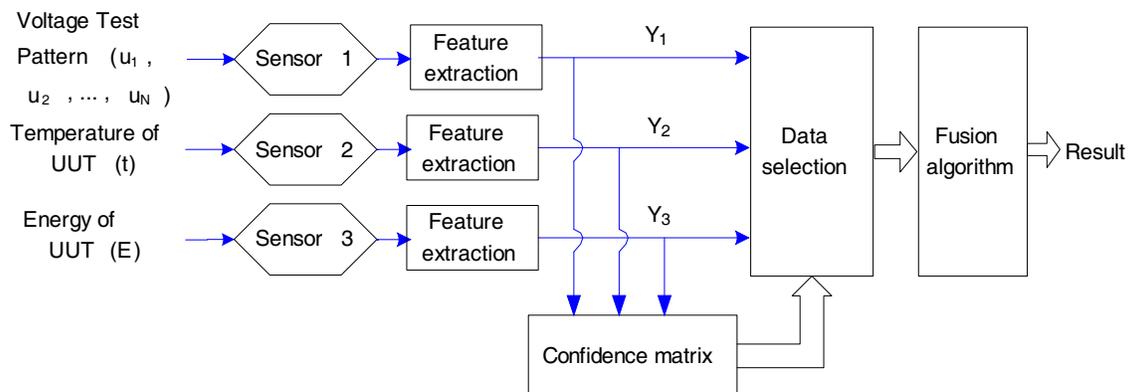


Figure 2: Scheme of decision-based data fusion.

Multi-sensor data fusion method consists of five steps:

- (1) Sensors collect individually different types of the data of system under test, then send them to Feature extraction part. As for electronic equipments, we select voltage test pattern, temperature and energy as its input. As researches have shown, they are comparatively main attributes of an electronic system which is changing with time and circumstances.
- (2) Feature extraction processes the raw data from every sensors, such as A/D conversion and outputs feature patterns Y_i .
- (3) The data in Confidence matrix represent the credibility of Y_i . It is the evidence of data selection. The credibility of Y_i depends on two factors:
 - The credibility of the sensor.
 - The credibility of data source.

Let the sensors are identical with each other in credibility, the data of contact measurement are more creditable than that of non-contact measurement, under current condition. Voltage and current are the data of contact measurement, though temperature and energy is not. So we can set voltage test pattern highest credibility, temperature, the lower, and energy, lowest.

- (4) Data selection module selects data according to the data in confidence matrix, then sends the selected data to fusion algorithm module.
- (5) fusion algorithm module is the crucial part of data fusion. A BP neural network fusion algorithm is shown as follows:

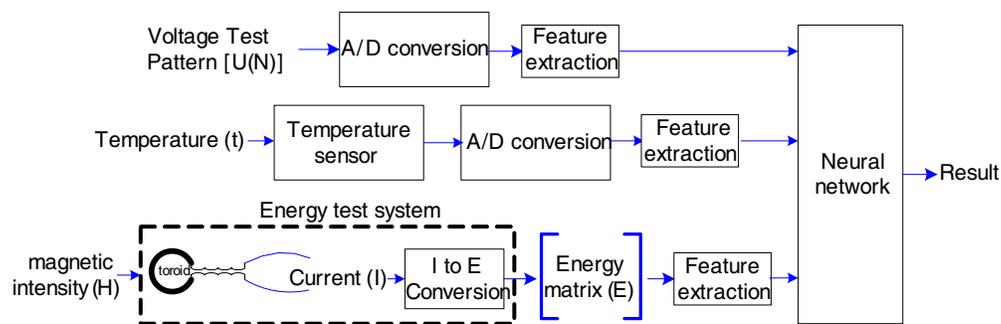


Figure 3: BPNN-based data fusion algorithm structure.

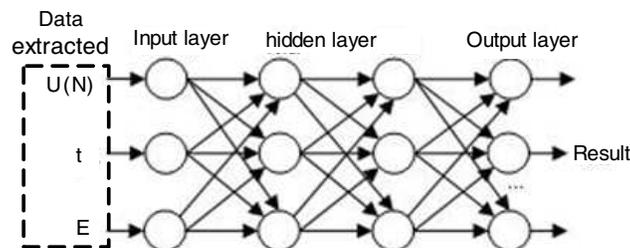


Figure 4: BP network architecture in PHM system.

BP neural network (Back-Propagated Delta Rule Networks) is a feed-forward multi-layer network. It has three fundamental layers: input layer, hidden layer and output layer. The number of hidden layers depends on the topological structure that the practical situation needs. It can set a nonlinear mapping from input vectors to output vectors. So it is widely used in pattern recognition, data compression and approximation of function.

(1) Input layer

The neurons of BP neural network in the input layer are voltage test pattern, temperature of the object and energy of it.

(2) Training

Data that have been pre-processed and featured are employed as the network's training data, while data from unknown fault models are as its prognostic data.

The trial and error procedure was employed to optimize the number of neurons in the hidden layer. The weights are adjusted by the feedback during the learning stage using the back-propagation learning algorithm that uses a gradient search technique to minimize the mean square error between the actual output pattern of the network and the desired output pattern. Then the architecture of BP neural network is constructed to ensure the minimum total prediction error.

(3) Prognosing

The BP model, after proper training, is used for predicting. Data from unknown fault models are as the network's prognostic input. Output is the future state of the system and its probable fault models.

4. CONCLUSION

The paper presents an integration of a PHM which combines every functional module into an organic and compositive system. It also puts up a BP network data fusion algorithm to reduce the uncertainty of data inputs. Based on current research, voltage, temperature and energy are three efficient inputs for prognose, but how extent the energy can be used in prognose to is still in discussion.

REFERENCES

1. Tan, Y., Y. He, C. Cui, and G. Qiu, "A novel method for analog fault diagnosis based on neural networks and genetic algorithms," *IEEE Transactions on Instrumentation and Measurement*, Vol. 57, No. 11, 2631–2639, November 2008.
2. Haan, B., "A model of the competitive advantage of prognostics and health management," *Reliability and Maintainability Symposium, RAMS 2009, Annual*, 442–447, 2009.
3. Xu, E., T. Weber, A.-M. Buibish, J. L. Lewis, and T. Hughes, "Communications solutions for the prognostics & health management architecture," *Military Communications Conference, MILCOM 2008, IEEE*, 1–9, 2008.
4. Ma, Z., "A new life system approach to the prognostic and health management (PHM) with survival analysis, dynamic hybrid fault models, evolutionary game theory, and three-layer survivability analysis," *IEEEAC*, paper #1700, Final Version Updated on January 7, 2009.
5. Brown, E. R., E.-E. Moore, E. R. Brown, and N. N. McCollom, "Prognostics and health management a data-driven approach to supporting the F-35 lightning II," *IEEEAC*, paper #1597, Version 3, Updated January 19, 2007.
6. Bagul, Y. G., I. Zeid, and S. V. Kamarthi, "A framework for prognostics and health management of electronic systems," *IEEEAC*, paper #1366, December 14, 2007.
7. Grasso, F., S. Manetti, and M. C. Piccirilli, "A method for the automatic selection of test frequencies in analog fault diagnosis," *IEEE Transactions on Instrumentation and Measurement*, Vol. 56, No. 6, 2322–2329, December 2007.
8. Janasak, K. M. and R. R. Beshears, "Diagnostics to prognostics — A product availability technology evolution," *The 53rd Annual Reliability and Maintainability Symposium (RAMS 2007)*, 113–118, 2007.
9. Briand, D. and J. E. Campbell, "Real time consequence engine," *Reliability and Maintainability Symposium, RAMS' 07, Annual*, 53–58, 2007.
10. Azzam, H., P. Knight, and J. Cook, "FUMSTM fusion and decision support for intelligent management of aircraft data," *IEEEAC*, paper #1197, Version 6, January 9, 2006.
11. Callan, R., B. Larder, and J. Sandiford, "An integrated approach to the development of an intelligent prognostic health management system," *IEEEAC*, paper #1104, Version 2, Updated January 4, 2006.
12. Fatih, C. G., S. Valentine, and K. Navarra, "Methodologies for integration of PHM systems with maintenance data," *IEEEAC*, paper #1191, Version 1, Updated September 20, 2006.
13. Sun, C., K. Nguyen, and L. Vu, "Prognostic/diagnostic health management system (PHM) for fab efficiency," *2006 IEEE/SEMI Advanced Semiconductor Manufacturing Conference*, 433–438, 2006.

14. Berenji, H., Y. Wang, and A. Saxena, “Dynamic case based reasoning in fault diagnosis and prognosis,” *Proceedings of IEEE International Conference on Fuzzy Systems*, 845–850, 2005.
15. Qiu, H., H. T. Liao, and J. Lee, “Degradation assessment for machinery prognostics using hidden Markov models,” *Proceedings of the ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference (DETC2005)*, 531–537, 2005.
16. Baruah, P. and R. B. Chinnam, “HMMs for diagnostics and prognostics in machining processes,” *International Journal of Production Research*, Vol. 43, No. 6, 1275–1293, 2005.
17. Biagetti, T. and E. Sciubba, “Automatic diagnostics and prognostics of energy conversion processes via knowledge based systems,” *Energy*, Vol. 9, No. 12–15, 2553–2572, 2004.
18. Byington, C. S., P. W. Kalgren, R. Johns, and R. J. Beers, “Embedded diagnostic /prognostic reasoning and information continuity for improved avionics maintenance,” *Proceedings of IEEE Systems Readiness Technology Conference (Autotestcon2003)*, 320–329 2003.
19. Hess, A. and L. Fila, “The joint strike fighter (JSF) PHM concept: Potential impact on aging aircraft problems,” *IEEEAC*, Vol. 6, 3021–3026, 2002.

A Study of Deformation Monitoring Using StaMPS Technique

Huan-Huan Liu^{1,2}, Jian-Ping Chen^{1,2}, Hong-Li Zhao^{1,2}, Jing-Hui Fan³, and Xiao-Fang Guo³

¹School of the Geosciences and Resources, China University of Geosciences (Beijing), Beijing 100083, China

²Beijing Land Resources Information Development Research Laboratory, Beijing 100083, China

³China Aero Geophysical Survey & Remote Sensing Centre for Land & Resources, Beijing 100083, China

Abstract— In this paper, StaMPS, a new PS method, which uses spatial correlation of interferogram phase to find PSs in almost all terrain without temporal variation models in the deformation is applied to monitor the substance in Tianjin area and landslide in Xintan area. After PSs selected and the deformation estimated by StaMPS, Kriging interpolation is applied to Tianjin test to get a continuous deformation sense and the results are shown in 3D views.

1. INTRODUCTION

Space-borne SAR is an active system, which can work day and night and is almost weather independent. By means of the analysis of the evolution of the SAR phase signals among different acquisitions and subtracting the topographical relief, information regarding the ground deformation can be extracted, that is the principle of DInSAR. But geometrical and temporal decorrelations degrade the accuracy of DInSAR, and even sometimes make the measurement impossible in deformation monitoring [1–10].

PSInSAR (permanent scatter InSAR) [2] is possible to avoid many of the limitations of conventional DinSAR by only analyzing certain pixels which behave like point scatters and retain some degree of correlations. The methods to identify PS pixels have been developed and practiced by some groups, such as Dehls et al. [3], Crosetto et al. [5], Lynons and Sandwell [6], Werner et al. [7], Yixian Tang [11], Daqing Ge [10], Jinghui Fan [8], etc. In almost all of these algorithms, an initial set of PS pixels is selected by amplitude variation. This kind of methods may fail in non-urban areas because the relationship between amplitude dispersion and phase stability will break down for low SNR scatters. Besides, approximate temporal models are required to refine PS, estimate and remove nuisance terms.

Stanford Method for PS (StaMPS) is a new PS method developed at Stanford University by Andy Hooper et al., which uses spatial correlation of interferogram phase to find a network of stable pixels in almost all terrain without prior knowledge of temporal variations in the deformation [1].

In this paper, StaMPS is applied to monitor the substance in Tianjin area and landslide in Xintan, using ENVISAT ASAR images, SRTM DEM and DEOS precise orbits. After interferogram series formation, PS are selected and the deformation information are estimated by StaMPS. Then for Tianjin test area, Kriging interpolation is applied to get a continuous deformation sense and the results are shown in 3D views.

2. DATA PREPARATION

There are 14 SLCs for Tianjin covering from Oct. 2003 to Dec. 2005, and 30 SLCs for Xintan from Aug. 2003 to Dec. 2007. The SLC senses acquired on 2005-08-27 and 2005-09-25 are respectively chose as the master images for the two test areas, and the temporal baseline and spatial perpendicular baselines are shown in Fig. 1.

For the generation of differential interferograms, Doris software (Delft Object oriented Radar Interferometric Software) was used and the workflow has been introduced by the guide documents of Doris and StaMPS software. It is worth mentioning that, to avoid high decorrelation, when coregistration, only master-slave pairs with baseline less than a specified maximum are directly coregistered, while the other slave scenes are coregistered to the 3 nearest slave scenes closer to the master, and the resulting coefficients are estimated by weighted least-squares inversion.

3. THEORY OF PS IDENTIFY AND SELECTION

PS are identified based on the analyzing of phase characters. For each pixel in the topographically corrected interferogram, its phase can be considered to the wrapped sum of 5 terms:

$$\phi_{int,x,i} = \phi_{def,x,i} + \phi_{top,x,i} + \phi_{atm,x,i} + \phi_{orb,x,i} + \phi_{n,x,i} \quad (1)$$

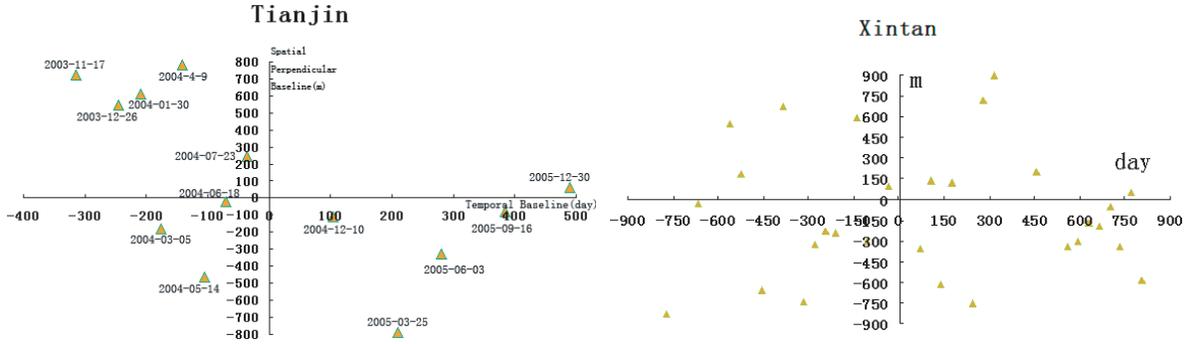


Figure 1: Temporal and spatial perpendicular baselines between slave and master images for two test series.

where $\phi_{def,x,i}$ is the phase change of the movement in the satellite line-of-sight (LOS) direction, $\phi_{top,x,i}$ is the residual phase caused by uncertainty in the DEM, $\phi_{atm,x,i}$ is the phase difference of in atmospheric delay of two passes, $\phi_{orb,x,i}$ is the residual phase due to orbit inaccuracies, and $\phi_{n,x,i}$ is the noise term due to variability in scattering from the pixel, thermal noise and coregistration errors. The pixels we seek as PS are those $\phi_{n,x,i}$ of which is small enough and doesn't completely obscure the signal.

The first four terms of the right side of (1) can dominate the noise term, and make it difficult to identify PS. So they have to be estimated and subtracted, leaving an estimate for $\phi_{n,x,i}$, which can then be assessed statistically and refined in an iterative way. As $\phi_{def,x,i}$, $\phi_{atm,x,i}$ and $\phi_{orb,x,i}$ can be assumed to be spatially correlated over a specified scale. Once we calculate the mean phase of the surrounding PS pixels, and subtract it from (1), we will get:

$$\phi_{int,x,i} - \bar{\phi}_{int,x,i} = \Delta\phi_{top,x,i} + \phi_{n,x,i} - \bar{\phi}'_{n,x,i} \quad (2)$$

The phase error caused by uncertainty in the DEM $\phi_{top,x,i}$ is approximately proportional to the spatial perpendicular baseline, that is $\Delta\phi_{top,x,i} = B_{\perp,x,i}K_{top,x,i}$ (3). Substitute (3) into (2), and $K_{top,x,i}$ is the only term correlates with baseline, which can be estimated in a least square sense, using all the available interferograms.

On the basis of above phase analyzing and assuming $\bar{\phi}'_{n,x,i}$ is very small, we get a variation of the residual phase for pixels, which can measure the temporal coherence:

$$\gamma_x = \frac{1}{N} \left| \sum_{i=1}^N \exp j \left(\phi_{int,x,i} - \bar{\phi}_{int,x,i} - \Delta\hat{\phi}_{top,x,i} \right) \right| \quad (3)$$

As the locations of PS are unknown, an iterative algorithm has to be used to identify PS in all locations simultaneously. For computational reasons, an initial selection of PS candidates based on amplitude dispersion is used for each test data. γ_x is gotten assuming $\bar{\phi}'_{n,x,i}$ is very small. By rejecting candidates with low gamma and recalculating $K_{top,x,i}$ and γ_x interactively, the $\bar{\phi}'_{n,x,i}$ will be smaller and smaller and γ_x gradually becomes dominated by $\phi_{n,x,i}$.

The next step is to select PS based on the calculated value of γ_x , and $\gamma^{threshold}$ should be set as a threshold that maximizes the number of real PS selected and keeping the fraction of random phase pixels selected (false positives) below a specified value q . False positives is defined as

$$\frac{(1 - \alpha) \int_{\gamma^{THRESH}}^1 p_R(\gamma_x) d\gamma_x}{\int_{\gamma^{THRESH}}^1 p(\gamma_x) d\gamma_x} = q \quad (4)$$

where $p(\gamma_x)$ is the probability density of γ_x , which is a weighted sum of probability density for PS pixels $p_{PS}(\gamma_x)$ and probability density for non-PS pixels $p_R(\gamma_x)$:

$$p(\gamma_x) = (1 - \alpha)p_R(\gamma_x) + \alpha p_{PS}(\gamma_x) \quad (5)$$

As for $\gamma_x < 0.3$, $p_{PS}(\gamma_x)$ can be assumed to be 0, which implies $\int_0^{0.3} p(\gamma_x) = (1 - \alpha) \int_0^{0.3} p_R(\gamma_x) d\gamma_x$. α can then be estimated by evaluating the integral of the test data on the left side and the integral

of a set of simulation data on the right side. As there is a correlation between amplitude variance $\hat{D}_{A,x}$ and phase stability γ_x , false positive can be calculated more accurately by estimating $\alpha(\hat{D}_{A,x})$ correlated with $\hat{D}_{A,x}$.

Besides, γ_x is expected to decrease with increasing $\hat{D}_{A,x}$, and empirically the relationship is linear. PS is then reselected whose $\gamma^{threshold} > k\hat{D}_{A,x}$, the best-fitting for k is calculated by least-squares inversion.

For groups of adjacent stable pixels, only pixels with the highest γ_x are selected; partially PS, which appear to be persistent only in certain interferograms, are rejected with a estimate of the variance of γ_x .

4. DEFORMATION ESTIMATION

The deformation information of PS could be calculated after the phase unwrapping and nuisance terms removal.

Before phase unwrapping, non-spatially correlated terms, mostly caused by look angle error and some terms caused from ‘master’, have to be subtracted from the wrapped differential interferogram phase.

For phase unwrapping, we first resample the wrapped phases to grids and filter them and then the grids are interpolated using a nearest neighbor approach. The following phase unwrapping is applied to the gridded interferograms in time and then space domain.

After unwrapping, there are still several terms masking the deformation phase. The spatially uncorrelated part of nuisance terms are modeled as noise and removed, while the spatially correlated parts are divided to two parts — time correlated and time uncorrelated part, both of which is estimated using a combination of temporal and spatial filtering. The time correlated part consists of the master contribution to atmosphere and orbit error (AOE), and it is estimated by a low-pass filter in time. The time uncorrelated part is estimated by a high-pass filter in time.

Unwrapping errors are more likely to occur in longer perpendicular baseline interferograms, because there is more noise and the phase due to spatially-correlated look angle (SCLA) error is larger. The unwrapping accuracy can be further improved by redoing phase unwrapping after subtracting SCLA, AOE phase in an interactive way until the unwrapped result becomes stable and reliable.

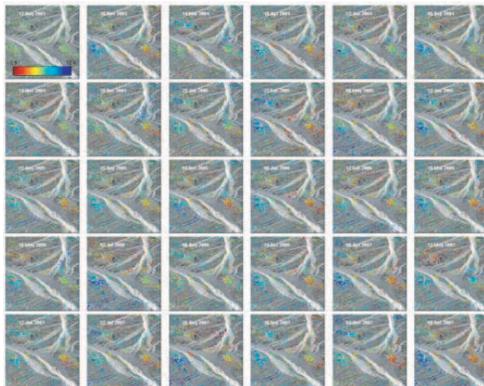


Figure 2: Xintan temporal series of the phase difference relative to the earliest differential interferogram.

5. RESULT SHOW AND ANALYSIS

For Xintan test site, the cropped master image is 200 lines *400 pixels. As the preliminary result is still being enhanced and not be validated, here we just show the temporal series of the phase difference relative to the earliest differential interferogram (Fig. 2). The deformation in LOS direction is proportional to the phase difference and we can infer from Fig. 2 the tendency and variation of surface deformation in this area. However, as the phase unwrapping starting point's stability is uncertain, each map just shows the relative deformation.

For Tianjin test site, the cropped master image is 16000 lines *3300 pixels. After the generation of unwrapped temporal series and the deformation in LOS direction, we then work out the mean deformation velocity of the PS for year 2004 and 2005, and plot them on the mean amplitude map

of the 14 slave ASAR images (Fig. 3). As the phase unwrapping didn't start from a phase stable point, the result is actually a relative displacement map. The red parts on map, with labels nearby, are corresponding to the places with serious subsidence.

As the PS are discrete and irregularly distributed the deformation for some interesting places may not be mapped. We then apply Kriging interpolation method to the discrete deformation velocity of PS, and get a continuous deformation field for the two years. Setting their highest value (also relative to the starting PS) as the base height respectively, contours with interval of 20 mm are generated and overlaid on the field (Fig. 4).

Take the displacement value as the third dimension, we can also get a 3D view for the annual mean deformation velocity in Tianjin test area (Fig. 5).

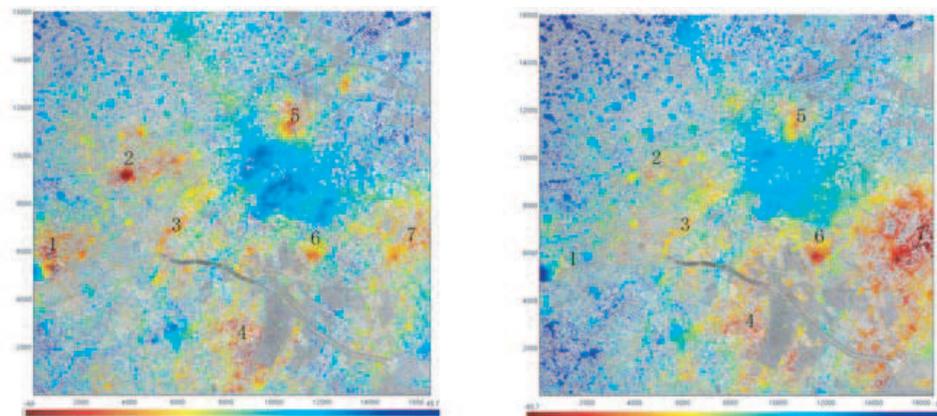


Figure 3: Distribution of mean deformation velocity of PS relative to PS in the southwest for 2004 and 2005.

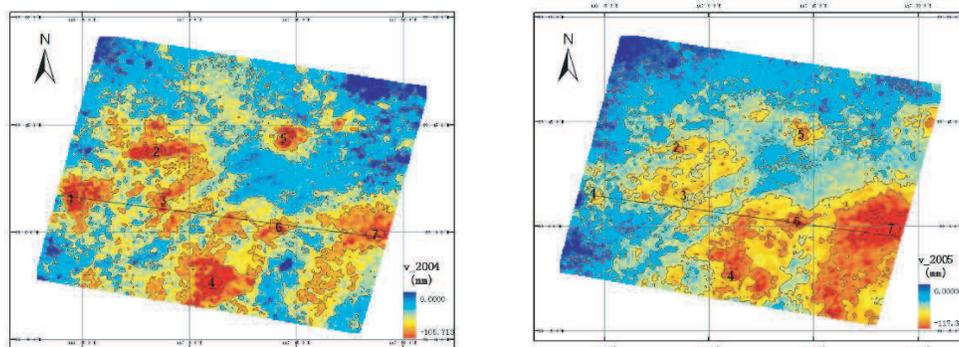


Figure 4: The mean deformation velocity map with contours for 2004 and 2005.

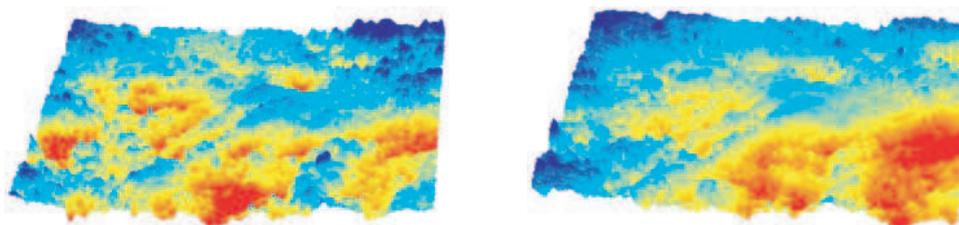


Figure 5: 3D view for annual mean deformation velocity for 2004 and 2005 in Tianjin test area.



Figure 6: 3D path profile maps of a line in the middle-lower part for 2004 and 2005 deformation velocity.

In the middle-lower part of the map, from west to east, the land subsidence tendency has changed a lot from 2004 to 2005. We set a profile line from (116.28, 39.07) to (117.46, 38.99), as the line shown in Fig. 4, and create the 3D path profile maps from 2004 and 2005 annual mean deformation velocity maps respectively (Fig. 6).

As for the result validation of Tianjin area test, we have got several land subsidence center labeled 1 to 7, the deformation trend for year 2004 agree with previous results [9,12], while the result for 2005 is still to be carefully checked and compared with other data.

6. CONCLUSION

By using StaMPS, subsidence in urban area can be mapped, with the deformation in time series not limited to certain model, such as linear speed. Although StaMPS can only calculate discrete deformation values, as long as the PS are with enough density, we can get the deformation for the whole area by Kriging interpolation and contour creating.

When it comes to non-urban area, the result is still not good, which may be caused the complex test area or other data preparation reason, for example, image coregistration. Besides, the input parameters in PS selection are very important but not easy to set.

In the future work, we will continue to enhance the work in Xintan landslide, and research more accurate and robust phase unwrapping methods for 3D data, especially for data discrete and irregular distributed in space.

ACKNOWLEDGMENT

This work was supported by the Ministry of Land and Resources (Project No. 200811053). The precise orbit datas are offered by Delft Institute for Earth-oriented Space Research (DEOS), and the DEM datas are from SRTM DEM. Data preparation is accomplished by Doris, and PS selection and deformation estimation by StaMPS.

REFERENCES

1. Andrew, J. H., "Persistent scatter radar interferometry for crustal deformation studies and modeling of volcanic deformation," A Dissertation Submitted to the Department of Geophysics and the Committee on Graduate Studies of Stanford University in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy, May 2006.
2. Ferretti, A., C. Prati, and F. Rocca, "Permanent scatters in SAR interferometry," *IEEE Transactions Geoscience and Remote Sensing*, Vol. 39, No. 1, 820 2001.
3. Dehls, J. F., M. Basilico, and C. Collesanti, "Ground deformation monitoring in the ranafjord area of norway by means of the permanent scatters technique," *IEEE Int. Geoscience and Remote Sensing Symposium*, 203–207, 2002.
4. Zebker, H. A. and J. Villasenor, "Decorrelation in interferometric radar echoes," *IEEE Transactions Geoscience and Remote Sensing*, Vol. 22, 2517–2520, 1995.
5. Crosetto, M., A. Arnaud, J. Dura, E. Biescas, and M. Agudo, "Deformation monitoring using remotely sensed radar interferometric data," *Proc. 11th FIG Symposium on Deformation Measurements*, Santorini, 2003.
6. Lyons, S. and D. Sandwell, "Fault creep along the southern San Andreas from interferometric synthetic aperture radar, permanent scatterers, and stacking," *Journal of Geophysical Research*, Vol. 108, No. B1, 2047–2070, 2003.
7. Werner, C., U. Wegmüller, T. Strozzi, and A. Wiesmann, "Interferometric point target analysis for deformation mapping," *IEEE Int. Geoscience and Remote Sensing Symposium*, 2003.
8. Fan, J.-H., "Monitoring subsidence in Tianjin area using interferogram stacking based on coherent targets," *Journal of Remote Sensing*, Vol. 12, No. 1, 111–118, Jan. 2008.
9. Fan, J.-H., "A preliminary study of the subsidence in Tianjin area using ASAR images based on PSInSAR technique," *Remote Sensing for Land & Resources*, No. 4, 23–28, Dec. 2007.
10. Ge, D.-Q., "Surface deformation monitoring with multi-baseline D-InSAR based on coherent point target," *Journal of Remote Sensing*, Vol. 11, No. 4, 574–560, Jul. 2007.
11. Tang, Y., "A study of deformation in Suzhou area using PSInSAR (in Chinese)," *Progress in Natural Science*, Vol. 08, 1015–1020, 2006.
12. Mu, J. and G. Yao, "Time series radar interferometry method for surface deformation monitoring," *Remote Sensing for Land & Resources*, No. 3, 28–30, Sep. 2007.
13. Zebker, H. A., "Phase unwrapping in three dimensions with application to InSAR time series," *Journal of the Optical Society of America A*, Vol. 24, No. 9, 2737–2747, 2007.

Compact Dual-band Balanced Handset Antenna for WLAN Application

A. G. Alhaddad¹, R. A. Abd-Alhameed¹, D. Zhou¹,
C. H. See¹, E. A. Elkhazmi², and P. S. Excell³

¹Mobile and Satellite Communications Research Centre, University of Bradford, UK

²The Higher Institute of Electronics, Bani Walid, Libya

³Centre for Applied Internet Research, Glyndwr University, Wrexham, UK

Abstract— In this paper, a balanced antenna for mobile handset applications with dual-frequency performance, covering the 2.4 GHz and the entire 5 GHz WLAN frequency bands, is investigated and discussed. The antenna is a thin-strip planar dipole with folded structure and a dual-arm on each monopole. For validation, the antenna prototype was fabricated and tested. The performance of this balanced antenna was verified and characterised in terms of the antenna return loss, radiation pattern, power gain and surface current distribution of the proposed antenna. The predicted and measured results show good agreement.

1. INTRODUCTION

Wireless communication has been characterized of the new modern move to make the mobile handsets small and light as possible; without compromising functionality. To miniaturize in line with consumer needs and aspiration and retain multiband functionality, mobile handsets development must be characterized by making all physical components as small as physically possible. The key concerns considered here on the design of antenna systems for small handsets relates to keeping the antenna performance unchanged or improved, even though the antenna size becomes small and reduces the degradation of antenna performance caused by the operator's adjacent effect [1]. A balanced structure is a genuine choice to avoid the aforementioned degradation of the antenna performance when held by users [2] since balanced currents only flow on the antenna element in this type of antenna, thus dramatically reducing the effect of current flow on the ground plane. As a result, balanced antennas should have good efficiency and more important to maintain their performance when in use adjacent to the human body [3]. In recent years, several novel mobile antennas designed with the balanced technique have demonstrated the enhanced stability of antenna performance, compared to the unbalanced type, when the handset is approximately placed next to the human head and/or hand [4–8].

A built-in planar metal plate antenna for mobile handsets with balanced operation is presented in this paper. The antenna was designed by folding a thin strip planar dipole with extra arm on each monopole. The antenna features balanced operation, is to reduce the current flow on the conducting surface of the handset body. The antenna design model intends to cover 2.4 GHz and the 5 GHz WLAN applications. In short, this paper presents and investigates a new design of a built-in dual-frequency balanced antenna for WLAN and short range wireless communications. The characteristics of this balanced folded dipole antenna with a novel dual-arm structure for mobile handsets are analysed, including calculating the return loss and the radiation patterns for comparisons.

2. ANTENNA DESIGN

The proposed balanced antenna structure is shown in Fig. 1. The antenna was constructed from a copper sheet with thickness of 0.15 mm. Fig. 1(b) shows one side of the folded dipole antenna, in which the copper plate was folded up to become a folded dipole antenna. The proposed antenna is achieved by using two tier processes in order to generate another resonant frequency.

Firstly, it was started by folding the monopole arm and having a slot inside each monopole with a cut on the bottom side, as shown Fig. 1(a). Secondly, an additional thin-strip arm was inserted into each arm of the planar dipole.

This folded element of the proposed antenna was designed to operate at 2.4 GHz with a single arm to generate the second resonant frequency for 5 GHz frequency band [9]. In order to achieve a low-profile folded (i.e., lower d) balanced antenna, while maintain the sufficient impedance bandwidth required at the two WLAN bands, a long slot is introduced on the each folded arm of the dipole

antenna. In this way, the equivalent wavelength of the surface current at 2.4 GHz is increased, compared to the case without the long slot. As a result, the folded antenna height (d) can be reduced by 50% and the low-profile design is therefore realised. The optimized dimensions of the proposed antenna to operate at required bands are: $a = 18.5$ mm, $b = 8$ mm, $d = 5$ mm, $c = 11.5$ mm, $w = 3$ mm, $h = 4$ mm, $t = 1.5$ mm, $f = 2$ mm, $g = 10.5$ mm. The antenna is mounted 1 mm above the ground plane with dimension of 90×40 mm.

Parametric study has been carried out to optimize the impedance matching bandwidth for the proposed antenna in order to achieve the required impedance matching covering the frequencies bands of interested at 2.4 GHz and 5 GHz bands for WLAN and short range communication systems. The antenna height (h and w) were considered to be the most sensitive parameters to control the impedance bandwidth of the proposed antenna for meeting the design goals. The parameter h was varied from 2 mm to 5 mm with 1 mm each step and w parameter was varied from 1 to 4 mm with also 1 mm step. To fully understand, the influence of these parameters based on its impedance bandwidth, the parametric study will be carried out here with only one parameter are varying at a time, while others were keep constant with the assume optimum value. The optimum value of h and w were found to be 4 mm and 3 mm as shown in Fig. 2 and Fig. 3 respectively. By modifying the length and location of the additional arm of the proposed antenna, it was able to let the antenna covers the required two frequency bands at acceptable return loss ≤ -10 dB. The proposed antenna features of the compact design used, has the size dimensions of $(l = 38) \times (w = 10) \times (h = 4)$ mm.

3. SIMULATION AND MEASUREMENT RESULTS

In order to effectively characterize the proposed antenna, a prototype antenna (see Fig. 4) was fabricated and tested. Two antenna properties were measured; return Loss and radiation pattern. In regards to return loss, two methods were undertaken for this measurement exercise, which includes using a balun and implementing the S -parameter method. By using the first method, a commercially hybrid junction from ET Industries [10] that operates from 2 to 12 GHz has been

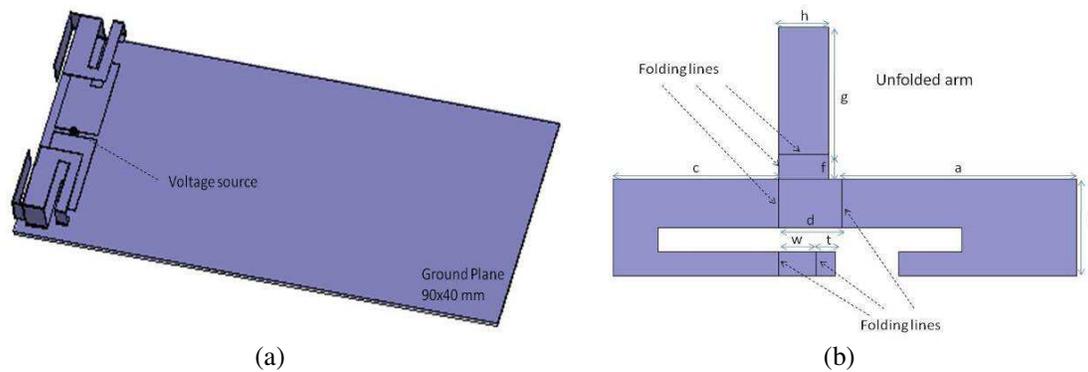


Figure 1: Balanced mobile antenna configuration studied; (a) the proposed antenna in this study and (b) unfolded arm with the important antenna parameters.

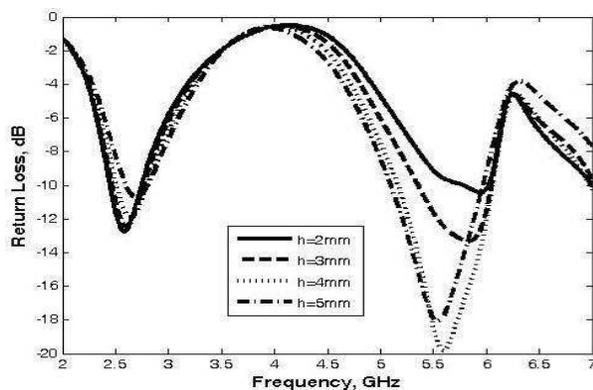


Figure 2: Variation of the parameter h on the effect of the return loss.

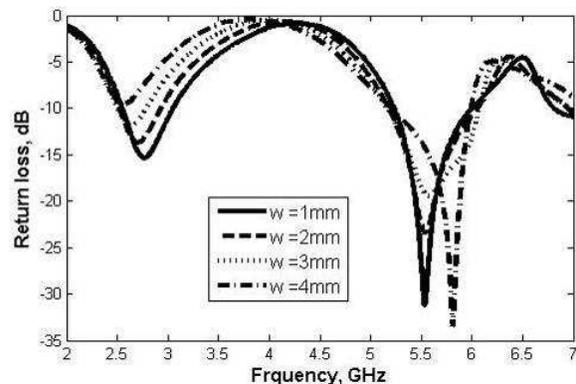


Figure 3: Variation of the parameter w on the effect of the return loss.

utilized in this measurement exercise, as a balanced feeding network (see Fig. 5). In this case, this balun is required as a support feeding network, to provide a balanced feed from an unbalanced source.

Figure 6 illustrates the measured return loss of the prototype antenna. In which measured return loss shows fairly good agreement with simulated result. Using the S -parameter method for measuring input impedance for the balanced antennas was the second method to verify the impedance of the proposed antenna. As for this methodology; balanced antennas are considered as two-port devices and the S -parameters can be obtained from a well-calibrated Network Analyzer. By employing some mathematical operation formula used by [11] the differential input impedance of the balanced antenna can be obtained. The results from this exercise also gave much confidence to



Figure 4: Photograph of fabricated prototype antenna.



Figure 5: Photograph of balun used in the measurement.

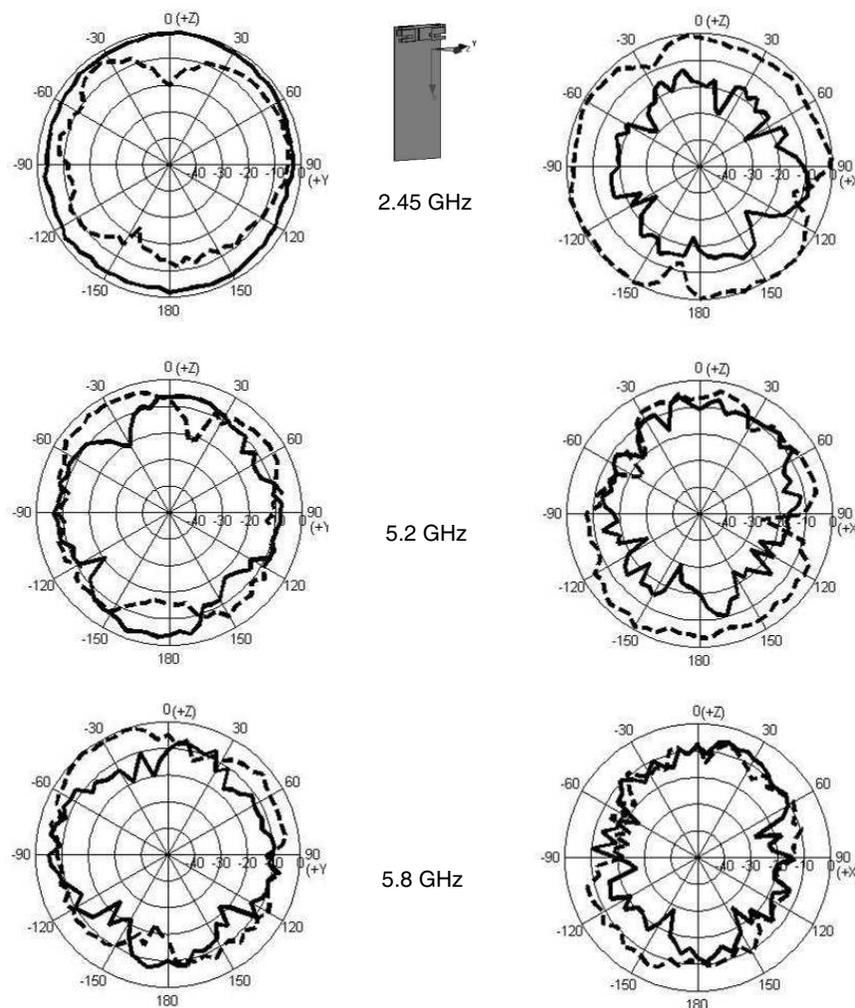


Figure 6: Radiation patterns of the proposed antenna for 2.45 GHz, 5.2 GHz and 5.8 GHz at: (left) xz plane; (right) yz plane, where '—' measured E_θ and '- - -' measured E_ϕ .

the measurement carried out, as the measured return loss in this method was relatively correlated with the first one, as can be observed in Fig. 6.

The radiation patterns of the proposed antenna were measured inside a far-field anechoic chamber by placing the antenna under test at one end of the chamber, while placing a standard gain horn antenna at the other end of the room. Two pattern cuts were taken for three WLAN operating frequencies that cover the designated whole bandwidth in this study. The radiation patterns in the xz plane and yz plane for the balanced antenna at 2.45 GHz, 5.2 GHz and 5.8 GHz were measured, as presented in Fig. 7. The antenna gain of the antenna was measured for the frequencies across the 2.4 GHz, 5.2 GHz and 5.8 GHz WLAN bands. It is notable that the insertion loss of the feeding network was subsequently compensated for each measured power gain over all bands. It was found that the maximum measured antenna gain at lower and upper WLAN bands were 4 dBi, 6 dBi and 5.6 dBi at the selected frequencies, respectively.

The surface current distribution on the mobile phone ground plane at three specific frequencies (including 2.4 GHz, 5.2 GHz and 5.8 GHz) was analyzed using the EM simulator and presented in Fig. 8, as can be seen in the Figure a high proportion of the current induced on the ground plane and the currents is mostly confined in the area underneath the proposed antenna. Moreover, a minimum current distribution appeared on the rest of the ground plane. This further explains the advantage of using balanced antenna design for future mobile handsets.

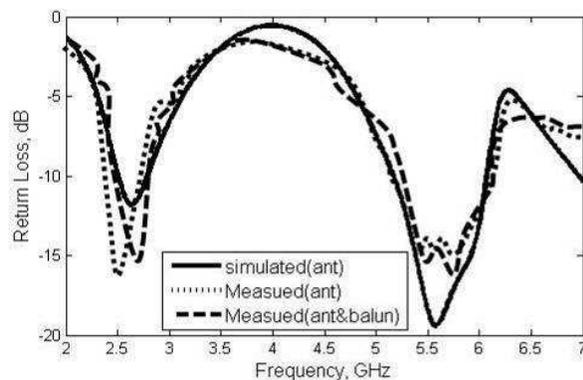


Figure 7: Comparison of simulated and measured return loss.

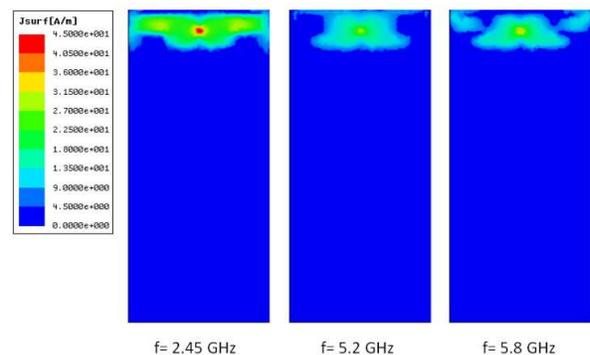


Figure 8: Surface current distributions for the proposed antenna.

4. CONCLUSION

A novel compact dual balanced handset antenna for mobile devices has been presented. The proposed antenna model was designed and measured to verify the design concept. It was shown that the proposed antenna covers 2.4 GHz (2.4–2.4835 GHz) and the 5 GHz (5.15–5.35 GHz & 5.650–5.925 GHz) WLAN applications. The characteristics of proposed balanced antenna was analysed in terms of antenna return loss, radiation pattern and power gain. The simulated results show good agreement with the measured one and therefore indicate that the proposed design can be recommended as a promising candidate mobile-antenna solution for WLAN applications.

REFERENCES

1. Morishita, H., H. Furuuchi, and K. Fujimoto, "Performance of balance-fed antenna system for handsets in vicinity of a human head or hand," *IEE Proc. — Microw. Antennas Propag.*, Vol. 149, No. 2, 859–891, April 2002.
2. Abd-Alhameed, R. A., P. S. Excell, K. Khalil, R. Alias, and J. Mustafa, "SAR and radiation performance of balanced and unbalanced mobile antennas using a hybrid formulation," *IEE Proceedings-science, Measurement and Technology, Special Issue on Computational Electromagnetics*, Vol. 151, No. 6, 440–444, November 2004.
3. Zhou, D., R. A. Abd-Alhameed, and P. S. Excell, "Wideband balanced folded dipole antenna for mobile handsets," *Proceedings of the European Conference on Antennas and Propagation: EuCAP 2007*, Paper No. MoPA.012, Edinburgh, UK, November 11–16, 2007.
4. Morishita, H., S. Hayashida, J. Ito, and K. Fujimoto, "Analysis of built-in antenna for handset using human (head, hand, finger) model," *Electronics and Communications in Japan, Part 1*, Vol. 86, No. 9, 35–45, 2003.

5. Kingsley, S., “Advances in handset antenna design,” *RF Design*, 16–22, May 2005.
6. Collins, B. S., S. P. Kingsley, J. M. Ide, S. A. Saario, R. W. Schlub, and S. G. O’Keefe, “A multi-band hybrid balanced antenna,” *IEEE 2006 International Workshop on Antenna Technology: Small Antennas; Metamaterials*, 100–103, White Plains, New York, March 6–8, 2006.
7. Zhou, D., R. A. Abd-Alhameed, C. H. See, A. G. Alhaddad, and P. S. Excell, “New mobile balanced mobile antenna with wide bandwidth performance,” *Proceeding of the European Conference on Antennas and Propagation, EuCAP 2009*, 549–552, Berlin, Germany, March 23–27, 2009.
8. Arenas, J. J., J. Anguera, and C. Puente, “Balanced and single-ended handset antennas: free space and human loading comparison,” *Microwave and Optical Technology Letter*, Vol. 51, No. 9, 2248–2254, September 2009.
9. Zhou, D., R. A. Abd-Alhameed, C. H. See, S. W. J. Chung, A. G. Alhaddad, and P. S. Excell, “Dual-frequency balanced mobile antenna for WLAN and short range communication systems,” *PIERS Proceedings*, 1264–1267, Beijing, China, March 23–27, 2009.
10. ET Industries, USA, <http://www.etiworld.com/>.
11. Meys, R. and F. Janssens, “Measuring the impedance of balanced antennas by an S -parameter method,” *IEEE Antennas and Propagation Magazine*, Vol. 40, No. 6, 62–65, December 1998.

Isolation Enhancement Based on Adaptive Leakage Cancellation

J. Y. Wang¹, B. Lv¹, W. Z. Cui², W. Ma², J. T. Huangfu¹, and L. X. Ran¹

¹Department of Information and Electronics Engineering, Zhejiang University
Hangzhou 310027, China

²Xi'an Institute of Space Radio Technology, Xi'an 710000, China

Abstract— In the design of direct-conversion architecture, an important issue is the leakage from the transmitter to the receiver, which can cause DC-offset problem. In this paper, an adaptive leakage cancellation method is proposed to improve the isolation performance. The fundamental concept of this method is to introduce an anti-phase signal, to cancel the leakage. In order to improve the performance, variable attenuator and phase-shifter are used in the system, and a closed-loop configuration is also employed. Both of the mathematical analysis and simulation about this method have been provided in this paper. And the results show that, good isolation performance can be achieved automatically and real-timely in this method.

1. INTRODUCTION

Direct-conversion architecture is widely used in the RF system nowadays. In the design of Direct-conversion system, there are some special issues should be pay attention to, an important one of which is the leakage problem due to the no perfect isolation performance. After the leakage signal is mixed with the LO signal, a zero frequency offset voltage is produced, which can corrupt the receiving signal and even saturate the following stages [1, 2].

Much effort has been taken to solve the leakage problem. The simplest way is to employ AC coupling in the receiving path. However, in several applications, such as the phase demodulation in [3], there is critical “DC information” in the received signal, which can not be removed by AC coupling. Another promising method is to introduce an anti-phase signal to cancel the leakage signal [4–6]. The drawback in [4] is that, the directional coupler proposed could not be tunable. And some changes in the transceiver might introduce unknown amplitude and phase mismatch between the leakage and cancellation signal, which could degenerate the isolation performance. After introducing variable components in [5, 6], the systems became more adaptive. However, calibration could not be made in such open loop systems, without the help of network analyzer. In this paper, a novel adaptive leakage cancellation method is proposed, with a closed-loop configuration to decrease the amplitude and phase mismatch automatically and real timely.

2. MATHEMATICAL ANALYSIS ON MISMATCH

The basic concept of the leakage cancellation method, is to introduce another signal, which has the same amplitude but anti-phase, to cancel the leakage signal, as shown in Figure 1. The main factor that affects the isolation performance, is the mismatch in amplitude and phase between the leakage and the cancellation signal. In this section, mathematical analysis is provided to evaluate the mismatch problem in this method.

The transmission signal $T(t)$ and leakage signal $L(t)$ can be described in (1) and (2), respectively. And the original isolation I is defined as the power ratio between them, as shown in (3).

$$T(t) = At \cos(\omega t) \quad (1)$$

$$L(t) = Al \cos(\omega t + \varphi_l) \quad (2)$$

$$I = -10 \lg \frac{\frac{1}{T} \int_0^T \frac{L(t)^2}{R} dt}{\frac{1}{T} \int_0^T \frac{T(t)^2}{R} dt} = -10 \lg \frac{A_l^2}{A_t^2} \quad (3)$$

If a cancellation signal $C(t)$ in (4) is introduced simultaneously, new isolation I' can be recal-

culated in (5) after combination.

$$C(t) = A_c \cos(\omega t + \varphi_c) \tag{4}$$

$$I' = -10 \lg \frac{\frac{1}{T} \int_0^T \frac{[L(t) + C(t)]^2}{R} dt}{\frac{1}{T} \int_0^T \frac{T(t)^2}{R} dt} = -10 \lg \frac{A_l^2 + A_c^2 + 2A_l A_c \cos(\varphi_l - \varphi_c)}{A_t^2} \tag{5}$$

If the amplitude of the $C(t)$ is equal to that of $L(t)$, while the phase difference between them is 180° , $L(t)$ will be cancelled completely after combination, and I' will become infinite large in theory. However, there is always some unknown attenuation and phase shift in the signal path, which results in mismatch in amplitude and phase. These two kinds of mismatch are defined in (6) and (7). And the isolation enhancement ΔI in (8), rather than I' itself, is used in following analysis.

$$\Delta AdB = -10 \lg \frac{A_c^2}{A_l^2} \tag{6}$$

$$\Delta\varphi = \varphi_l - \varphi_c - 180^\circ \tag{7}$$

$$\Delta I = -10 \lg \left[1 + \left(\frac{A_c}{A_l} \right)^2 + \left(\frac{2A_c}{A_l} \right) \cos(\varphi_l - \varphi_c) \right] \tag{8}$$

Figure 2 shows the relationship between mismatch and ΔI . Some important conclusions can be drawn here. a) The mismatch may introduce disastrous problem when the system is badly arranged. ΔI can even be a minus figure, in which case the cancellation signal itself acts as another leakage. b) Because the attenuation and phase shift are unknown in the signal path, variable components should be introduced to decrease the mismatch. c) The maximum point in Figure 2 doesn't equal infinite large. That's because the matched point, which means $\lg(0)$ in mathematics, can not be calculated in the program. The optimum point is determined by the minimum step in X and Y axis, which means optimum ΔI will be affected by the resolution of the variable attenuator and phase shifter. d) The Optimum point can be easily achieved, after closed-loop architecture is employed.

3. SYSTEM IMPLEMENT AND SIMULATION

With the analysis above, a closed-loop configuration is proposed in this section, as shown in Figure 3.

This closed-loop system is composed of two directional couplers (DC1 & DC2), a variable attenuator (VA), a variable phase-shifter (VP), a power combiner (PC), a power detector (PD), and an acquisition and processing unit (APU). DC1 is settled to derive cancellation signal from the transmission signal, a -10 dB coupling type may be employed to minimize the insertion loss introduced. VP and VA, which are both controlled by APU, are used to decrease the phase and

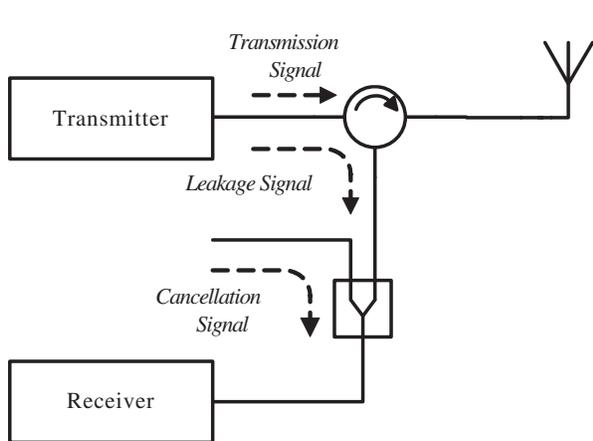


Figure 1: Basic concept of leakage cancellation.

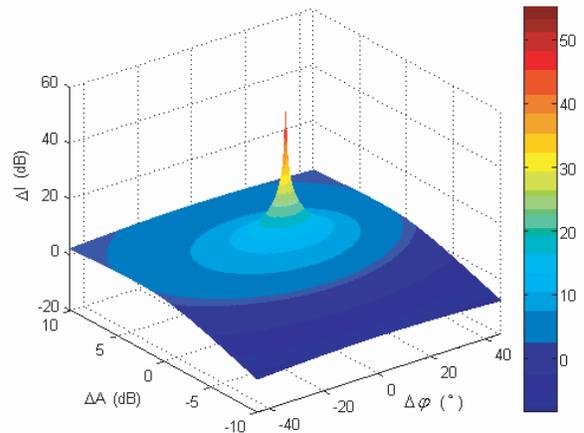


Figure 2: Relationship between mismatch and ΔI .

amplitude mismatch. After the cancellation signal and leakage signal are combined by PC, DC2 will couple -10 dB power of the combined signal to PD. This coupled power, which indicates the isolation level, is converted to a voltage on a linear scale by PD. Then, APU will find the optimum input for VP and VA automatically, due to the output of PD. After calibration, APU should go to sleep mode to save power consumption, until another calibration is carried out.

To verify this system, a simulation has been carried out. The procedure is shown in Figure 4. Firstly, random isolation (ranging from 18 dB to 32 dB) and phase difference between cancellation and leakage signal (ranging from 135° to 225°) were generated by the software. Secondly, the phase difference was changed in 32 steps, which was used to model the function of VP with 32 bits resolution. In each step, the power of combined signal would be recalculated, and then be converted to the output of PD on a linear scale. Thirdly, after a comparison procedure, optimum phase would be determined due to the minimum output of PD. Fourthly, the amplitude of cancellation signal would be changed in 32 steps, which was used to model the function of VA. And the output of PD

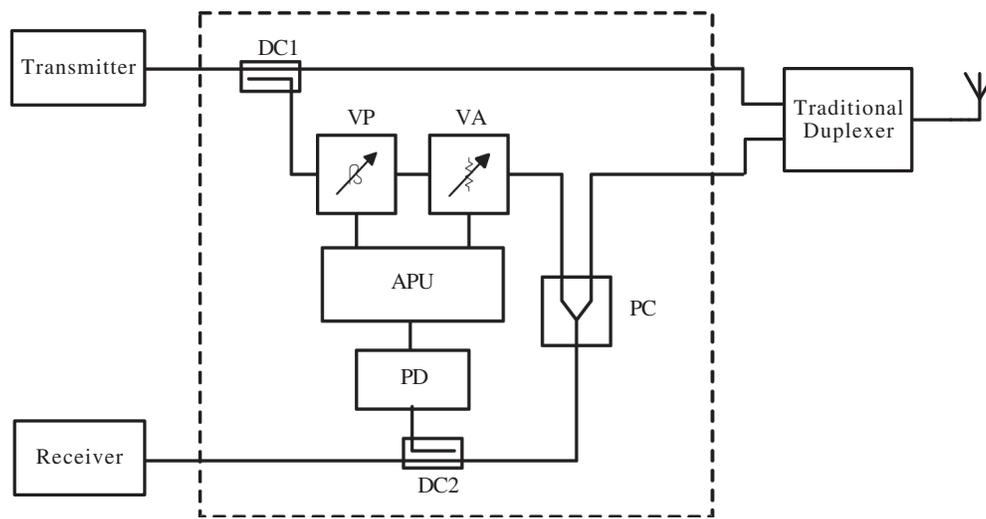


Figure 3: System configuration.

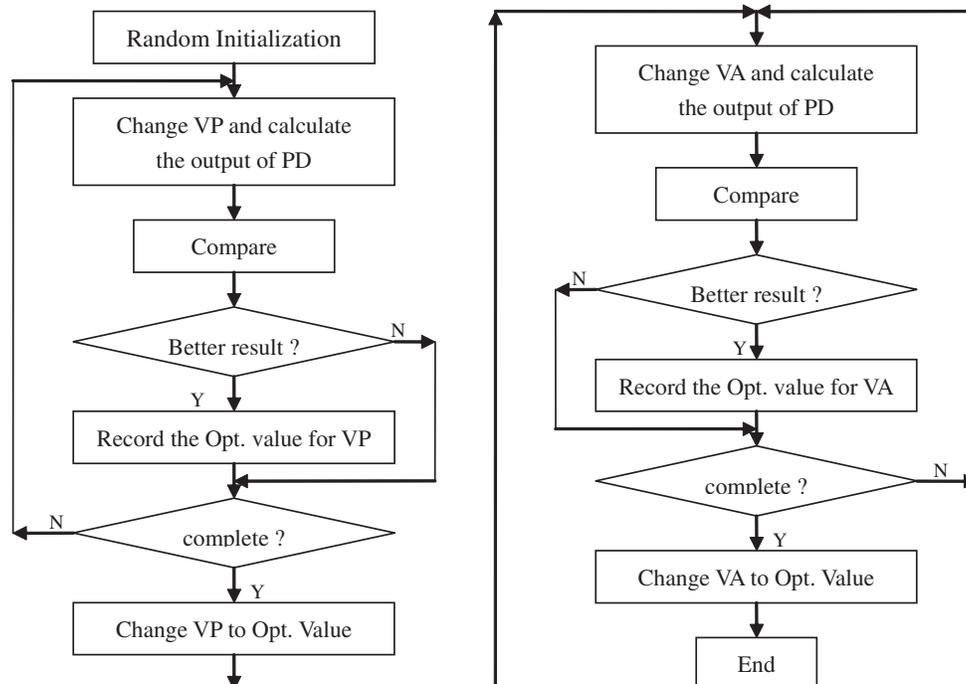


Figure 4: Simulation flow chart.

would be recalculated in each step as well. Fifthly, the optimum amplitude would be determined due to the output of PD. Finally, the best isolation was achieved under the optimum arrangement of VP and VA.

Figure 5 shows a simulation result for one time. The original isolation and phase difference were 21.98 dB and 155.8°, respectively. The curve between point A and point C on Figure 5(a), were the output of PD with different phase shift of cancellation signal. Point B has the minimum value in this section, which was the optimum phase shift point for isolation enhancement. After the optimum phase shift was determined, Figure 5(a) jumped to point D, which had the same value of B. The curve between point D and F, were the output voltage with different attenuation of cancellation signal. And the optimum attenuation was achieved on point E. Figure 5(b) shows the optimum isolation enhancement procedure in this simulation. The points on it are corresponding to that of Figure 5(a).

The same experiment has been carried out for 1000 times, and a histogram of the simulation result is shown in Figure 6(a). Enhancements were about 30 ~ 40 dB in most of these experiments (about 85%). To achieve better performance, another 1000 times experiments have been carried out, with the resolution of VP and VA changed to 256, as shown in Figure 6(b). The enhancements were about 17 dB higher in average.

The only drawback with higher resolution is that the operating time will be longer. Fortunately, the total number of instructions in the simulation is less than 50 times the resolution. For an ordinary 5 MIPS MCU, the operating time is only about 2.56 ms, when the resolution equals 256. In other words, the real time performance will not be influenced with high resolution.

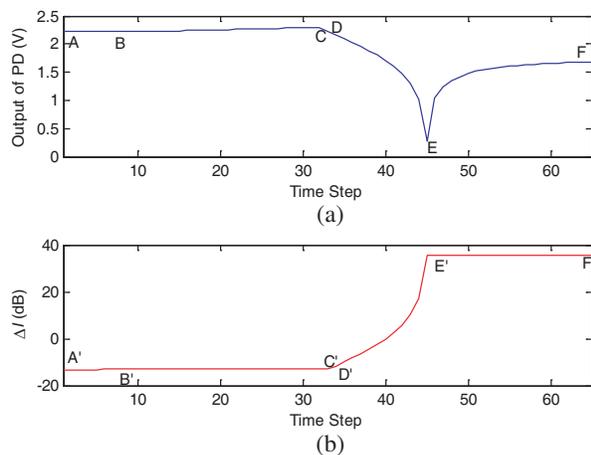


Figure 5: Simulation result. (a) Output of PD, and (b) ΔI .

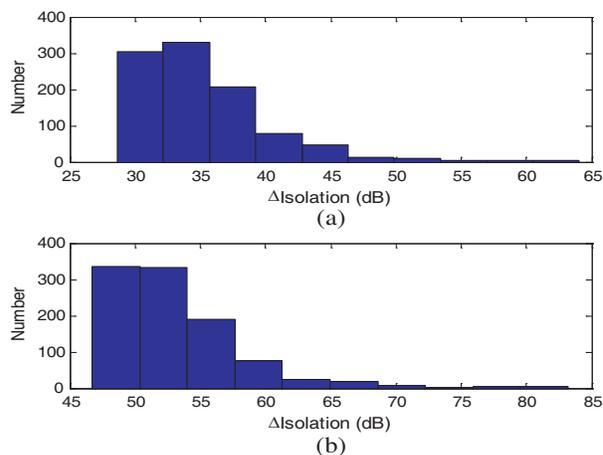


Figure 6: Histogram of simulation results with (a) resolution = 32, and (b) resolution = 256.

4. CONCLUSION

An adaptive leakage cancellation method, with closed-loop configuration has been presented in this paper. The simulation result shows that, at least 28 dB of isolation enhancement can be achieved if the resolution of VP and VA equals 32. Better performance can be achieved with higher resolution. The advantage of this configuration is that, the APU will find the optimum isolation automatically and real timely. More over, the components in the dashed box in Figure 3 can be integrated as an individual system, regardless of the selection of duplexer.

ACKNOWLEDGMENT

This work is sponsored by NSFC (Nos. 60531020, 60671003 and 60701007), 863 Project (No. 2009A-A01Z227), NCET-07-0750, ZJNSF (No. Y1080715), ZJSTP (No. 2009C31141), Ph.D. Programs Foundation of MEC (No. 20070335120) and the National Key Laboratory Foundation (Nos. 9140C5-304020901 and 9140C5304020704).

REFERENCES

1. Razavi, B., "Design considerations of direct-conversion receivers," *IEEE Trans. on Circuits and Systems II: Analog and Digital Signal Processing*, Vol. 44, No. 6, 428–435, 1997.

2. Bae, J.-H., W.-K. Choi, J.-S. Kim, G.-Y. Choi, and J.-S. Chae, "Study on the demodulation structure of reader receiver in a passive RFID environment," *Progress In Electromagnetics Research*, PIER 91, 243–258, 2009.
3. Park, B., O. B. Lubecke, and V. M. Lubecke, "Arctangent demodulation with DC offset compensation in quadrature doppler radar receiver systems," *IEEE Trans. on Microwave Theory and Techniques*, Vol. 55, No. 5, 1073–1079, 2007.
4. Chan, Y. K., B. K. Chung, and H. T. Chuah, "Transmitter and receiver design of an experimental airborne synthetic aperture radar sensor," *Progress In Electromagnetics Research*, PIER 49, 203–218, 2004.
5. Bai, P., Y. Z. Yin, and X. Yang, "A novel RX-TX front-ends for passive RFID reader with high isolation," *IEEE International Symposium on Microwave, Antenna, Propagation, and EMC Technologies for Wireless Communications*, 332–335, Hangzhou, China, August 2007.
6. Jung, J. W., H. H. Roh, J. C. Kim, H. G. Kwak, M. S. Jeong, and J. S. Park, "TX leakage cancellation via a micro controller and high TX-to-RX isolations covering an UHF RFID frequency band of 908–914 MHz," *IEEE Microwave and Wireless Components Letters*, Vol. 18, No. 10, 710–712, 2008.

Superluminal Phase Velocity in the Dispersive Media

Dexin Ye¹, Yuhua Wang², Shan Qiao², Jiangtao Huangfu¹, and Lixin Ran¹

¹Department of Information and Electronic Engineering, Zhejiang University
Hangzhou 310027, China

²Department of Information Engineering, Zhejiang University City College
Hangzhou, Zhejiang 310015, China

Abstract— We introduce a method to realize superluminal phase velocity of electromagnetic wave which can be five thousand times larger than c (c is the speed of light in vacuum) in dispersive uniaxial media. Theoretical analysis on dispersion relation for plasma media reveals that the phase velocity is much larger than C when the refractive index is smaller than that in vacuum at special frequency. Based on the theory, we designed a metamaterial construction which is dispersive uniaxial and exhibiting Drude-model-like frequency dependence to realize the superluminal phase velocity. We also verified it through the experiment.

1. INTRODUCTION

The velocity of electromagnetic wave propagating in the media has always captured the attentions of the researchers [1]. For most of the unartificial media, the phase velocity v_p is always smaller than c (the speed of light in vacuum). In order to get a superluminal phase velocity, we must get a kind of media with refractive index smaller than that in vacuum, such as plasma media around plasma frequency. Theoretical analysis from the K-surface reveals that the EM wave can propagate in plasma media with superluminal phase velocity. In this paper, we designed a kind of artificial material with the same property to plasma media at some frequency to realize the superluminal phase velocity.

Metamaterial is a sort of artificial structural composite possessing extraordinary electromagnetic (EM) properties, such as negative effective permittivity and/or permeability [2, 3], which have excited the imaginations of physicists and engineers in the past few years. Recently, metamaterials with “near-zero” permittivity and/or permeability have also attracted a lot of interest [4–7]. With the near-zero constitutive parameters and therefore the near-zero refractive index of the metamaterial, the corresponding wave vector will be near to zero. According to the formula of phase velocity $v_p = \omega/k$, the phase velocity will be large enough. In this paper, we show by theoretical analysis, simulation and experiments that the phase velocity will be much larger than c while the frequency is approaching to the frequency of near-zero refractive index.

2. THEORETICAL ANALYSIS

We consider a uniaxial media with a scalar permittivity $\varepsilon = \varepsilon_0$ and a permeability tensor

$$\bar{\bar{\mu}} = \mu_0 \begin{pmatrix} \mu_{rx} & 0 & 0 \\ 0 & \mu_{ry} & 0 \\ 0 & 0 & \mu_{rz} \end{pmatrix}$$

which is given in the principal coordinates. Here, we assume $\mu_{rx} = \mu_{ry} = 1$ and μ_{rz} obey the frequency dependence of Drude model as below

$$\mu_{rz} = 1 - \omega_p^2 / [\omega(\omega + i\gamma_e)],$$

where ω_p is the plasma frequency and γ_e is damping constant. Consider a TE-polarized line source in the middle of the media with wave vector $\vec{k} = \hat{x}k_x + \hat{z}k_z$, the dispersion relation in the proposed uniaxial media can be express as

$$k_x^2 / \mu_{rz} + k_z^2 = k_0^2 = (\omega/c)^2$$

We can find when μ_{rz} is positive and near to zero, the $k_x = \sqrt{\mu_{rz}((\omega/c)^2 - k_z^2)}$ will be much smaller than the k_0 , which means the wave propagates along x axis has a much larger phase velocity than that in vacuum. On the other hand, the wave will radiate in the z direction with high directivity in this case [8, 9]. Strictly speaking, the phase velocity will be infinity when μ_{rz} is zero. If we consider the loss of media when γ_e is not zero, the real part of k_x will be a small value but not zero, that means the phase velocity can't be infinity but a large enough value contrasting to the speed of light in vacuum if the media is not lossless.

3. SIMULATION RESULT

We use a practical metamaterial slab to verify the aforesaid analysis. On account of the limit of experiment equipment, we design a kind of slab to carry out the simulation and experiment in microwave region without losing universality. The slab we designed is constructed with the unit cell as shown in Fig. 1(b).

This kind of metallic pattern is first proposed in [10] for fabricating metamaterial showing effective negative permeability. In our realization, the metallic patterns are printed in alignment on a 1-mm-thick FR4 substrate with a relative permittivity 4.6. One piece of the metamaterial sample have 40 unit cells in the x direction with a periodicity of 10 mm and 15 unit cells in the y direction with a periodicity of 6 mm, respectively. Eight pieces of such sample are aligned along the z direction with an interval of 16 mm to obtain a slab-like sample. In such a slab, for EM wave incidence, a magnetic resonance can be induced by the resonant currents flowing along the metallic rings, yielding an effective negative permeability along the z direction but no magnetic resonances in the x and y direction, so μ_{rx} and μ_{ry} can be regarded as unit. We can get the equivalent parameter using homogenization approach proposed in [11, 12] as Fig. 2 shows. It is seen that the real part of μ_{rz} is negative over a frequency range from 9.2 GHz to 9.45 GHz and is equal to zero at 9.45 GHz.

Before the experiment, we perform finite-element method (FEM) simulation to verify the performance of the metamaterial slab. The dimensions and parameters are just the same as described previously, and the antenna is located in the center of the slab along the x direction. The simulation result is shown in Fig. 3 at the frequency of 9.45 GHz. We can see that the phase varies slow along x direction in the slab while fast along z direction which means that the phase velocity in x direction is much larger than the z direction, according to the formula $v_p = \omega/k = \omega/(\Delta\phi/\Delta d) = \omega\Delta d/\Delta\phi$ (Δd is the distance between two measure points and the $\Delta\phi$ is the difference of phase between the two points).

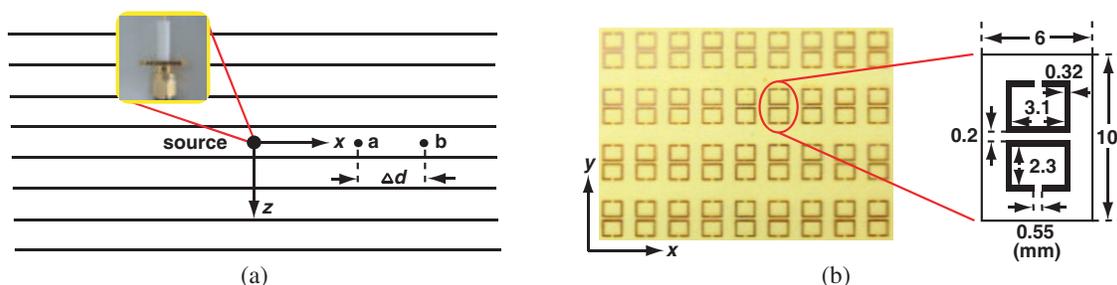


Figure 1: (a) The sketch map of experiment, point a and point b are the place which we choose to measure the EM-wave phase. (b) The experiment metamaterial slab and unit cell of the metamaterial sample.

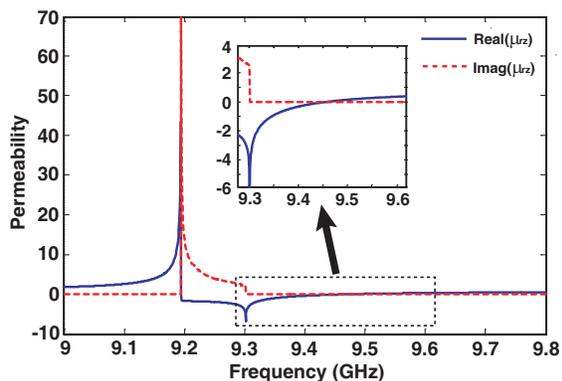


Figure 2: The retrieved real and imaginary parts of the permeability μ_{rz} .

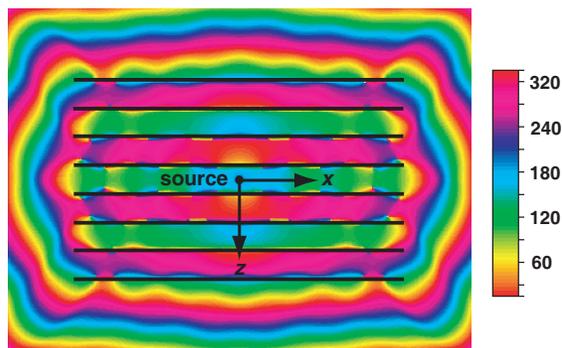


Figure 3: Simulation results of the phase of electric field distribution in the metamaterial slab at the frequency 9.45 GHz.

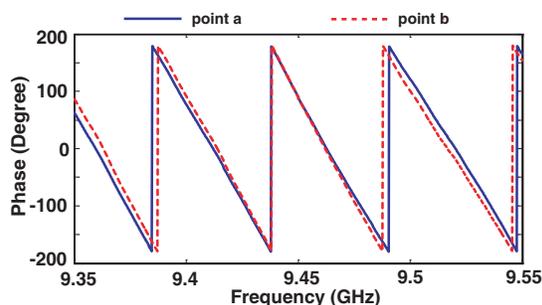


Figure 4: The phase of electric field at three different measure points from 9.35 GHz to 9.55 GHz.

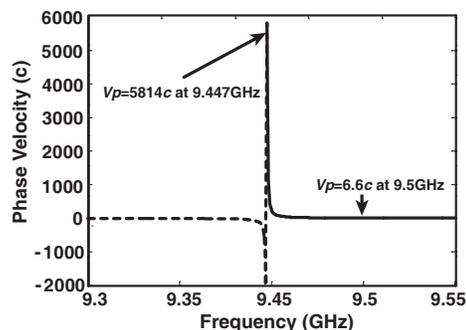


Figure 5: The phase velocity of electric field between point a and point b using the formula $v_p = \omega/k = \omega/(\Delta\phi/\Delta d) = \omega\Delta d/\Delta\phi$.

4. EXPERIMENT RESULT

The experiment is carried out in a microwave anechoic chamber. We put a monopole antenna in the center of the metamaterial slabs as a source and use another monopole antenna to measure the phase. Fig. 1(a) shows the sketch map photo of the experiment. During the experiment, a microwave signal from 9.35 GHz to 9.55 GHz is fed into the monopole. And we put a monopole antenna on two points (point a and point b) to measure the phase which are along with z axis and the distance between two points Δd is 12 mm as Fig. 1 shows. The result is showed as Fig. 4.

The two lines denote that the phase varies at different frequency at different point. But at 9.447 GHz, the two lines almost overlap. That means the phases are almost the same at the two points which denote the phase velocity is very large in x direction at the 9.447 GHz. Based on the formula $v_p = \omega/k = \omega/(\Delta\phi/\Delta d) = \omega\Delta d/\Delta\phi$, we calculate the phase velocity and the result is showed as Fig. 5.

When the frequency is below the 9.447 GHz, the permeability of the metamaterial slabs is negative, and the EM wave is evanescent. So the phase velocity is insignificant. At the 9.447 GHz, the largest phase velocity is almost 5814 times larger than the speed of light in vacuum, and when frequency moves toward higher frequencies, the phase velocity drastically decreased. As one example, at 9.5 GHz, the phase velocity is only 6.5 c and the phase velocity decreases as the frequency increase because that the permeability and k_x are increasing in this case. The variation trend of the phase velocity versus frequency clearly reflects a Drude-model-like dispersion, therefore we can conclude that the measured results fit well with theoretical analysis.

5. CONCLUSION

Through simulation and measurement, we can get a superluminal phase velocity of 5814 c in a metamaterial slab which consisting of periodically arranged split-ring resonators and exhibiting Drude-model-like frequency dependence along its primary axis in microwave frequency band. The experimental results fit well with theoretically analysis. The superluminal phase velocity along the x direction can be used to perfectly explain how the highly directivity antennas based on such uniaxial metamaterial work, and has potential in microwave applications such as in-phase wave divider and coherent wave sources. In optical band structural materials, such as in photonic crystals, we can also find lots of similar dispersion behaviors similar to Drude model the band gap diagrams, indicating that superluminal phase velocity can also be found easily in optics, and therefore can be used in optical applications. In the future, we will keep on this study and try to extend similar concept to two-dimensional which means that the EM-wave propagates with superluminal phase velocity in all direction in these materials.

ACKNOWLEDGMENT

This work is sponsored by NSFC (Nos. 60531020, 60671003 and 60701007), 863 Project (No. 2009A-A01Z227), NCET-07-0750, ZJNSF (No. Y1080715), ZJSTP (No. 2009C31141), Ph.D. Programs Foundation of MEC (No. 20070335120) and the National Key Laboratory Foundation (Nos. 9140C-5304020901 and 9140C5304020704).

REFERENCES

1. Brillouin, L., *Wave Propagation and Group Velocity*, Academic Press Inc., New York, 1960.
2. Veselago, V. G., "The electrodynamics of substances with simultaneously negative values of ϵ and μ ," *Sov. Phys. Usp.*, Vol. 10, 509, 1968.
3. Smith, D. R., W. J. Padilla, D. C. Vier, S. C. Nemat-Nasser, and S. Schultz, "Composite medium with simultaneously negative permeability and permittivity," *Phys. Rev. Lett.*, Vol. 84, 4184, 2000.
4. Ziolkowski, R. W., "Propagation in and scattering from a matched metamaterial having a zero index of refraction," *Phys. Rev. E*, Vol. 70, 046608, 2004.
5. Zhang, F., S. Potet, and J. Caubonell, "Negative-zero-positive refractive index in a prism-like omega-type metamaterial," *IEEE Trans. Microwave Theory Tech.*, Vol. 56, 2566, 2008.
6. Alu, A., M. G. Silveirinha, and N. Engheta, "Transmission-line analysis of epsilon-near-zero-filled narrow channels," *Phys. Rev. E*, Vol. 78, 016604, 2008.
7. Edwards, B., A. Alu, M. E. Young, M. Silveirinha, and N. Engheta, "Experimental verification of epsilon-nearzero metamaterial coupling and energy squeezing using a microwave waveguide," *Phys. Rev. Lett.*, Vol. 100, 033903, 2008.
8. Yuan, Y., L. Shen, L. Ran, T. Jiang, and J. Huangfu, "Directive emission based on anisotropic metamaterials," *Phys. Rev. A*, Vol. 77, 053821, 2008.
9. Jiang, T., Y. Luo, Z. Wang, L. Peng, J. Huangfu, W. Cui, W. Ma, H. Chen, and L. Ran, "Rainbow-like radiation from an omni-directional source placed in a uniaxial metamaterial slab," *Opt. Express*, Vol. 17, 7068, 2009.
10. Pendry, J. B., A. J. Holden, D. J. Robbins, and W. J. Stewart, "Magnetism from conductors and enhanced nonlinear phenomena," *IEEE. Trans. Microwave Theory Tech.*, Vol. 47, 2075, 1999.
11. Smith, D. R., S. Schultz, P. Markos, and C. M. Soukoulis, "Determination of effective permittivity and permeability of metamaterials from reflection and transmission coefficients," *Phys. Rev. B*, Vol. 65, 195104, 2002.
12. Chen, X. D., T. M. Grzegorzczuk, B.-I. Wu, J. Pacheco, and J. A. Kong, "Robust method to retrieve the constitutive effective parameters of metamaterials," *Phys. Rev. E*, Vol. 70, 016608, 2004.

Application of EH4 in the Shihu Gold Deposit of Western Hebei, China

Mingyan Wang^{1,2}, Tagen Dai³, Chaozhuang Xi³, Xiaoming Fu³, and Danyan Huang³

¹Institute of Mineral Resources of the Chinese Academy of Geological Science, Beijing 100037, China

²Department of Mineral Resources, Hunan Non-ferrous Metals Holding Group Co., Ltd
Changsha, Hunan 410015, China

³School of Geosciences and Environmental Engineering, Central South University
Changsha, Hunan 410083, China

Abstract— EH4 geophysical systems belong to electromagnetic survey method. As the most effective electromagnetic method for metalliferous mineral explorations with high detecting resolution and appropriate depths of penetration, EH4 is widely used in solid minerals, especially in finding concealed ore deposit. In this paper, the author adopted EH4 method to make a synthetic study on Shihu gold deposit in Hebei. The result showed that the geophysical method was quite useful in the exploration of concealed ore deposit.

1. INTRODUCTION

Shihu gold deposit is under the administrative divisions of Chenzhuang town Lingshou Country Hebei Province. And it is located 66 km to the west-north of Lingshou Country. The mining geographic coordinates are: Longitude 114°03'15"–114°04'21" and Latitude 38°39'04"–38°40'19". The mining area is 4.1 km².

Our predecessor had done a comprehensive detailed study in this area, and our focus was to give a depth and margin prediction on the 101 ore belt, which was a kind of local metallogenic prediction compared with the regional one.

The 101 ore belt in Shihu gold deposit was stick controlled by faulted and shattered zone, and almost all of the ore body was limited to wide range of less than 100 m. It had a high degree of alteration in faulted and shattered zone. It was Cataclasite of high degree silicification or filled by quartz sulfide ore body, between which there was obvious physical difference with Archaean-group biotite plagioclase gneiss, homblende-biotite plagioclase gneiss and amphibolite

The methods such as electrical method, magnetic method, earthquake method and EH-4 high frequently mt sounding were in the choice of geophysical methods. Our main aim of this research was to find out the downward extension of the 101 ore belt. The mining had reached to 500 m, which the conventional geophysical methods were difficult to survey. The EH-4 high frequently mt sounding, with high-precision, large-depth and strong anti-jamming, was our choice, which achieved good results on exploration in the resources crisis mine in recent years.

2. FIELD WORK METHOD AND TECHNOLOGY OF EH-4 SYSTEM

- (1) Choice of operating frequency: The gneiss, with resistivity more than 1000 Ω·m, was the main bedrock in this project area. In order to meet the design requirement, we chose 10 Hz, the lowest operating frequency while detecting. Three frequency groups like the second, the third and the fourth were in our choice, ranging from 10 Hz to 99 kHz.
- (2) Arrangement of observation points: The geophysical section line of survey area in Shihu gold deposit was located by our study team. We adopted GPS position and fix distance with measuring wire.
- (3) Systematic inspection: The parallel test was carried out the day before we began. It was required that the two magnetic bar parallel on the ground, 2–3 m apart. The two electric dipole were arranged parallel, and be vertical to the bar. We observed the time series signal of electric field and magnetic field channels to make sure if it works properly.

3. DATA PROCESSING METHOD

The time series collected in field were pretreated firstly, and then changed to FFT. The imaginary and real component, and the phase data were obtained through it. The one-dimensional BOSTICK inversion was performed to transform from frequency to depth. Based on one-dimensional inversion,

the two-dimensional electromagnetic was imaged fast and automatically by the two-dimensional imaging software used abroad.

The data obtained by inversion was done Kringing gridding in the XZ plane. Meeting the actual situation of electrical anisotropy, X -axis and Y -axis were to taken anisotropy method to search radius.

4. WORKLOAD COMPLETED AND QUALITY ASSEMENT

Choice of the Investigation Profile

Only three representative profiles were arranged this time because of high investigation cost (See Figure 5(b)). They were Line 19, Line 29, and Line 35 respectively. The workload completed was list in Table 1 in details.

Table 1: The workload table of survey in Shihu gold deposit.

Survey-Line Number	Points Number	Dot Pitch (m)	Polar Distance (m)	Profile Length (m)
Line 19 Section	26	20	20	500
Line 29 Section	26	20	20	500
Line 35 Section	26	20	20	500
Total observation points	78			

The profile layout obeyed the following three principles. Firstly, it can basically reflect deep extension in the production midcourse of 101 ore belt. Secondly, it can reflect the extension of the southern section of the ore belt roughly. Thirdly, it was favorable for verifying with pitting and drilling in near future. Base on these principles, the Line 19, Line 29, and Line 35 were layout.

5. ANOMALY INTERPRETATION

In general, the gold-bearing quartz veins showed high resistance to electrical conductivity, but the ones in this area were of low resistance due to the strong water abundance in the faulted and shattered zone. And the quartz vein-type of gold ore bodies in this area were almost hosted in faulted and shattered zone, which was also of low resistance.

5.1. Anomaly Interpretation of Line 19

Line19, with 500 meters length of plant profile and 20 meters of dot pitch, had a total of 26 measuring points. It can be recognized the following understanding.

Firstly, it was supposed that faulted and shattered zone named F1 was located between the surface No. 260 measuring point and No. 280. The faulted and shattered zone extended down from the surface (the elevation of 710 m) to the surface No. 560 measuring point (the elevation of 50 m), with the inclination of about 40° . It was found out that the resistance curve trended to be downward through analysis, and was likely to down to 50 meters beneath the No. 620 measuring point.

Secondly, there existed a wide range of low-resistance zone on both sides of the fault in the range of elevation from 650 m to 350 m that faulted and shattered zone named F1 passed, indicating the fault fracture zone here was of intensive or on large-scale.

Thirdly, the faulted and shattered zone named F2 was located between the surface No. 470 measuring point and No. 500, and continued to trend to down. It can reach the elevation of 250 meter of No. 700 measuring point from the analysis for resistivity contour.

Fourthly, the faulted and shattered zone named F3, with a vertical occurrence, was located on the surface No. 620 measuring point. It downed to the 400-meter elevation in No. 600 measuring point, and was likely to extend to 400-meter elevation to intersect with F2.

5.2. Anomaly Interpretation of Line 29

Line 29, with 500 meters length of profile and 20 meters of dot pitch, had a total of 26 measuring points. It can be recognized the following understanding.

Firstly, the faulted and shattered zone F1 was located in the vicinity of No. 220 measuring point and it was low resistance, with no less than $2000\ \Omega\cdot\text{m}$, on the negative 150-meter elevation down to the No. 400 measuring point, which showed that F1 was of more development. It was likely to be a band for two types of rocks with different resistance on the 300-meter elevation down. The

resistance on the left side (or the west side) of the rocks was low, while the ring side (or the east side) was high.

Secondly, the faulted and shattered zone F2 was between No. 360 measuring point and No. 400, and extended down to the 450-meter elevation of the No. 420. There existed large difference between the faulted and shattered zone, with the resistance less than $2500 \Omega\cdot\text{m}$, and its surrounding rocks on both sides, with the resistance more than $5000 \Omega\cdot\text{m}$, from which it was concluded that the faulted and shattered zone F2 had a high fragmentation degree and fracture development.

Thirdly, the faulted and shattered zone F3, located near the No. 560 measuring point, had no obvious surface characteristics, mainly in the 575-meter elevation to 675-meter elevation of measuring point from No. 540 to No. 560, away 150 meters to surface depth. Although the zone was shallow, it had high degree of fragmentation. As a hidden zone, this faulted and shattered zone should be paid attention to be examined.

5.3. Anomaly Interpretation of Line 35

Line 35, with 500 meters length of profile and 20 meters of dot pitch, had a total of 26 measuring points. It can be recognized the following understanding.

Firstly, the faulted and shattered zone F1, located near surface No. 220 measuring point, extended from the surface to the 500-meter elevation of No. 260 measuring point and thinned out. It was on a small scale, with angle of about 40°C .

Secondly, the faulted and shattered zone F2, located near No. 340 measuring point, extended from the surface to the 170-meter elevation of No. 620 measuring point. There formed a low resistance zone due to the low resistance from the 700-meter elevation to 550-meter elevation of the passed measuring point from No. 320 to No. 400. Another low resistance zone was formed from the 350-meter elevation to 470-meter elevation of the measuring point from No. 440 to No. 560. These two low resistance zones may be a development area for the faulted and shattered zone, or a secondary zone.

Thirdly, the vertical faulted and shattered zone F3, on a small scale, located near No. 700 measuring point, extended to the 550-meter elevation of No. 620 measuring point. The east side of the zone is unknown due to the location or the different lithologic contact zone.

6. CONCLUSION

It was the first time to apply EH-4 electrical conductivity imaging system, and take high-frequency magneto telluric sounding method in deep part ore prediction for the mining area in Shihu gold deposit. It was identified that distribution and extension of the deep structure in 101 ore belt, and pointed out the favorable ore-forming parts. Then quantitative imputation for its depth was made. All the interpretation indicated that as following. In the deep part of Line 19 and Line 29 of the mining area, the 101 ore belt had been extended to the elevation from -50 m to -150 m . The faulted and shattered zone was in development of large-scale in Line 35 which in the southern of 101 ore belt, and it showed that 101 ore belt continued to extend southward without thinning out, with stable extension from the shallow to the deep. The deep extension was up to above 50-meter elevation. However, this method reflected vaguely to the information in the shadow surface and generally had low resistivity, which could be explained by strong interference of surface natural field or a better aquosity in near-surface, so the upper boundary of ore belt was fixed above 400 meters in metallogenic prediction in the sixth chapter. In short, the geophysical results basically coincided with the geological analysis, which provided reliable evidence for predicting the depth and margin boundary of 101 ore belt.

ACKNOWLEDGMENT

This paper is supported by National “973” program (2002CG412607) and National Science Foundation Project of Hunan Province (07JJ6071).

REFERENCES

1. Zhang, Y. X., S. L. Chen, and S. L. Peng, “Study on the gold ore structural controlling over TS ore-field,” *Geotectonical Et Metallogenia*, Vol. 20, No. 1, 71–80, Lingshou, Hebei, 1996.
2. Chen, G. D., *China Metallogenic Geotectonic Map*, Central South University of Technology Press, Changsha, 1999.
3. Yang, D. F. and G. S. Li, “Study on the ore-forming conditions and genesis of Tuling and Shihu gold deposit in the Taihang mountain region,” *Journal of Changchun University of Earth Science*, Vol. 21, No. 1, 47–53, 1991.

4. Yang, D. F., G. S. Li, and K. S. Jia, “The analysis of constraint character to ore deposit and forecasting of ore-forming in Shihu gold mine,” *Journal of Hebei College of Geology*, Vol. 15, No. 5, 516–520, Hebei, 1992.
5. Yang, D. F., R. F. Liu, and G. S. Li, “The geochemical characteristics of the Tuling and Shihu gold mines in the Taihangmountain area and the indicators for prospecting for mineral deposits,” *Geology of Jilin*, Vol. 4, 70–78, 1991.
6. Han, J. C., “Analysis structure controlling condition of Tuling-Shihu gold depositin,” *Gold*, Vol. 18, No. 8, 7–10, Hebei, 1997.
7. Chen, J. R., “Geological characteristics and genesis of gold deposit in Tangxian Dashiyu and peripheral region,” *Hebei Gold Geological Science and Technology*, Vol. 34, No. 4, 21–26, 1992.
8. Cui, Y. H., “Chemical composition of pyrite from the Tuling-Shihu gold district and its genetic significance,” *Acta Petrologic et Mineralogic*, Vol. 12, No. 4, 371–381, 1993.
9. Xi, C. Z., T. G. Dai, and W. Liu, “Geological-geochemical characteristics of the Shihu gold deposit in western Hebei Province,” *Acta Geoscientica Sinica*, Vol. 29, No. 4, 451–458, Aug. 2008.
10. Xi, C. Z., T. G. Dai, W. Liu, and H. J. Zhang, “Petrogeochemical characteristics of the intrusive bodies of mapeng granitoids in western Hebei,” *Acta Petrologica Et Mineralogica*, Vol. 27, No. 2, 113–120, Mar. 2008.

An Optimized Monopole Microstrip Patch Antenna with Gradual Steps for Ultrawideband Applications

Reza Khalilpour¹, Javad Nourinia², and Changiz Ghobadi²

¹Telecommunication Company of Iran, Iran

²Department of Electrical Engineering, Urmia University, Iran

Abstract— An optimized monopole microstrip Patch Antenna for UWB (Ultra Wideband) Applications is designed. The proposed antenna is fed by a single microstrip line and designed to operate from 2 to 14 GHz. It consists of a patch with gradual steps, a single slot on the patch, and a partial ground plane with special slit. Experiments are carried out to investigate its return loss and its radiation, which exhibit good radiation pattern and impedance bandwidth over the entire band of frequency.

1. INTRODUCTION

A UWB antenna should provide a gain and impedance that is relatively constant for bandwidth from 3.1 to 10.6 GHz. For dominant applications [1], these antennas must be compact, low priced, and must present omnidirectional radiation patterns and ease of fabrication. The demand for ultra wideband systems is increasing rapidly in the communication systems, particularly wireless multimedia system with high data rate. The technique of a UWB wireless communication is to transmit digital information signals transformed into impulse or nonsinusoidal signals with an extremely short pulse below nanoseconds. This technique is able to communicate very fast in hundreds Mbps like an optical communication.

The microstrip patch antenna may be a good choice as it is compact, low profile and easy to integrate with monolithic microwave integrated circuits (MMICs). The various types of UWB antennas, such as the planar volcano-smoke slot antenna, bowtie patch antenna, and the modified bowtie antenna with a triangular shape, have been developed for UWB systems. For many years, ultrawide band (UWB) antennas have had many applications in communication systems with broadband and spread-spectrum features in radar systems [2]. Guillanton et al. [3] proposed a new balanced antipodal Vivaldi antenna for UWB applications with ultrawide bandwidth (1.3–20 GHz) with respect to 10-dB impedance. Recently, a rectangular patch with two steps and a single slot on the patch, and a partial ground plane has been designed for UWB applications [4], and the transient responses of a logarithmic periodic dipole and a Vivaldi antenna are presented for the characterization of a UWB antenna [5]. Many researches have been attempted to widen the bandwidth of the conventional printed antennas. Although the bowtie antenna is one of the promising techniques for UWB antenna systems [6], it is not sufficient to cover the UWB frequency band. Thus, various structures have been proposed to overcome the narrow bandwidth. Parasitic elements around the antenna bring about broad bandwidth operation, but increase the size of the antenna [7].

In this paper, we propose an optimized ultrawide band antenna for UWB applications. The proposed monopole antenna consists of a patch with complex of gradual increasing and gradual decreasing steps, a single rectangular slot on the patch, and a partial ground plane with special slit with improvement in impedance and good gain bandwidth. Investigations based on experiments and simulations are conducted. The simulation is performed using the commercially available simulation software HFSS 11. The proposed antenna is successfully implemented and the simulated results show reasonable agreement with the measured results. In this design, a 3.2 ~ 14 GHz frequency range for $S_{11} < -10$ dB is obtained. Radiation patterns are also examined.

2. ANTENNA GEOMETRY AND DESIGN

Printed monopole antennas radiate in broadside direction. To meet the demands of a reduced antenna size, a higher dielectric substrate is required. The FR substrate with a dielectric constant of 4.4 can be used for this application. Patch antennas exhibit narrowband characteristics making it challenging to enhance their bandwidth. To account for this a novel antenna shape has been designed, shown in Figure 1, which is a tradeoff between size, bandwidth and substrate type. The proposed antenna is fed by single microstrip line and designed to operate from 2 to 14 GHz. It

consists of a patch with gradual steps, a single slot on the patch, and a partial ground plane with the special slit. The antenna, which has dimensions of $30 \times 36 \text{ mm}^2$, is printed in the substrate FR4 of thickness 1.6 mm and relative permittivity 4.4. The dimension of the slot is $0.9 \times 10.5 \text{ mm}^2$ and the dimension of the ground plane is chosen to be $30 \times 11 \text{ mm}^2$ in this study. The excitation is a 50Ω microstrip line printed on the slitted partial grounded substrate. To design the UWB antenna, we have applied four techniques to the proposed antenna: the use of (i) gradual steps, (ii) a partial ground plane, and (iii) a single slot on the patch, and (iv) special slit on ground plane which can lead to a good impedance matching. By selecting these parameters, the proposed antenna can be tuned to operate in the $3.2 \sim 14 \text{ GHz}$ frequency range.

The proposed antenna was constructed and studied. Using an Agilent Technologies, (N5230A) Analyzer, the measured return loss of the proposed antenna is obtained at $2 \sim 14 \text{ GHz}$, as depicted in Figure 2. Figure also shows the simulated results from HFSS 11 for comparison. The measured return loss S_{11} reasonably agrees with the results. A $3.2 \sim 14 \text{ GHz}$ frequency range below -10 dB of S_{11} is obtained.

3. SIMULATION RESULTS FOR THE MONOPOLE ANTENNA

The proposed antenna fed by a microstrip line is shown in Figure 1, which is printed in a FR4 of thickness 1.6 mm and relative permittivity 4.4. The microstrip line width T_S is 2.5 mm for 50Ω impedance, and monopole patch with gradual steps is connected to the microstrip line. The special slit of width X on the groundplane, and the rectangular slot of width W_2 are introduced. By optimizing these parameters, a compact antenna is produced. Simulations are performed using the HFSS 11 simulation software.

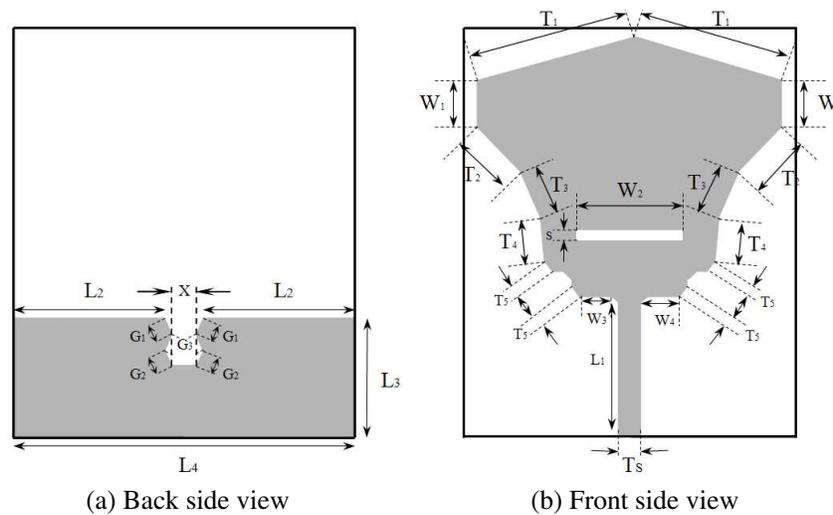


Figure 1: Geometry of the proposed antenna.

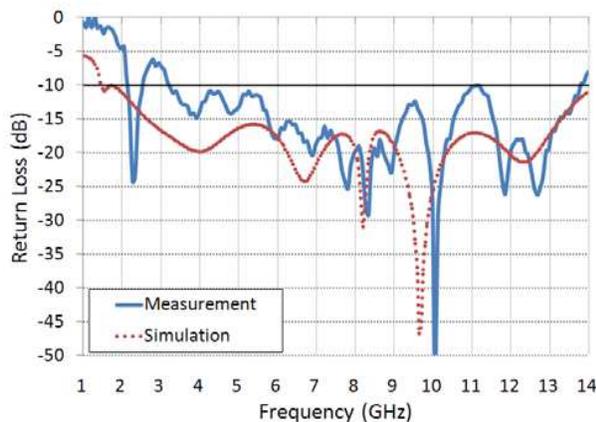


Figure 2: Measured and simulated return loss for the proposed antenna.

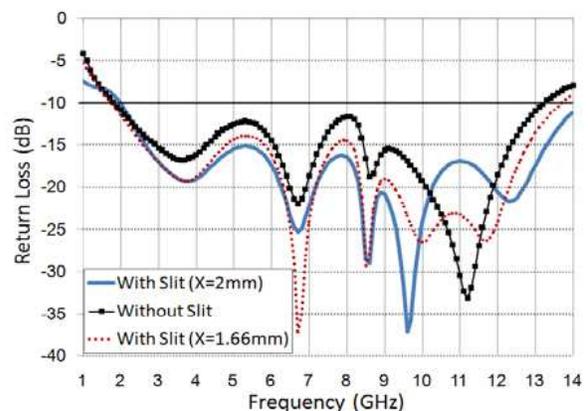


Figure 3: Simulated return loss for different values of slit dimensions X for comparison.

3.1. The Effect of the Slit in the Ground Plane

Another technique to broaden the bandwidth is to adjust the width of the slit near the starting point of the radiating element. To compact the proposed UWB antenna and increase the impedance bandwidth, a slit is introduced into groundplane to alter the input impedance characteristics. As shown in Figure 3 the small changes of slit width have significant effects on the bandwidth of the proposed antenna. It is found that the return loss of lower-edge frequency will decrease when the slit width $X = 1.66$ mm, and the return loss also decreases with increasing slit width. The width of the slit can be used to extend the return loss of lower-edge frequency or the upper-edge frequency of the impedance bandwidth.

The upper-edge frequency without slit is 11.20 GHz and with slit is 9.65 GHz. This is a useful extension of the upper-edge frequency. The optimized slit width X was found to be 2 mm, which provides a very wide bandwidth. The slit also influences the bandwidth of the antenna.

3.2. The Effect of the Gradual Steps

The proposed antenna with eight steps instead of a rectangular shape is shown in Figure 1(a). We can find the tendency of a return loss value regard to the geometry of the antenna. The gradual steps are shaped with eight steps in the proposed antenna compared with the return loss results of the patch antenna without them, and the modified antenna brings about more frequency resonances and an improvement of an impedance matching than in case of the staircases with 0 and 8 steps. Therefore, it achieves enhanced bandwidth as a result as shown in Figure 4.

3.3. The Effect of the Ground Plane Size

The groundplane dimensions are very important parameters in the design of these antennas, because of strong dependence of bandwidth on groundplane size. Figure 5 illustrates the return loss for different values of groundplane length L_3 without slit ($L_4 = 30$ mm). It is seen that the bandwidth is heavily dependent on groundplane width. To design a compact UWB antenna, the groundplane L_4 is selected as 30 mm. It was also found that the return loss is sensitive to the groundplane length L_3 . In this paper, the L_3 is selected as 11 mm.

To understand the behavior of the antenna model and obtain the optimum parameters, simulations were performed with the HFSS software. The optimized values of each physical dimension of the proposed antenna are shown in Figure 1 and Table 1.

With the dimensions of the Table 1, a prototype antenna is simulated by HFSS software and is

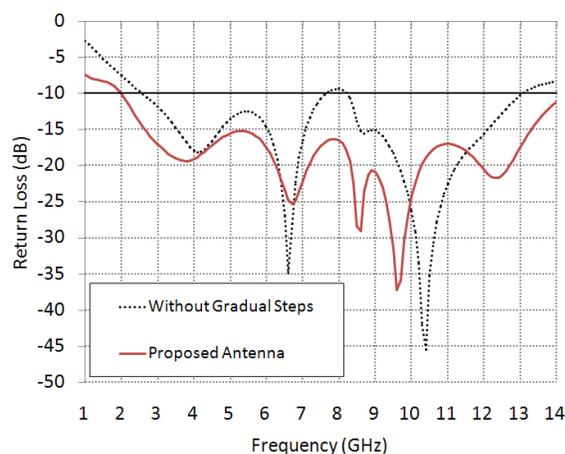


Figure 4: Simulated return loss with gradual steps and without them for comparison.

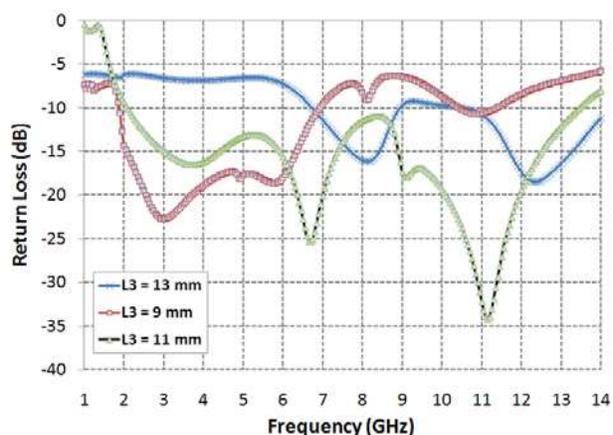


Figure 5: Simulated return loss for different values of groundplane width (without slit).

Table 1: The dimensions of the proposed antenna. All dimensions are in millimeters.

Parameter	L_1	S	W_1	W_2	W_3	W_4	T_1	T_2	T_3
Value	12	0.9	4	10.5	2.9	3.1	14.56	5.657	4.752
Parameter	L_2	L_3	L_4	G_1	G_2	G_3	T_4	T_5	T_S
Value	13.5	11	30	1.424	1.393	1.269	3.769	0.938	2.5

also fabricated. These dimensions are obtained after performing an optimization. The measured and simulated results are shown in Figure 2. As shown in Figure 2, there exist some discrepancies between measured data and the simulated results. In order to achieve the accurate return loss characteristics for the designed antenna, it is recommended that the manufacturing and measurement process need to be performed carefully.

Figure 7 and Figure 8 show the measured co- and cross-polarized radiation patterns across the entire frequency band in the H -plane (xz -plane) and the E -plane (yz -plane) respectively. The H -plane radiation patterns are relatively omnidirectional, especially at lower frequency bands. However, the radiation patterns start to change in high frequencies and show higher directivities in other directions. The resonant frequency of a patch antenna mainly depends on relative permittivity of substrate material ϵ_r , and the length of the patch. The width of the patch will not affect the resonant frequency to a large extent but it will have some effect on the impedance of the patch

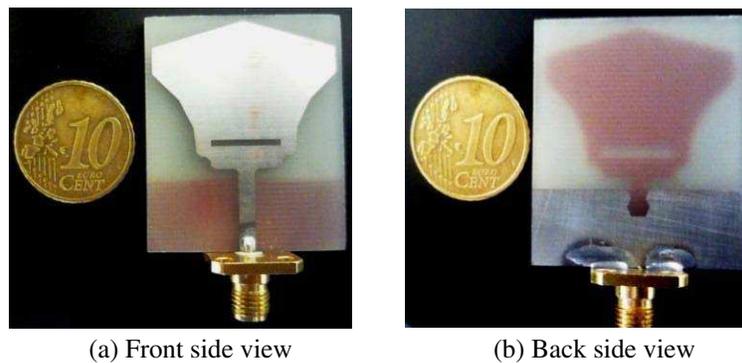


Figure 6: Photograph of the realized antenna.

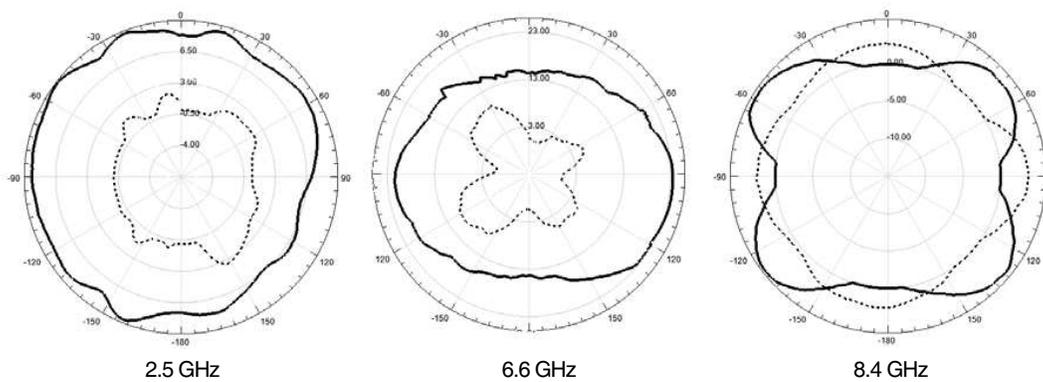


Figure 7: Measured radiation pattern of proposed antenna in H -plane ($\varphi = 0^\circ$). The solid line is co-polar and the dash line is cross-polar component.

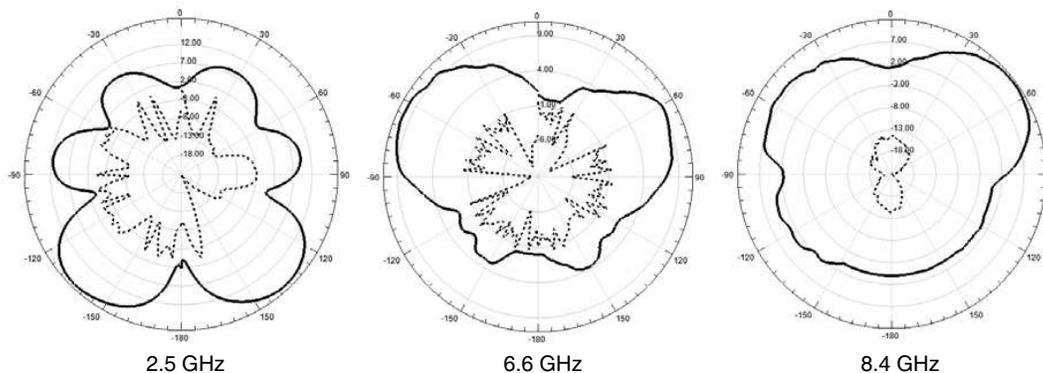


Figure 8: Measured radiation pattern of proposed antenna in E -plane ($\varphi = 90^\circ$). The solid line is co-polar and the dash line is cross-polar component.

antenna, and thus the bandwidth of the antenna. In high frequencies (i.e., from 7 GHz onwards), it seems that the patch of antenna length almost equals to the half wavelength and starts to introduce spurious radiation.

Figure 6 presents the photograph of a realized antenna on a FR4 substrate with SMA connector.

4. CONCLUSIONS

A new small UWB antenna has been designed, simulated, measured and fabricated. The proposed antenna with different width of slit, gradual steps and slot have been studied. The measured and simulated return loss of the proposed antenna suggests that this antenna is a potential UWB candidate, that can be used both in lower band as well as the upper band for pulse based UWB technology. From the simulated radiation patterns it can be predicted that this antenna strongly radiates in the elevation plane perpendicular to the antenna radiating element. This kind of radiation, which has nulls in the direction of $\theta = 90^\circ$ and $\theta = -90^\circ$ (Broadside radiation).

ACKNOWLEDGMENT

The authors are thankful to Iran Telecommunication Research Center (ITRC) for its financial support and also the Antenna Lab of the Khaje Nasir Toosi University of Technology (Tehran, Iran) where the proposed antenna has been tested.

REFERENCES

1. "FCC Report and Order for Part 15 Acceptance of Ultra Wideband (UWB) Systems from 3.1–10.6 GHz", Washington, DC, 2002.
2. Young, J. and L. Peters, "A brief history of GPR fundamentals and applications," *Proc. 6th Int. Conf. Ground Penetrating Radar*, 5–14, Sendai, Japan, 1996.
3. Guillanton, E., J. Y. Dauvignac, C. Pichot, and J. Cashman, "A new design tapered slot antenna for ultra-wideband applications," *Microwave Opt. Technol. Lett.*, Vol. 19, 286–289, 1998.
4. Bories, S., C. Roblin, and A. Sibille, "Ultra-wideband monocone antenna for UWB channel measurements," *XXVIII URSI Conv. Radio Sci. and FWCW Mtg.*, University of Oulu, Finland, 2003.
5. Sorgel, W., C. Waldschmidt, and W. Wiesbeck, "Antenna characterization for ultra-wideband communications," *XXVIII URSI Conv. Radio Sci. and FWCW Mtg.*, University of Oulu, Finland, 2003.
6. Eldek, A. A., A. Z. Elsherbeni, and C. E. Smith, "Wide-band modified printed bow-tie antenna with single and dual polarization for C- and X-band applications," *IEEE Transactions on Antennas and Propagation*, Vol. 53, No. 9, 3067–3072, Sept. 2005.
7. De Jean, G., R. L. Li, J. Laskar, and M. M. Tentzeris, "Circularly polarized loop antennas with a parasitic element for bandwidth enhancement," Georgia Institute of Technology, Atlanta, GA 30332-0250, USA, 2005.
8. Choi, S. H., J. K. Park, S. K. Kim, and J. Y. Park, "A new ultra-wideband (UWB) antenna for UWB applications," Microsystem Group, Materials and Devices Lab, LG Electronics Institute of Technology, Wiley Periodicals, Inc., 2004.
9. Kim, K.-H., Y. J. Cho, S.-H. Hwang, and S.-O. Park, "A band-rejected UWB planar monopole antenna with a ring-shaped parasitic patch," *Microwave Conference Proceedings, APMC 2005, Asia-Pacific Conference Proceedings*, 2005.
10. Cho, Y. J., K. H. Kim, D. H. Choi, S. S. Lee, and S.-O. Park, "A miniature UWB planar monopole antenna with 5-GHz band-rejection filter and the time-domain characteristics," *IEEE Transactions on Antennas and Propagation*, Vol. 54, No. 5, May 2006.
11. Akdagli, A., C. Ozdemir, and S. Yamacli, "A review of recent patents on ultra wide band (UWB) antennas," *Electrical Engineering*, Vol. 1, 68–75, Bentham Science Publishers Ltd., 2008.

Utilization of Effective Apparent Resistivity in Magnetotelluric Data Processing and Interpretation

Ai-Yong Li, Jian-Xin Liu, Xiao-Zhong Tong, Wei Zhang, and Chuang-Hua Cao
School of Info-Physics Geomatics Engineering, Central South University, Changsha 410083, China

Abstract— Effective apparent resistivity is correspondent to the modulus of magnetotelluric response impedance tensor matrix. It is a invariable under the coordinate rotation. Under one dimensional (1-D) condition, It is equal to the normal apparent resistivity ρ_a , when to two dimensional condition, it is the geometry average of apparent resistivity of TE mode (ρ_{TE}) and TM mode (ρ_{TM}), which has the dimensionality reduction property. The parallel moving algorithm can used to do static correction, as effective apparent resistivity curl shape was not affected by the static effect. At the same time, it is irrelative to coordinate rotation, so the inverse deviation can be avoided in the effective apparent resistivity two-dimensional (2-D) inversion, because of the incorrect polarization mode discrimination. So the utilization of effective apparent resistivity in processing and interpretation of the magnetotelluric sounding data was recommended.

1. INTRODUCTION

With the improving of the performance of instrument hardware system and the designing thoughts, especially applying GPS to the time synchronization of signal acquisition, Magnetotelluric sounding method (referred to as MT) field data collection makes remote reference technology into reality and greatly improved the quality of field data [1]. In the aspect of data processing and interpretation, The introduction of Robust statistical methods [2], perfecting and maturity of one-dimensional and two-dimensional inversion method as well as development of three-dimensional forward and inversion method continuously improving the accuracy of data interpretation [3].

There are still many problems of MT method to be solved, such as the static correction and problems of multiple solutions in the inversion process, especially that geological conditions encountered in the actual MT exploration are different and complicated, not as simple and stand as the assumed conditions in the theoretical study. Such theoretical assumptions (such as 1D and 2D conditions) does not correspond exactly with the real geological conditions, which will deviate the interpretation results, or even get wrong results. The most common problem that we will encounter in the actual exploration is: MT survey line in the field is often not completely vertical to the direction of geological structure, at this time, different inversion results will be achieved by two-dimensional inversion of ρ_{TM} and ρ_{TE} in principal axis and inversion of ρ_{xy} and ρ_{yx} in the direction parallel or vertical to the survey line.

For the problems of MT methods, although MT scholar at home and abroad have done a great deal of research work, some problems still have not been fundamentally solved, which can not be avoided in actual data processing and interpretation. Processing and interpretation staffs must take different measures to solve these problems. Since different person may use different interpretation methods, using the same inversion software for the same raw data may lead to different results, not to mention the further geological inference. Therefore, how to effectively weakening various effects appears to be meaningful and practical. The effective apparent resistivity method discussed in this article is the attempt in this aspect.

2. EFFECTIVE APPARENT RESISTIVITY

In magnetotelluric sounding exploration, after time-frequency Fourier transformation along electromagnetic fields of two orthogonal directions measured on the ground, the relationship between electric field components (E_x , E_y) of a certain frequency (f) and magnetic field components (H_x , H_y) is as following:

$$\begin{aligned} E_x &= Z_{xx}H_x + Z_{xy}H_y \\ E_y &= Z_{yx}H_x + Z_{yy}H_y \end{aligned} \quad (1)$$

In matrix form:

$$\mathbf{E} = \mathbf{ZH} \quad (2)$$

where:

$$\begin{aligned}\mathbf{E} &= (E_x, E_y)^T \\ \mathbf{H} &= (H_x, H_y)^T \\ \mathbf{Z} &= \begin{pmatrix} Z_{xx} & Z_{xy} \\ Z_{yx} & Z_{yy} \end{pmatrix}\end{aligned}$$

Z is impedance tensor which is the magnetotelluric response of measured station, Z_{xx} , Z_{xy} , Z_{yx} and Z_{yy} are impedance elements which are all complex. Two apparent resistivity curves ρ_{xy} and ρ_{yx} given by field real-time processing are as follows:

$$\begin{aligned}\rho_{xy} &= \frac{1}{5f} |Z_{xy}|^2 \\ \rho_{yx} &= \frac{1}{5f} |Z_{yx}|^2\end{aligned}$$

where: f is frequency.

Effective apparent resistivity (ρ_{eff}) is defined as [4]:

$$\rho_{eff} = \frac{1}{5f} |Z| \quad (3)$$

$|Z|$ is the matrix of impedance tensor matrix Z , there is:

$$|Z| = |Z_{xx}Z_{yy} - Z_{xy}Z_{yx}| \quad (4)$$

impedance phase (φ_{eff}) is:

$$\varphi_{eff} = 0.5 \times \arg(Z_{xx}Z_{yy} - Z_{xy}Z_{yx}) \quad (5)$$

From (4) and (5), we can see that four impedance tensor elements are involved in the calculation of efficient apparent resistivity, which are three-dimensional parameters.

Since the impedance elements relates to the direction of measuring coordinates, and different coordinates direction will arrive at different impedance element values of. Based on different purposes, Z should be rotated when doing data processing. After rotation, the relationship between impedance tensor Z_θ and Z is:

$$Z_\theta = R(\theta)ZR(\theta)^T$$

where:

$$R = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}$$

$R(\theta)^T$ is the transposed matrix for $R(\theta)$ of the, θ is rotation angle.

Impedance elements changed after rotating of impedance tensor, but the matrix of impedance tensor is unchanged, which is: $|Z| = |Z_\theta|$

It shows that effective apparent resistivity is a rotation invariant.

In 1D condition, $Z_{xx} = Z_{yy} = 0$, $Z_{xy} = Z_{yx} = Z_{1D}$, so:

$$\rho_{eff} = \rho_{1D} \quad (6)$$

Therefore, under 1D condition, conventional apparent resistivity is a special case of effective apparent resistivity.

When the geological structure is consistent with 2D condition, rotate Z to the principal axis, $Z_{xx}(\theta) = Z_{yy}(\theta) = 0$, there is:

$$\begin{aligned}\mathbf{Z}_{2D} &= \begin{pmatrix} 0 & Z_{TE} \\ -Z_{TM} & 0 \end{pmatrix} \\ \rho_{eff} &= (\rho_{TE} \times \rho_{TM})^{1/2}\end{aligned} \quad (7)$$

Under 2D condition, the effective apparent resistivity ρ_{eff} is the geometric mean of ρ_{TE} and ρ_{TM} .

3. APPLICATION OF DATA PROCESSING

Two-dimensional (2D) inversion technique of magnetotelluric sounding data is becoming more perfect and mature at present, and can be practically used. Since the actual ground conditions complex, there are some practical problems to be considered when applying 2D inversion method into data processing and interpretation, have to the. The first problem is identification of TE and TM mode. 2D inversion is to fitting ρ_{TE} and ρ_{TM} , after principal axis rotation, the measured apparent resistivity ρ_{xy} , ρ_{yx} should be decided which is ρ_{TE} and which is ρ_{TM} . Since the direction of principal axis may be the same with or vertical to the geological structure, it is sometimes very difficult to determine. If wrongly discriminated the inversion results will be inevitably affected. The second problem is that the principal axis is mostly a curve changing with, the difference between principal axis with high and low frequency can even reach to nearly 90 degrees, which indicates that the strike direction of local structures is inconsistent with regional. This is a normal geological phenomena, but 2D inversion of such data will bring bias, because at high-frequency it is TE mode, but low-frequency is TM mode. The third problem is the MT survey line can not be completely orthogonal with geological structure, but forward calculation during 2D inversion is on theoretical apparent resistivity and impedance phase in two directions parallel or vertical to the survey line while the fitting object is apparent resistivity and impedance phase in another direction (principal axis), which is obviously unreasonable. The above-mentioned problems are caused by the uncertainty of principal axis, and the former text also mentioned that effective apparent resistivity is a rotation invariant, so to do 2D inversion directly on effective apparent resistivity can avoid the above problems.

Figure 2 is an inversion and forward inversion example when survey line and structure strike is not perpendicular to each other, the survey line and positive and structure intersects with an angle of 45 degrees. Figure 2(a) is the profile of geoelectric section in the survey line, after calculating the theoretical response of the model on the survey, then do 2D inversion on different apparent resistivity formed, ρ_{xy} and ρ_{yx} are apparent resistivity parallel and vertical to the survey line respectively, ρ_{TE} and ρ_{TM} are apparent resistivity in the direction of principal axis, Figures 2(b)–(d) are inversion resistivity sections of ρ_{xy} , ρ_{yx} , ρ_{TE} and ρ_{TM} respectively, and clearly inversion results of ρ_{eff} has the highest accuracy.

On the other hand, in 2D inversion, when the above problems exist, 2D conditions are not satisfied, so fitting two apparent resistivity curve at the same time may achieve bad result.

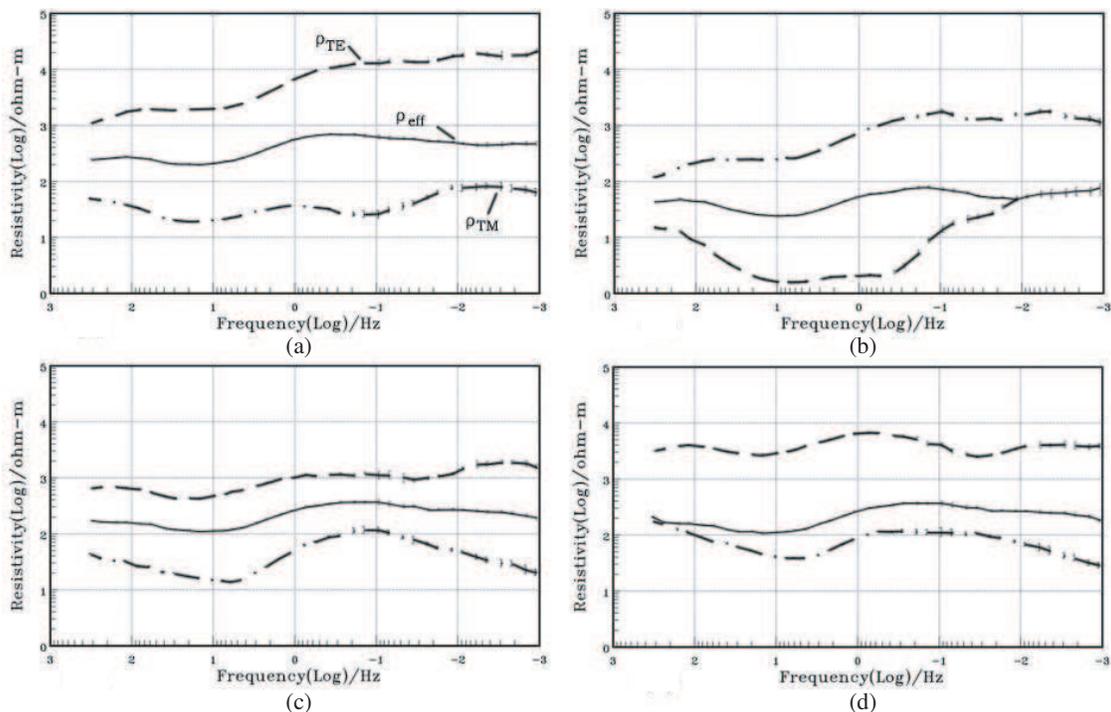


Figure 1: Apparent resistivity curves'changing features of 4 different continuous stations. (a) No. 206, (b) No. 207, (c) No. 208, (d) No. 209.

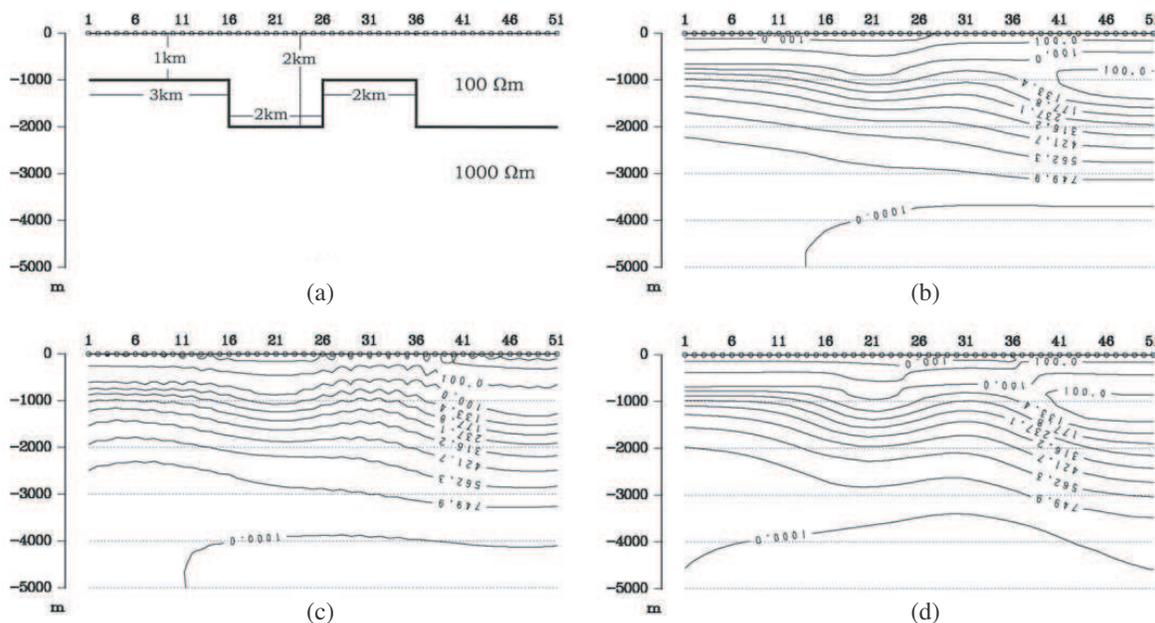


Figure 2: Inversion results of 2D geoelectric model and different apparent resistivity. (a) 2D geoelectric section on the survey line, (b) ρ_{xy} , ρ_{yx} joint inversion result, (c) ρ_{TE} , ρ_{TM} joint inversion result, (d) ρ_{eff} inversion result.

4. CONCLUSION

Effective apparent resistivity is the modulus of response magnetotelluric impedance tensor matrix divided by the $5f$, which is a rotation invariable and equals to conventional apparent resistivity ρ_a in 1D condition. In 2D condition, the effective apparent resistivity equals to the geometric mean of ρ_{TE} and ρ_{TM} . Since the actual geological structure is rather complex, and TE and TM mode of measured data is sometimes very difficult to determine with errors often occur. Even if judged correctly, the electrical principal axis is often a curve changing with frequency. All these factors will affect the 2D inversion results of sub-mode (TE mode, TM mode or joint mode). Inversion directly to the effective apparent resistivity, does not involve the issue of principal axis, so the error of principal axis determine does not affect inversion results.

ACKNOWLEDGMENT

This paper is supported by National Natural Science Foundation Project (60672042).

REFERENCES

1. Chen, L.-S., R. Liu, and T.-S. Wang, *The Magnetotelluric Sounding Data Processing and Interpretation Methods*, Petroleum Industry Press, Beijing, 1989.
2. Sutarno, D. and K. Vozoff, "Robust M-estimation of magnetotelluric impedance tensors," *Expl. Geophys.*, Vol. 20, 383–398, 1989.
3. Bostick, F. X., "Electromagnetic array profiling (EMAP)," *56th Annual SEG Meeting*, 42–100, Houston, 1986.
4. Groom, R. W. and R. C. Bailly, "Analytical investigations of the effects of near-surface three-dimensional galvanic scatters on MT tensor decompositions," *Geophysics*, Vol. 56, 496–518, 1991.
5. DeGroot-Hedlin, C. and S. Constable, "Occam's inversion to generate smooth, two-dimensional models from magnetotelluric data," *Geophysics*, Vol. 55, 1613–1624, 1990.
6. Smith, J. and J. Booker, "Rapid inversion of two- and three-dimensional magnetotelluric data," *Journal of Geophysical Research*, Vol. 96, 3905–3922, 1991.
7. Rodi, W. and R. Mackie, "Nonlinear conjugate gradients algorithm for 2-D magnetotelluric inversion," *Geophysics*, Vol. 66, 174–187, 2001.

Research and Application on Supergain Property of Arrays for Target Detection

Zhanlin Xie and Yingmin Wang

Institute of Acoustic Engineering, Northwestern Polytechnical University, Xi'an, China

Abstract— The method of supergain for a column array and its integrate robust beamforming are presented. The upright super-directive array gain of the column array outperforms a lot over that of a conventional beamforming in isotropic noise fields when the inter-element spacings are less than one-half wavelength. It is the phenomenon of supergain at the top direction to line array. At the horizontal direction, the fact of supergain also exists in the circle array of its cross section. However, in practice, the performance of optimum beamforming algorithms are known to degrade if some of underlying assumptions on the sensor array violated. It is impossible to get the theoretic supergain through optimum disposal. Therefore, the integrate supergain of robust beamforming algorithm (ISBA) is used to improve the robustness of supergain beamforming against random errors. Under the inhibit of beamforming weight vector norm, the column array shows great performance at target search and resist disturbance in two dimensions of orientation and depth.

1. INTRODUCTION

There are tow key factors to deal with in sonar. The first is the orientation of the target, and the second is to improve the space processing gain, which can increase the ratio of signal to noise in the input unit of receiver (the output unit of the array system). Now most arrays usually adopt conventional array of equable weight. The processing of signals are simple only by delay, weight and summation. To the conventional array, it means increases the sum of array elements to improve the array gain. Correspondingly, the aperture of the array is raised, that will lead to more difficulty for practical application. Under the equable isotropic noise field, when d/λ tends towards 0, however, the optimum array gain usually doesn't tend toward to 1, it will represent some special characteristics which are called supergain [1]. Only a small scale array can get higher array gain than usual big one. So, it is important to research the supergain of small scale array.

As early as 1983, Zhao Ning Zhen, Da Wei Xiang already depicted the conception of supergain in their papers [2]. Yet, within the immature of digital signal processing and electronic technique, underdeveloped manufacture technics, the supergain technique is despised by most scholars. Near the years, Yuan Liang Ma and Bao Gao Zhang gave concrete analysis to the robust of supergain beamforming, primarily demonstrated the supergian processing can be realized with the development of modern electronic technique and the level of manufacture technics at the time [3, 4]. She Feng Yan made deep and systemic research in supergain characteristics and signal processing method of hydrophone arrays [5]. By simulation and validation of experiment data processing, it made a strong impact on engineering application of this technique. After several decades dreariness, it is possible to get a new life for supergain processing technique. The purpose of this paper is brought under the background.

Along with the decreasing noise radiation of underwater vessels, it appears a big challenge to the passive sonar. In this paper, we study a supergain hydrophone column array to improve the detection performance of passive sonar. Comparison of performance is made between the conventional and supergain array. Beamforming in temporal domain and frequency domain is implemented for wideband signal of the array, and corresponding signal proecessing method is investigated.

2. THEORETICAL ANALYSE

2.1. Signal Field

Firstly, it is supposed that a plane wave signal of single frequency is received by the arrays, after the characteristics of signal is obtained, the same can be deduced to the broadband signals. To an array composed by M sensors, the hypothesis is that the sound press received by No. i sensor is:

$$S_i = p_0 e^{j(\omega t - \varphi_i)} \quad (1)$$

where $e^{j\omega t}$ denotes that sound press signal is sine wave signal which is changed with time, we call that quiver item. Since each signal received by the sensors includes the item, if we ignore $e^{j\omega t}$ in

formula (1), then it can be rewritten to:

$$S_i = p_0 e^{-j\varphi_i} \quad (2)$$

where $\varphi_i = \omega\tau_i$ is used to represent the delayed phase of No. i sensor unit to a reference wave surface, the function $\tau_i = \frac{1}{c}(r_i, e)$ is the corresponding delayed time, the item p_0 represents the correlated sound press amplitude.

After series of mathematics operation and transform, signal field R_s becomes:

$$R_s = [S_i S_j^*]_{N \times N} = S S^H \quad (3)$$

where $[\cdot]^H$ is the Hermitian transpose.

If we transform the correlation function of covariance matrix each element into normalization function, then can obtain signal normalization covariance matrix which every element is normalization correlation function of signal. The formula is :

$$p = V V^H = [V_i V_j^*]_{N \times N} \quad (4)$$

where function $V = [V_1, V_2, \dots, V_M]^T$.

Usually, we only pay attention to the formula when we analyse signal of plane wave:

$$p = V V^H \quad (5)$$

Function V is called arrange matrix of normalization signal, and $V V^H$ is called combined vector, which is the mathematics model of signal field of single frequency plane wave.

2.2. Noise Field

Noise is defined as the unexpected signal consisted in the receiver port which prompts the sensors. Essentially, it is some kind of signal with certain wave (usually called disturbance), or some ruleless random noise. Generally, it is used to assume that noise is a random course. So, generally, noise model is the concept of statistic meaning.

We always consider that the sea spontaneous noise is isotropic. In isotropic and uniform characteristic noise field, it is treated the noise received by sensors as the plus of different plane wave from all directions of space. These noise uniformly distribute and irrelate. They are smooth and random noise. It will satisfying describe the noise model if the second-order moments is calculated by statistic and average method.

The covariance matrix of noise is described as:

$$R_n = E\{n n^H\} = E \left\{ \sum_l^L \sum_k^K n_l n_k^H \right\} \quad (6)$$

Suppose that noise source is irrelated, then $n_l n_k = 0$, when $l \neq k$, the function is rewritten as:

$$R_n = E \left\{ \sum_{l=1}^L n_l n_l^H \right\} = \left[E \left\{ \sum_{l=1}^L n_l n_l^H \right\} \right]_{ij} \quad (7)$$

$i, j = 1, 2, \dots, N$ (N is the sum of sensors).

At last, the noise field formula can be deduced:

$$R_n = [R_{ij}]_{N \times N} = N \left[\frac{\sin k r_{ij}}{k r_{ij}} \right]_{N \times N} = N \left[\frac{\sin \frac{2\pi}{\lambda} r_{ij}}{\frac{2\pi}{\lambda} r_{ij}} \right]_{N \times N} \quad (8)$$

where, $i, j = 1, 2, \dots, N$, function k is the wavenumber, λ is the wavelength.

It can be concluded that when $\frac{r_{ij}}{\lambda}$ is integer double of $1/2$, $R_{ij} = 0$. Then it is considered that the noise received by sensors is irrelated, when the space between sensors is half wavelength.

2.3. Array Supergain

The data vector received by the sensors can be described as:

$$X(t) = S(t) + N(t) \tag{9}$$

That means the data vector received by the sensors are the sum of signal vector S and noise vector N . It is usually considered that the signal and noise received by the sensors is unrelated. Then,

$$E\{nS^H\} = E\{Sn^H\} = 0 \tag{10}$$

And

$$R_x = E\{SS^H\} + E\{nn^H\} = R_s + R_n \tag{11}$$

When we receive data vector by array, often give weights to the sensors. And the weight to the whole array is expressed by vector W [6, 7]. The formula is:

$$W = [w_1, w_2, \dots, w_n]^T \tag{12}$$

The whole output system's function is equivalent to that the received data vectors pass through a linear system with vector filter W . Its signification can be depicted by Figure 1:

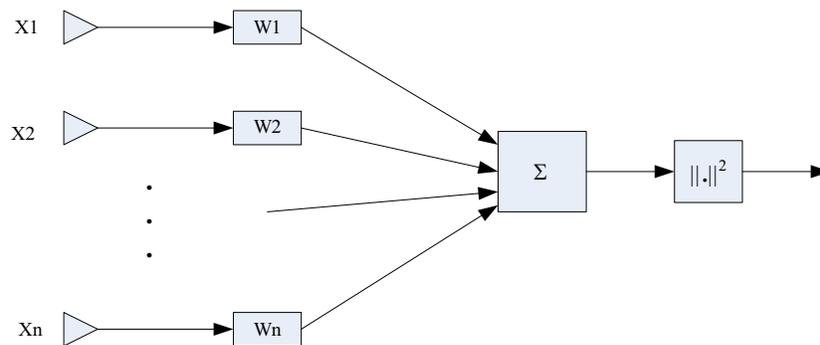


Figure 1: Sketch of array signal processing.

From Figure 1, it can be concluded that the final output power of array system is the sum of part of signal power and noise power. It is changed with the altering of weight vector and characteristics of signal and noise. If we expect to obtain the max space processing gain, which means to find a weight vector W :

$$G = \frac{|W^T V|^2}{W^T Q W^*} \rightarrow \max$$

To derivation it by W^* , which means $dG/dW^* = 0$, then can obtain $Q^{-1} p W^* = G W^*$. If it can content the characteristic equation:

$$\begin{cases} W^T = V^H Q^{-1} \\ G = V^H Q^{-1} V \end{cases} \tag{13}$$

That will means the system has the maximum output of S/N . Now it will have the best array gain G . The G is the supergain of the array, and the weight is the the supergain weight. Beamforming with the weight is supergain beamforming. When the interval of each sensor is half of the wavelength, the noise coefficient of each sensor received is 0. Then $Q = I$ which is the unit matrix. And the weight degenerates to the conventional weight [8]. To beamforming with it is the conventional beamforming. The maximum array gain of it is $10 \lg N$ (N is the number of sensors).

3. SIMULATION EXPERIMENT

Considering the frequency processed by the modern sonar is lower than ever, commonly from several decades Hz to several decades kHz. And the interval of each sensor is demanded far shorter than half wavelength by supergain. What is more, the column array has the characteristic of symmetry. We design a column array, its diameter is 2.0 meter, 32 sensors is placed in it, the interval of each sensor is $\lambda/4$ (λ is the wavelength). The planform of the array is shown in Figure 2:

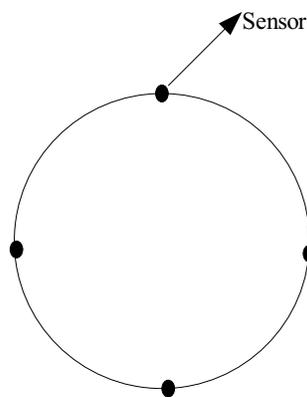


Figure 2: The planform of 32 sensors array.

To this array, we give these hypothesis:

- (1) Each sensor is uniform placed in the circle with four layers, the position excursion is inexistent.
- (2) Signal received by the sensors is isotropic, and all the sensors have the same sensitivity.
- (3) Whole sensors have the preferable frequency coherence during the design frequency.
- (4) The array is located in the uniform noise field of space.
- (5) The array frame is transparent to sensors.

Using this algorithm, we obtain the supergain weight of this column array, calculate out supergain of different frequency and the conventional gain of the same frequency. The gains of every frequency is compared and drawn in Figure 3.

From Figure 3, obviously, we can find that supergain processing can obtain higher space gain than conventional one, especially in the low frequency segment. During 200–800 Hz frequency segment to this array, it can get higher space gain of 5.5 dB than conventional. From this fact, it shows that supergain processing has advanced merit in little radius and wavelength of column array.

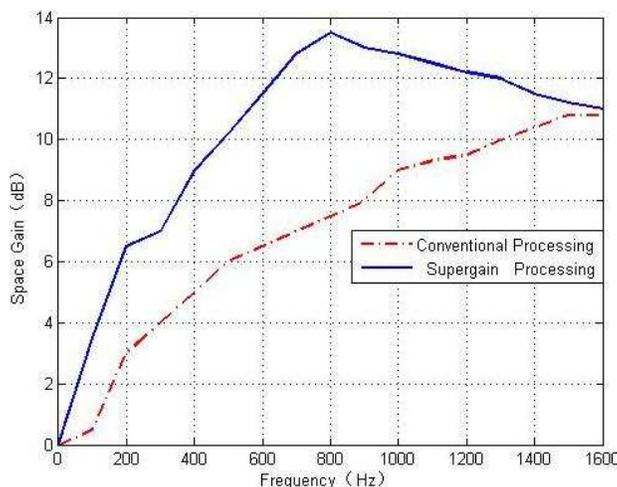


Figure 3: Comparison of conventional and supergain processing.

4. CONCLUSION

To sonar signal processing, if the sonar's frequency is low, the column of array will be huge to obtain certain space gain with conventional processing. It is not convenient to practice and engineering implement. However, by supergain processing, a relative small scale array can obtain enough space gain. It will be very useful to improve the characteristic of passive sonar of low frequency in target diction. The technique will have wider application view.

REFERENCES

1. Li, G. B., *Design Principle of Sonar Array*, Ocean Press, Beijing, 1993.
2. Zhen, Z. N. and D. W. Xiang, *Theory of Underwater Acoustic Signal Passive Detection and Parameter Estimation*, Science Press, Beijing, 1983.
3. Ma, Y. L. and B. S. Zhang, "Investigation to the supergain hydrophone arrays for wideband application in isotropic ambient noise field," *Proceedings of ECUA 2000*, 1043–1048, Lyon, France, July 10–13, 2000.
4. Zhang, B. G., "Optimal design of broadband hydroacoustic receiving array," Eng. Ph.D. Dissertation, Northwestern Polytechnical University, 1999.
5. Yan, S. F., "Study on supergain property and signal processing of hydrophone arrays," MEng Dissertation, Northwestern Polytechnical University, 2001.
6. Knight, W. C., R. G. Pridham, and S. M. Kay, "Digital signal processing for sonar," *Proceedings of the IEEE*, Vol. 69, No. 11, 1451–1506, November 1981.
7. Applebaum, S. P., "Adaptive arrays," *IEEE Trans. Antennas Propag.*, Vol. 24, 585–598, 1988.
8. Van Veen, B. D. and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE ASSP Magazine*, 4–24, April 1998.

Discrete Time Synergetic Control for DC-DC Converter

Qian Wang, Tao Li, and Jiuchao Feng

School of Electronic and Information, South China University of Technology
Guangzhou 510641, China

Abstract— A discrete-time version of synergetic control is introduced and applied to design a nonlinear synergetic controller for DC-DC converters such as buck converters. The performance of the discrete-time synergetic controller is compared with that of a conventional Proportional-Integral-Differential (PID) controller through the input voltage step test and the step load test on Simulink. In particular, two modified control laws are provided to improve performances of controller. Further, the two control laws are evaluated in comparison with the first simple discrete-time synergetic controller.

1. INTRODUCTION

Digital technology has great impacted the design of controls for power electronics with the advent of digital signal processors (DSP) and high-performance processors. The synergetic approach to control theory (SACT) is a novel nonlinear control method that the nonlinearities of system are considered in the control design and a systematic design procedure which yields control laws suitable for digital implementation are offered [1–3]. Moreover, synergetic control not only gives constant switching frequency operation, but also provides asymptotic stability with respect to the required operating modes, invariance to load variations and robustness to variation of converter parameters [4, 5]. Although the synergetic control is widely utilized in various fields [2–11], these synergetic controllers are all continuous-time.

In this paper, an alternative discrete-time extension of the SACT design procedure is at first provided, which can not be obtained from its continuous counterpart by means of simple equivalence. The discrete-time extension of the SACT procedure is used to design and implement a SACT controller for buck converters. The performance of the discrete-time SACT controller for this application is compared with that of a conventional Proportional-Integral-Differential (PID) controller through the input voltage step test and the step load test on Simulink. In particular, two modified control laws are introduced to avoid transient overshoot, reduce steady-state error and obtain fast transient response. Further, the two controls are evaluated by comparison with the first simple discrete-time synergetic controller.

2. BACKGROUD THEORY OF DISCRETE TIME SYNERGETIC CONTROL

A discrete version of the SACT procedure is important when the implementation of the control is realized by computers or other digital processors. Additionally, some control problems can be better approached if the plant is discretized. Similarity to the continuous-time SACT procedure [1–4], the discrete-time SACT design procedure consists of the following steps:

- 1) Start by defining a macro-variable vector as a function of the state variables. Consider a simple discrete-time system as follows

$$\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k) \quad (1)$$

where $\mathbf{x}(k)$ is the state vector of the system in \mathfrak{R}^n , \mathbf{u} is a control vector in \mathfrak{R}^m ($m \leq n$), $\mathbf{A}_{n \times n}$ and $\mathbf{B}_{n \times m}$ are system matrices. Then a macro-variable vector is defined as follows:

$$\psi(k) = \psi(\mathbf{x}, k) \quad (2)$$

where ψ is a macro-variable vector in \mathfrak{R}^l ($0 < l \leq m$). By controlling, the system will be forced operating on these invariant manifolds

$$\psi = 0 \quad (3)$$

and run to the specified equilibrium point along these manifolds. The designer can select the characteristics of the macro-variable vector according to the control objectives (e.g., limitation in the current, and so on). In general case, the macro-variable vector is a simple linear combination of the state variables. The dimension of the macro-variable vector does not exceed the number of control channels.

- 2) In continuous-time SACT procedure, $\mathbf{T} \cdot \dot{\psi} + \psi = 0$ defines the speed and trajectory of convergence to the invariant manifold [4]. Considering sampling period T_s , the discrete counterpart is derived as follows:

$$\mathbf{T} \cdot \left[\frac{\Psi(k+1) - \Psi(k)}{T_s} \right] + \psi(k) = 0, \quad T_i > 0, \quad i = 1 \dots l \quad (4)$$

where \mathbf{T} is a design parameter vector that specifies the convergence speed to the manifolds (3). (4) can be rewritten in the following form:

$$\left[\frac{\mathbf{T}}{T_s - \mathbf{T}} \right] \psi(k+1) + \psi(k) = 0 \quad T_i > 0, \quad i = 1 \dots l \quad (5)$$

- 3) The control vector \mathbf{u} is calculated according to (5) and (1), which gives a control ensuring the specified properties and decomposing the system by contraction of its state space. The order of the decomposed system on the manifolds is $n - l * m$.

3. THE BUCK CONVERTER CASE: AN EXAMPLE OF CONTROL SYNTHESIS

3.1. Controlling Synthesis

We now synthesize a discrete-time controller for a DC-DC Buck Converter as shown in Fig. 1. First, the state space averaged model is used to the buck converter that operates in a very high frequency continuous conduction mode. The output voltage transfer function of the buck converter is described by

$$G_P(s) = \frac{V_o}{u} = \frac{V_{DC}R}{RLCs^2 + Ls + R} \quad (6)$$

where u is the duty cycle ($0 \leq u \leq 1$). The synthesis object is to obtain a control law u , which forces the system to operate at the required values of converter output voltage $V_o = V_{oref}$ and current $i_L = i_{Lref}$.

The z -domain transfer function $G_P(z)$ from (6) with a sampling period T_s is

$$G_P(z) = \frac{V_o(z)}{u(z)} = \frac{b_0z + b_1}{z^2 + a_1z + a_2} \quad (7)$$

where b_0, b_1, a_1 and a_2 are coefficients. Then the difference equation of (7) is

$$V_o(k+1) = -a_1V_o(k) - a_2V_o(k-1) + b_0u(k) + b_1u(k-1) \quad (8)$$

Second, considering that there is one control channel u in the system, a macro-variable is defined as

$$\psi(k) = \kappa \cdot e(k) + e(k-1) \quad (9)$$

$$e(k) = V_o^*(k) - V_o(k) \quad (10)$$

where κ is the SACT controller parameter, $e(k)$ is the output voltage error, $V_o(k)$ is the output voltage and $V_o^*(k)$ is the voltage reference.

Substituting (9) and (10) into (5), yields

$$V_o^*(k+1) - V_o(k+1) = \left[1 - \frac{T_s}{T} - \frac{1}{\kappa} \right] [V_o^*(k) - V_o(k)] + \frac{1}{\kappa} \left[1 - \frac{T_s}{T} \right] [V_o^*(k-1) - V_o(k-1)] \quad (11)$$

To consider $V_o^*(k+1) = V_o^*(k)$, substituting (8) into (11) obtains the control law as follows:

$$u(k) = \frac{1}{b_0} \left[V_o^*(k) - b_1u(k-1) + a_1V_o(k) + a_2V_o(k-1) - \left(1 - \frac{T_s}{T} - \frac{1}{\kappa} \right) e(k) - \frac{1}{\kappa} \left(1 - \frac{T_s}{T} \right) e(k-1) \right] \quad (12)$$

The expression for u is the control action for the converter controller, which forces the plant (7) from an arbitrary initial condition to the invariant manifolds (3).

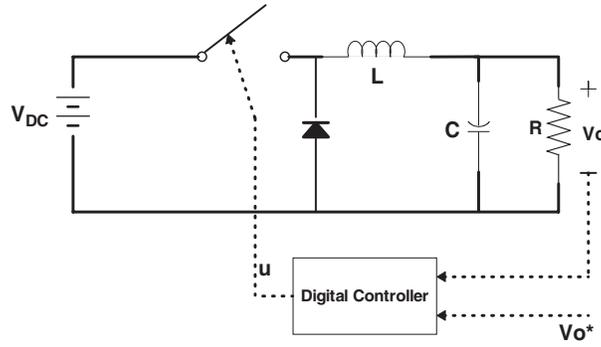


Figure 1: Buck converter control system.

3.2. The Stability of the Closed Loop System

Another form of (11) using the output voltage error e is as follows

$$e(k+1) - \left(1 - \frac{T_s}{T} - \frac{1}{\kappa}\right) e(k) - \frac{1}{\kappa} \left(1 - \frac{T_s}{T}\right) e(k-1) = 0$$

Therefore the characteristic equation of the close-loop system is

$$z^2 - \left(1 - \frac{T_s}{T} - \frac{1}{\kappa}\right) z - \frac{1}{\kappa} \left(1 - \frac{T_s}{T}\right) = 0 \quad (13)$$

For the close-loop system to be stable, it is necessary that (14) should be satisfied.

$$|z_p| < 1 \quad (14)$$

4. SIMULATION RESULTS

The simulations were performed in Matlab/Simulink environment. System parameters are shown in Table 1. Choose $\kappa = 20$, $T = 8 * 10^{-5}$, $T_s = 50 \mu\text{s}$ according to (14). Generally it is interesting to evaluate the control performance under an input voltage variation and a step-load variation. Therefore the control performance was tested under the following two situations.

- ◆ R changed from 10.4Ω to 5.2Ω at $t = 6 \text{ ms}$, and then stepped to 10.4Ω at $t = 12 \text{ ms}$.
- ◆ V_{DC} changed from 10 V to 6.5 V at $t = 8 \text{ ms}$, and then stepped from 6.5 V to 10 V at $t = 15 \text{ ms}$.

One of the traditional ways to design a PID controller for the above buck converter is to use the zero-pole deploying method. According to the method, the PID controller parameters are fine tuned as $K_p = 0.1001$, $K_i = 83.106$, $K_d = 1.1853e - 4$. The simulation results under the above two situations for both PID and SACT controls are shown in Fig. 2 and Fig. 3. From Figs. 2 and 3, the transient time is about 2 ms for synergetic control but about 5 ms for PID control. To be noticed, when the step load is from 10.4Ω to 5.2Ω , the synergetic control operates satisfactorily, but PID control goes unstable. The similar result happens when input voltage steps down. Therefore, synergetic control has faster response during transients than PID control. Also, synergetic control exhibits good performance under the large-signal transients than a conventional PID controller.

Table 1: Buck converter parameters.

Parameter	Symbol	Value
Input voltage	V_{DC}	10 V
Output filter inductance	L	1.9 mH
Output filter capacitance	C	100 μF
Nominal load	R	10.4 Ω

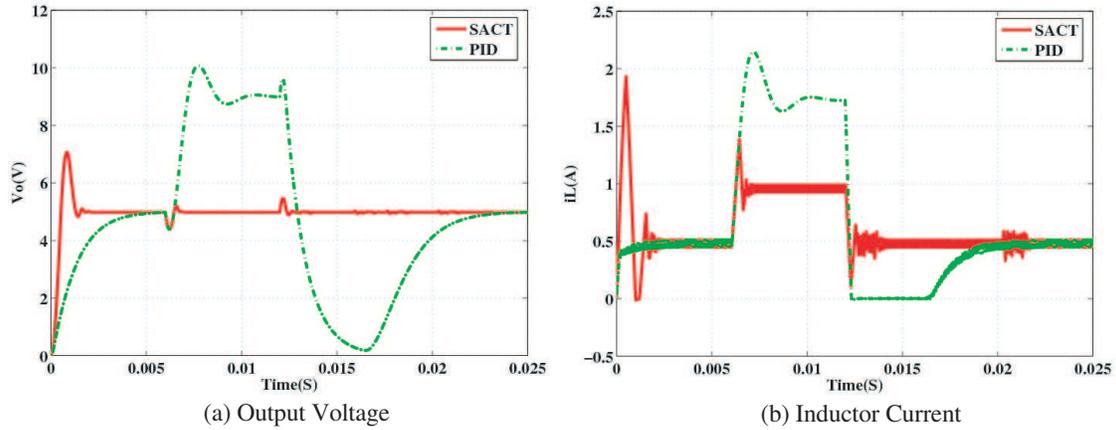


Figure 2: Inductor current and output voltage response of Buck converter under load variations.

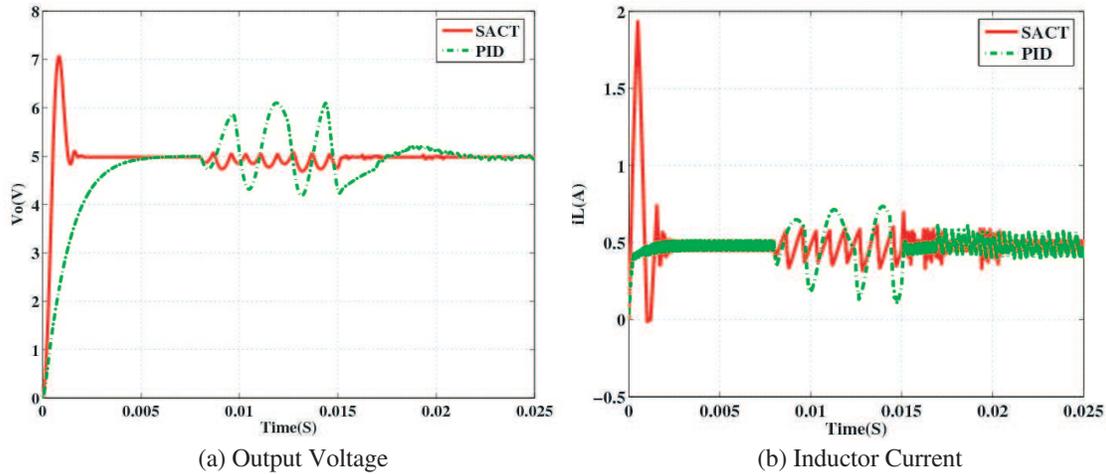


Figure 3: Inductor current and output voltage response of Buck converter under input voltage variations.

5. SYNTHESIS OF IMPROVED CONTROLS

The previous case has illustrated a simple discrete synergetic controller synthesis, which simulation results have been compared with PID controller. However, the control law (12) has not feedback of the inductor current so that current overshoot can not be limited. Therefore more complex macro-variable definitions will next be introduced to implement improved control laws. The procedure to derive the new control laws is analogous to the one described above. In the following part of the paper two different macro-variable definitions will be introduced and their performance evaluated in simulation.

5.1. Current Limit Implementation

The macro-variable is defined as follows

$$\begin{aligned}\psi(k) &= \kappa \cdot e_i(k) + e_v(k) \\ e_i(k) &= i_L^*(k) - i_L(k) \\ e_v(k) &= V_o^*(k) - V_o(k)\end{aligned}\quad (15)$$

where $i_L(k)$ is the inductor current and $i_L^*(k)$ is the current reference.

Repeating the usual synthesis procedure, the control law can be obtained as

$$\begin{aligned}u(k) &= \frac{1}{b_{v0} + \kappa b_{i0}} \{V_o^*(k) + a_{v1}V_o(k) + a_{v2}V_o(k-1) - (b_{v1} + \kappa b_{i1})u(k-1) \\ &\quad + \kappa [i_L^*(k) + a_{i1}i_L(k) + a_{i2}i_L(k-1)] - \left(1 - \frac{T_s}{T}\right) [e_v(k) + \kappa e_i(k)]\}\end{aligned}\quad (16)$$

where b_{v0} , b_{v1} , a_{v1} and a_{v2} are the coefficients of z -domain transfer function of $V_o(k)$, b_{i0} , b_{i1} , a_{i1} and a_{i2} are the coefficients of z -domain transfer function of $i_L(k)$.

Using the same simulation parameters as the Section 4 except for $\kappa = 5$, the simulation results under macro-variable (9) and (15) are presented in Fig. 4, in which the data1 (red) represents macro-variable (9), the data2 (green) represents macro-variable (15). Fig. 4(b) shows the peak inductor current of the synergetic control under macro-variable (9) is about 2A but 1A for the one under macro-variable (15). Furthermore, no voltage overshoot is under synergetic control of macro-variable (15) but the peak output voltage of the synergetic control under macro-variable (9) is about 7 V as shown in Fig. 4(a). However, the steady-state error under synergetic control of macro-variable (15) is a little bigger than that of the synergetic control under macro-variable (9).

5.2. Control Law with Dynamic Parameter κ

The parameter κ involved in (15) plays a significant role in control strength. A geometric interpretation of the control law in the phase plane is shown in Ref. [3]. The steady state operation point is the origin, where the error goes to zero with slope $-1/\kappa$. The parameter κ determines the strength that the errors in the output voltage are canceled by using the error on the current. Decreasing κ makes the control faster and stronger, forces the system to the origin more quickly and reduces steady output voltage errors more, but results in a higher current overshoot during the transient and makes the closed-loop system more possibly unstable. Therefore the choice of the value of control parameter κ involves a tradeoff: during transient a relatively large value of κ is desirable, because it avoids large over-currents and excessive stress on the switches. On the other hand, a small value is desirable in steady state, because this reduces the steady-state output voltage error. Based on this consideration, κ in (15) can be chosen as follows:

$$\kappa = \alpha + \beta \bullet |e_v(k)| \quad (17)$$

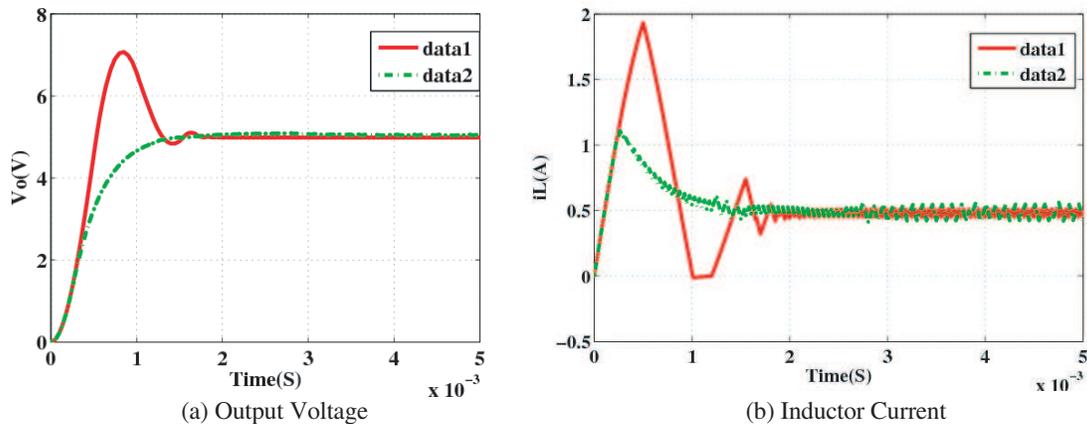


Figure 4: Comparison of simulation response under the definition of macro-variable (9) and (15).

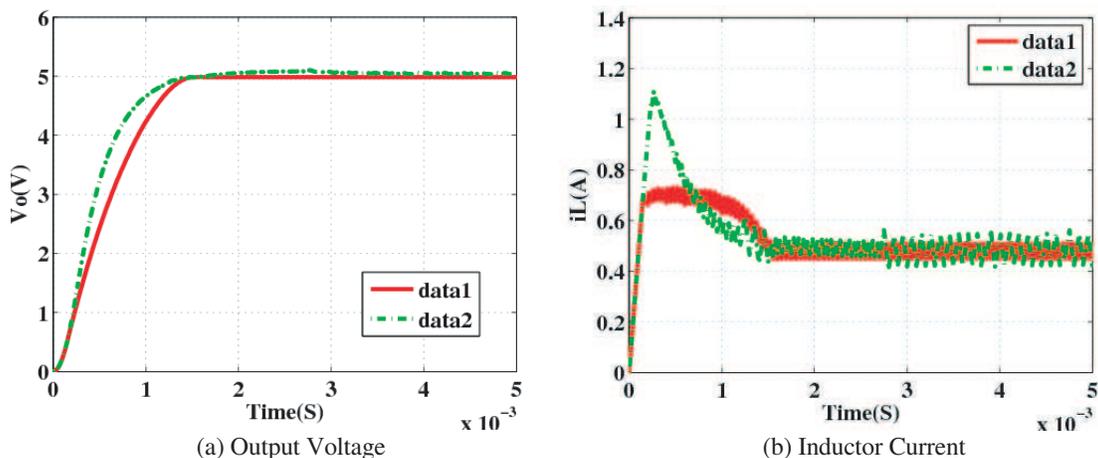


Figure 5: Comparison of simulation response under the different definition of κ .

The simulation parameters are the same as the Section 4. The simulation responses under macro-variable (15) of different definition of κ are shown in Fig. 5, in which the data1 (red) represents κ given by (17) for $\alpha = 0.05$ and $\beta = 4$, the data2 (green) represents κ is 20. From Fig. 5, under κ given by (17), not only the steady-state error but also the overshoot of current is reduced.

6. CONCLUSION

The application of discrete-time synergetic control to a buck converter has been studied. The advantages of synergetic control over PID are presented through the input voltage step test and the step load test on Simulink. Moreover, two different control options demonstrate the flexibility of the synergetic control method. Simulation results for control laws show that their dynamic performances of the synergetic control system can be improved by suitable choices for control parameters and macro-variables.

ACKNOWLEDGMENT

This work was partially supported by the National Natural Science Foundation of China under Grant 60872123, and by the Joint Fund of the National Natural Science Foundation and the Guangdong Provincial Natural Science Foundation, China, under Grant U0835001.

REFERENCES

1. Kolesnikov, A., G. Veselov, A. Kolesnikov, et al., *Modern Applied Control Theory: Synergetic Approach in Control Theory*, TSURE Press, Moscow-Taganrog, 2000.
2. Kolesnikov, A., G. Veselov, A. Kolesnikov, et al., "Synergetic synthesis of DC-DC boost converter controllers: Theory and experimental analysis," *Proc. of IEEE Applied Power Electronics Conference*, 409-415, Dallas, Texas, 2002.
3. Santi, E., A. Monti, D. H. Li, et al., "Synergetic control for DC-DC boost converter: Implementation options," *IEEE Trans. on Industry Applications*, Vol. 39, No. 6, 1803–1813, 2003.
4. Wang, Q., J. Feng, and T. Li, "Analysis of the synergetic control based on variable structure and application of power electronics," *International Conference on Information Engineering and Computer Science*, Wuhan, China, 2009.
5. Santi, E., A. Monti, D. H. Li, K. Proddatur, and R. Dougal, "Synergetic control for power electronics applications: A comparison with the sliding mode approach," *J. of Circuits, Systems, and Computers*, Vol. 13, No. 4, 737–760, 2004.
6. Davoudi, A., A. M. Bazzi, and P. L. Chapman, "Application of synergetic control theory to non-sinusoidal PMSMs via multiple reference frame theory," *Proc. of IEEE the 34th Annual Conference on Industrial Electronics*, 2794–2799, 2008.
7. Kondratiev, I. and R. Dougal, "Current distribution control design for paralleled DC/DC converters using synergetic control theory," *Proc. of IEEE Power Electronics Specialists Conference*, 851–857, 2007.
8. Kondratiev, I., E. Santi, and R. Dougal, "Robust nonlinear synergetic control for m-parallel-connected DC-DC boost converters," *Proc. of IEEE Power Electronics Specialists Conference*, 2222–2228, 2008.
9. Yuan, X. F., J. Y. Wen, Z. C. Zhou, and S. J. Cheng, "A synergetic control scheme for HVDC transmission system," *Proc. of 2006 International Conference on Power System Technology*, 1–5, 2006.
10. Jiang, Z., "Design of power system stabilizers using synergetic control theory," *Proc. of IEEE Power Engineering Society General Meeting*, 1–8, 2007.
11. Yu, X. W., Z. H. Jiang, and H. Zhang, "A synergetic control approach to grid-connected, wind-turbine doubly-fed induction generators," *Proc. of IEEE Power Electronics Specialists Conference*, 2070–2076, 2008.

Novel Optical Neuronal Cell and Data Recognition-generation Circuits in RFID Tags

Norimitsu Wakama and Yukio Iida

FSC, and Dept. of Electric and Electron. Eng., Faculty of Engineering Science
Kansai University, Japan

Abstract— The optical neuronal cell composed of silicon structure is also studied, and novel cell is proposed. The circuits that recognizes tag call signal and generates the sending data using the analog threshold element with a nonlinear characteristic are studied. In addition, single-mode silicon optical switch with T-shaped SiO₂ optical waveguide as a control gate is proposed [5]. In this report, the neuronal cell composed of silicon waveguide and electronic active elements is studied, and novel neuronal cell is proposed. The electron-hole pairs generated by illumination with short-wavelength light are used here.

1. INTRODUCTION

The research to model animals of the higher orders' nervous systems of the sight by composing the analog threshold element with a nonlinear characteristic in multilayer has been conducted [1, 2]. Recently, there are many reports from two dimension pattern recognition view point though an active research is performed.

On the other hand, RFID (radio frequency identification) technology has fast been noticed and has a great many possible applications [3]. We are proposing the RF recognition system to communicate by using faint radio wave with a little power consumption [4, 5].

In addition, single-mode silicon optical switch with T-shaped SiO₂ optical waveguide as a control gate is proposed [6].

In this paper, the neuronal cell composed of silicon waveguide and electronic active elements is studied, and novel neuronal cell is proposed. The electron-hole pairs generated by illumination with short-wavelength light are used here. In this paper, we also propose the circuit that recognizes the IC tag call signal and the sending data generation circuit by using the analog threshold element with a nonlinear characteristic. It is expected that these are effective to making it to small electric power or more.

2. NEURONAL CELL MODEL

Figure 1 shows the outline of the neuron cell that we are assuming. Signal light propagates from each SiO₂ optical waveguide in the left end of figure. The light having an $h\nu$ value larger than the band gap energy of silicon is absorbed to silicon, and converted into a free carrier. The free carrier enters self-pulsation electric circuit, and is converted into the pulse light with the light emission element. The converted light goes along in the waveguide and is sent to the following neuron cell.

3. SIMULATION RESULT

The finite-difference time-domain (FDTD) method is used for present electromagnetic field simulations. The computation model is shown in Fig. 2. The cell size for computations is $\Delta = 24.4$ nm. The typical design parameters are as follows: width and thickness of SiO₂ waveguide are about 880 nm, gap between SiO₂ and silicon is about 4.9 nm, and wavelength of signal light is 0.85 μ m. The signal light is excited at the left side of waveguide.

Figure 3 shows the simulation results. Left and right side figures are electromagnetic field distributions for electric conductivity $\sigma = 0$ and 20,000 S/m, respectively. From right figure, we can see that the light is absorbed. Fig. 4 shows the absorption characteristic for the electric conductivity. The unit of vertical axis is dB.

4. ESTIMATION OF CURRENT VALUE

When the silicon region is illuminated with the light having an $h\nu$ value larger than the band gap energy of silicon, the irradiated photons are absorbed to create electron-hole pairs. When the light power P_0 is supplied to the silicon, the number of photons per second is given by

$$\frac{P_0}{E_{\text{photon}}} = \frac{P_0}{h\nu} \quad [\text{photons/s}], \quad (1)$$

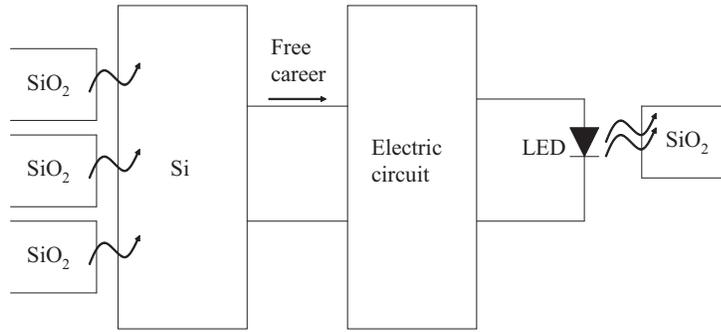


Figure 1: Outline of the neuron cell.

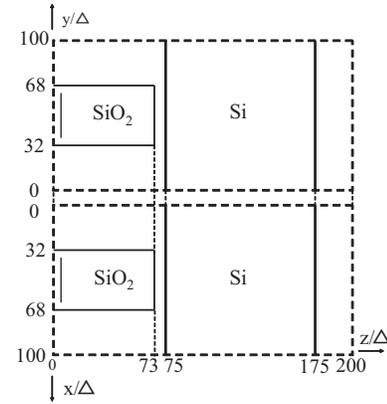


Figure 2: Computation model.

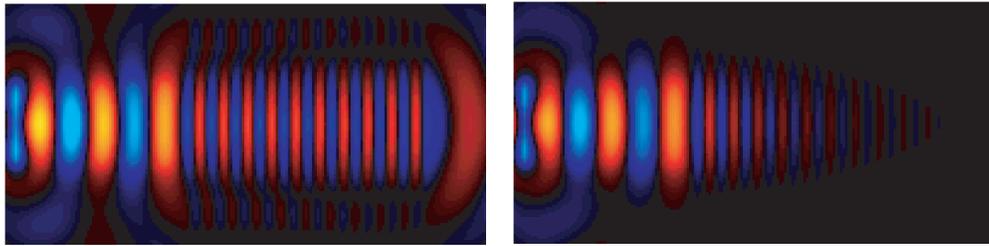


Figure 3: Electromagnetic field distributions for electric conductivity $\sigma = 0$ (left) and 20,000 S/m (right), respectively.

where h is Planck's constant and E_{photon} is the photon energy. We assume that the energy transferred by the SiO_2 waveguide is absorbed by a fraction of A in the silicon region. The number of photons contributing to the generation of electron-hole pairs (N_{photon}) is given by

$$N_{photon} = A \frac{P_0}{h\nu}. \tag{2}$$

We assume that the number of generated electron-hole pairs is N . In the steady state, the recombination rate per second (N/τ) is given by

$$\frac{N}{\tau} = A \frac{P_0}{h\nu} \tag{3}$$

The transferred energy to maintain the steady state can be estimated from Eq. (3).

When 50% of the light energy is consumed, it is estimated that the electron-hole pairs generated is about 2.14×10^8 for the light of $1 \mu\text{W}$; here we assume that the lifetime of minority carriers is $10 \mu\text{sec}$. When 10% is consumed to flow the external current and 40% is consumed to maintain the carrier concentration, we obtain the current $0.69 \mu\text{A}$ enough for the operation of an external circuit.

Table 1: Identification results.

X_1	-1	+1	-1	-1
X_2	-1	-1	+1	-1
A	+1	-1	-1	-1
B	-1	+1	-1	-1
C	-1	-1	+1	-1
D	-1	-1	-1	+1

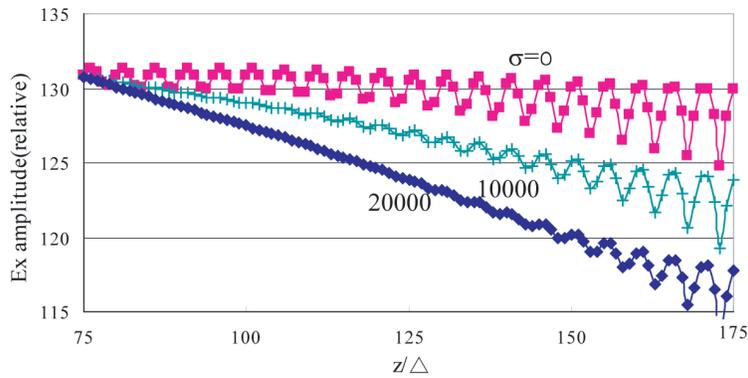


Figure 4: Absorption characteristics for various σ values.

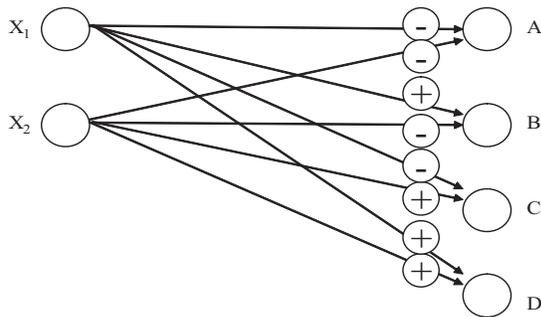


Figure 5: Call signal identification circuit.

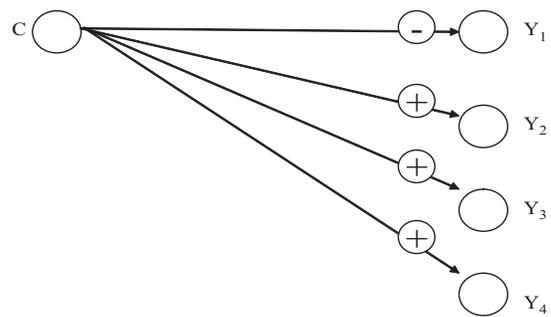


Figure 6: Sending data generation circuit.

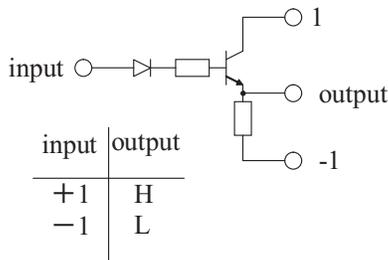


Figure 7: Computation circuit of route (+).

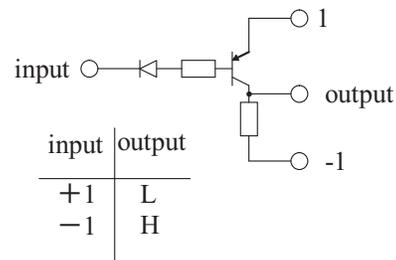


Figure 8: Computation circuit of route (-).

5. TAG CIRCUIT

Figure 5 shows the simple circuit that recognizes the IC tag call signal. Route + multiplies by $(1/2)$. Route - multiplies by $(-1/2)$. The sum of two routes is taken as the output signal. For instance, when -1 is input to x_1 , and 1 is input to x_2 , C becomes 1. The example of two bit input is shown for easiness. Table 1 shows the recognition results of the circuit. One of the outputs A, B, C, and D becomes the result of +1 corresponding to serial input (x_1, x_2) . If it wants to identify only condition $(x_1, x_2) = (-1, 1)$, only the circuit that leads to output C is necessary.

Figure 6 shows the sending data generation circuit. For example, if one input C of a parallel signal is 1, the serial output (Y_1, Y_2, Y_3, Y_4) becomes to $(1, -1, 1, 1)$. The operation of IC tag is relatively simple. We want to achieve it with an easy circuit as much as possible. It is thought that this was effective to making it to small electric power or more.

6. BASIC CIRCUIT ELEMENT

Figures 7 and 8 show the example of a basic element when route (+) and route (-) of Figs. 5 and 6 are composed of the transistor and the diode. The transistor works if + is input in the circuit of Fig. 7. H level is output as a result. If - is input oppositely, L level is output without the transistor working. In the circuit of Fig. 8, the transistor works when input is -, and H level is

output. L level is output in +.

7. CONCLUSION

The optical neuronal cell composed of silicon structure is also studied, and novel cell is proposed. A simple recognition circuit and the sending information data generation circuit were studied. A simple structural recognition circuit and the sending data generation circuit by using the analog threshold element with a nonlinear characteristic were studied, and were proposed. It is thought that this was effective to making it to small electric power or more.

ACKNOWLEDGMENT

This work was supported by “Academic Frontier” Project for Private Universities: Matching Fund Subsidy from MEXT, 2007-2011. The computations for this research were performed by the Large-Scale Computer System at the Osaka University Cyber media Center.

REFERENCES

1. Fukushima, K., “Feature extraction by multi-layered network of analog threshold elements,” *IECE Trans.*, Vol. 51-C, No. 7, 319–326, 1968.
2. Watanabe, O. and K. Fukushima, “A neural network model of near and far cells,” *IEICE Trans.*, Vol. J79-D-11, No. 9, 1404–1413, Sept. 1995.
3. Finkenzeller, G. K., *RFID Handbook, Radio-frequency Identification Fundamentals and Applications*, John Wiley & Son, Chichester, 1999.
4. Iida, Y., “Proposal of RF recognition system prepared for demand expansion and use of super-regenerative circuit,” *Proceedings of 2007 IEICE Society Conference*, B-5-88, Japan, Sep. 2007.
5. Iida, Y., “Novel super regenerative transponder (SRGT) for RFID tags and ASK signals,” *PIERS Online*, Vol. 5, No. 2, 187–190, 2009.
6. Kobayashi, H., Y. Iida, and Y. Omura, “Single-mode silicon optical switch with T-shape SiO₂ waveguide as n control gate,” *Jpn. J. Appl. Phys.*, Vol. 41, Part 1, No. 4B, 2563–2565, Apr. 2002.

3-D Analysis of Magnetic Flux Density in Modular Toroidal Coil Using Cubic Meshing

M. R. Alizadeh Pahlavani, A. Shiri, and A. Shoulaie

Department of Electrical Engineering, Iran University of Science and Technology (IUST), Tehran, Iran

Abstract— In this paper, equations for calculation of the magnetic flux density of the Modular Toroidal Coil (MTC) applicable to Tokamak reactors are presented. The MTC is composed of several solenoidal coils (SCs) connected in series and distributed in the toroidal and the symmetrical form. These equations are based on the Biot-Savart equation. The numerical analysis of the integrations resulting from these equations is solved using the extended three-point Gaussian algorithm. At the end, the Cubic Meshing approach is employed to present an algorithm to study the three dimensional leakage flux distribution pattern of the coil and to draw the magnetic flux density lines of the MTC. The presented algorithm, due to its simplicity in analysis and ease of implementation of the non-symmetrical and three dimensional objects, is advantageous to the commercial software such as ANSYS, MAXWELL, and FLUX.

1. INTRODUCTION

Recent research work in the area of plasma reactors (e.g., Tokamak), Superconductor Magnetic Energy Storage (SMES), and nuclear fusion reactors have studied different coils [1]. Tokamak reactors consist of coils with different structures, such as the modular toroidal, the helical toroidal, the solenoidal, and the poloidal coils as seen in Fig. 1 [2]. The capability of modular implementation of the MTC is one of the main advantages of this coil over the helical toroidal coil. Considering that the MTC has been studied less than other coils, in this paper the MTC is mathematically modeled and analyzed. Since the optimal design of the MTC with different objective functions such as the maximization of the stored magnetic energy, the minimization of the leakage flux, the stabilization of Tokamak reactors, and the elimination of stress requires an accurate calculation of the inductances of this coil, in this article we first derive the analytical equations of the inductances of this coil. The structure of this coil is shown in Fig. 2. As it is observed, this structure is composed of S solenoid coils (SCs) connected in series distributed in the toroidal and the symmetrical form. In this figure, S is assumed to be equal to 8. In Fig. 2(b), the projection of this coil on $x-y$ plane is shown. In this figure, R_{mk} is the distance between the symmetry center of the k th SC from z -axis. Also, the angle between R_{mk} direction and the latitudinal axis of the k th SC is shown by ν_{mk} . Moreover, the toroidal angle of the k th SC is defined by φ_{mk} , and the distance between the longitudinal axis of the k th SC and plane $z = 0$ is shown by H_{mk} . As mentioned above, in the MTC the values of R_{mi} and ν_{mi} , $i = 1, \dots, S$ are defined as $R_{mi} = \text{const}$, $\nu_{mi} = 0$. The dependency of analytical equations of inductance of the MTC on the geometrical parameters of the SCs such as R_{mi} , ν_{mi} , φ_{mi} , H_{mi} , $i = 1, \dots, S$, when the dimensional parameters of the SCs are known, shows that these parameters can be used as the degrees of freedom of the objective function and thus can be manipulated to satisfy the optimization function.

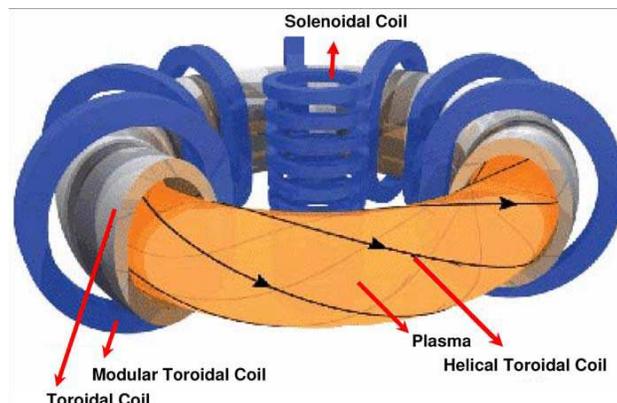


Figure 1: A typical tokamak reactor.

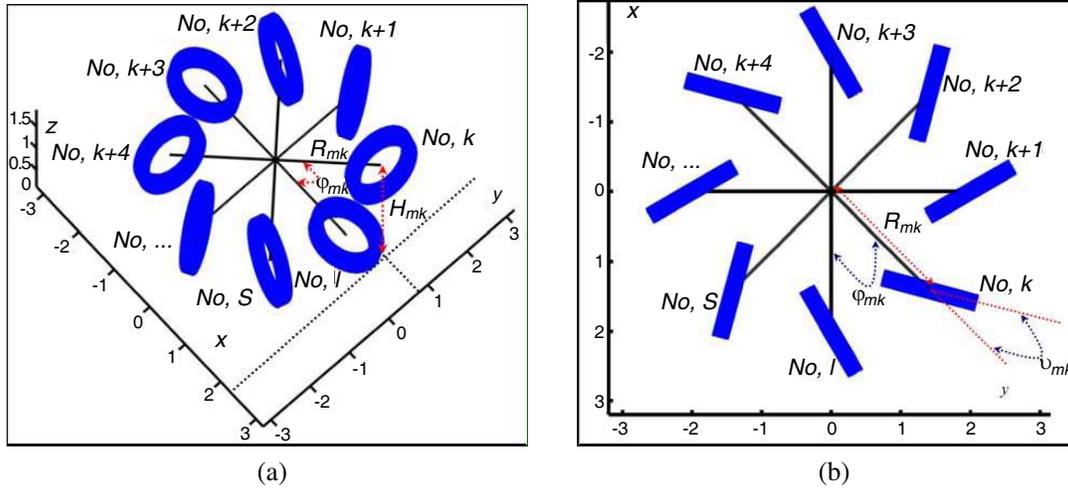


Figure 2: The structure of the modular toroidal coil with 8 solenoid coils. (a) Three dimensional diagram of the modular toroidal coil. (b) Projection of the structure of the modular toroidal coil on x - y plane.

This paper is organized as follow: In Section 2, an appropriate coordinate system to simplify the mathematical equations is presented, the longitudinal components of the rings element of each SC in this coordinate system are introduced, and the relations between the geometrical parameters of the k th SC with the geometrical parameters of the rings located in the same coil are derived. In Section 3, based on the Cubic Meshing a new algorithm is presented to compare the magnetic flux density lines of the MTC for $S = 4$ and $S = 8$.

2. COORDINATE SYSTEM AND LONGITUDINAL COMPONENTS

Figure 3 shows the i th and the j th hypothetical rings of the MTC with the geometrical parameters of $v_i, \varphi_i, R_i, a_i, H_i$ and $v_j, \varphi_j, R_j, a_j, H_j$, respectively. To consider these parameters, the coordinate system should be non-orthogonal, three dimensional and rotational in order to simplify the mathematical equations. An arbitrary point such as P in the space is defined by ρ, θ , and φ (see Fig. 3). In this coordinate system, it is typical to use $\vec{a}_\rho, \vec{a}_\theta$, and \vec{a}_φ to represent unit vectors in directions ρ, θ , and φ , respectively. In Fig. 3, the unit vectors of this coordinate system for the point P , located on the i th coil, is shown. The Cartesian coordinate system of this point can be expressed as Equations (1)–(3) using the projection of the i th ring on the x - y plane as shown in Fig. 4. The longitudinal components of this coordinate system are defined by Equation (4).

$$x_P = (R_i \cos(\varphi - \varphi_i) + a_i \cos \theta \cos(v_i - \varphi + \varphi_i)) \cos \varphi \quad (1)$$

$$y_P = (R_i \cos(\varphi - \varphi_i) + a_i \cos \theta \cos(v_i - \varphi + \varphi_i)) \sin \varphi \quad (2)$$

$$z_P = H_i + a_i \sin \theta \quad (3)$$

$$\begin{aligned} dl_i &= \vec{a}_\varphi (R_i \cos(\varphi - \varphi_i) + a_i \cos \theta \cdot \cos(v_i - \varphi + \varphi_i)) d\varphi + \vec{a}_\theta a_i d\theta + \vec{a}_\rho d\rho \\ &= \vec{a}_x dl_{ix} + \vec{a}_y dl_{iy} + \vec{a}_z dl_{iz} \end{aligned} \quad (4)$$

The longitudinal components of the i th ring element in the Cartesian coordinate system are obtained using Table 1 of Ref. [5] as Equations (5)–(7). Since the i th ring's geometric loci is given by $\rho = a_i$, the longitudinal components of the i th ring by substitution of $d\rho = 0$ can simplify Equations (4) to (7). Furthermore, the relation between θ and φ for the i th ring and the differential of φ , using Fig. 4 can be expressed as Equations (8) and (9). The dependency of the presented equations on the geometrical parameters of the i th ring shows that the relation between the geometrical parameters of k th SC and the geometrical parameters of the rings of the same SC should be defined. In Fig. 5 the latitudinal cross section of the k th SC of the MTC with the geometrical parameters of this coil i.e., $v_{mk}, R_{mk}, \varphi_{mk}, H_{mk}$ is shown. In this figure, the numbers on each ring indicates the sequence of the rings connected in series

$$\begin{aligned} dl_{ix} &= dl_i \cdot \vec{a}_x = -a_i \sin \theta \cos(v_i - \varphi + \varphi_i) \cdot \cos \varphi d\theta + \cos \theta \cos(v_i - \varphi + \varphi_i) \cdot \cos \varphi d\rho \\ &\quad - \sin \varphi (R_i \cos(\varphi - \varphi_i) + a_i \cos \theta \cos(v_i - \varphi + \varphi_i)) d\varphi \end{aligned} \quad (5)$$

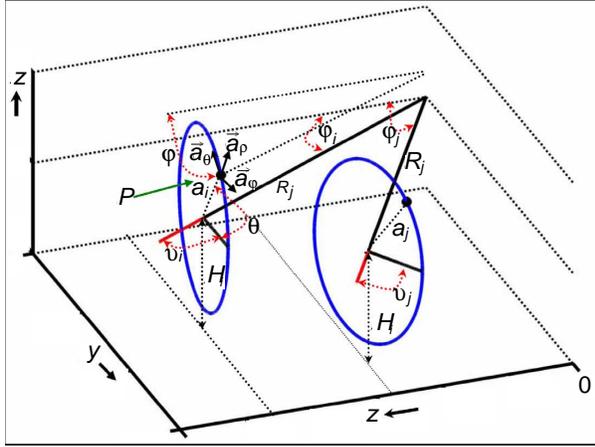


Figure 3: The coordinate system and the two hypothetical rings of the modular toroidal coil.

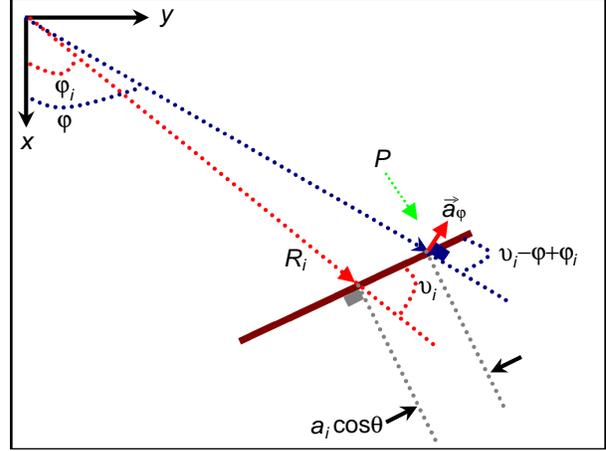


Figure 4: The projection of the \$i\$th ring on the \$x-y\$ plane.

$$dl_{iy} = dl_i \cdot \vec{a}_y = -a_i \sin \theta \cos(\nu_i - \varphi + \varphi_i) \sin \varphi \cdot d\theta + \cos \theta \cos(\nu_i - \varphi + \varphi_i) \sin \varphi d\rho + \cos \varphi \cdot (R_i \cos(\varphi - \varphi_i) + a_i \cos \theta \cos(\nu_i - \varphi + \varphi_i)) d\varphi \quad (6)$$

$$dl_{iz} = dl_i \cdot \vec{a}_z = a_i \cos \theta d\theta + \sin \theta d\rho \quad (7)$$

$$\varphi = \varphi_i + \tan^{-1} \left(\frac{a_i \cos \theta \sin \nu_i}{R_i + a_i \cos \theta \cos \nu_i} \right) \quad (8)$$

$$d\varphi = \frac{-a_i R_i \sin \nu_i \sin \theta}{R_i^2 + a_i^2 \cos^2 \theta + 2a_i R_i \cos \theta \cos \nu_i} d\theta \quad (9)$$

As seen in this figure, the \$k\$th SC is composed of \$N_k M_k\$ rings where \$M_k\$ is the number of the layers and \$N_k\$ represents the number of the rings in each layer. Using filament current method, if each ring of the \$k\$th SC is modeled as filament current, then the inductance matrix is \$N_k M_k\$ by \$N_k M_k\$ [3]. Knowing the dimensional parameters of the \$k\$th SC (\$M_k, N_k, d_k, H_{zk}, R_{ouk}, R_{ink}\$) and the geometrical parameters of the same coil (\$\nu_{mk}, R_{mk}, \varphi_{mk}, H_{mk}\$), the geometrical parameters of the \$i\$th ring i.e., \$\nu_i, \varphi_i, R_i, a_i\$ and \$H_i\$ can be derived as Equations (10)–(14). It is noted that using the filament current method the number of rings of the MTC consisting of \$S\$ SCs, is \$A\$ as given in Equation (15). Consequently, the inductance matrix of the MTC is \$A\$ by \$A\$.

The relation between the geometrical parameters of each ring of the MTC with the geometric and the dimensional parameters of the corresponding SCs are given as Equations (10)–(14), when \$i = 1, \dots, A\$.

$$\eta_i = \tan^{-1}(h_i \cos \nu_{mk} / (R_{mk} - h_i \sin \nu_{mk})) \quad (10)$$

$$\varphi_i = \varphi_{mk} + \eta_i \quad (11)$$

$$\nu_i = \nu_{mk} - \eta_i \quad (12)$$

$$R_i = (R_{mk} - h_i \sin \nu_{mk}) / \cos \eta_i \quad (13)$$

$$H_i = H_{mk} \quad (14)$$

$$A = \sum_{i=1}^S M_i N_i \quad (15)$$

3. MAGNETIC FLUX DENSITY ANALYSIS OF THE MTC

In this section the algorithm of drawing the magnetic flux density lines of the MTC is presented in order to study the leakage fluxes. This algorithm can be easily implemented in MATLAB®. For this purpose, the analytical equations of the magnetic flux density components at any point in the three dimensional space (\$x_q, y_q, z_q\$) using the Biot-Savart equations are presented as Equation (16). Then, by cubic meshing of the three dimensional spaces for all different magnetic materials, the magnetic flux density components can be calculated for any node located on the cubes' vertex,

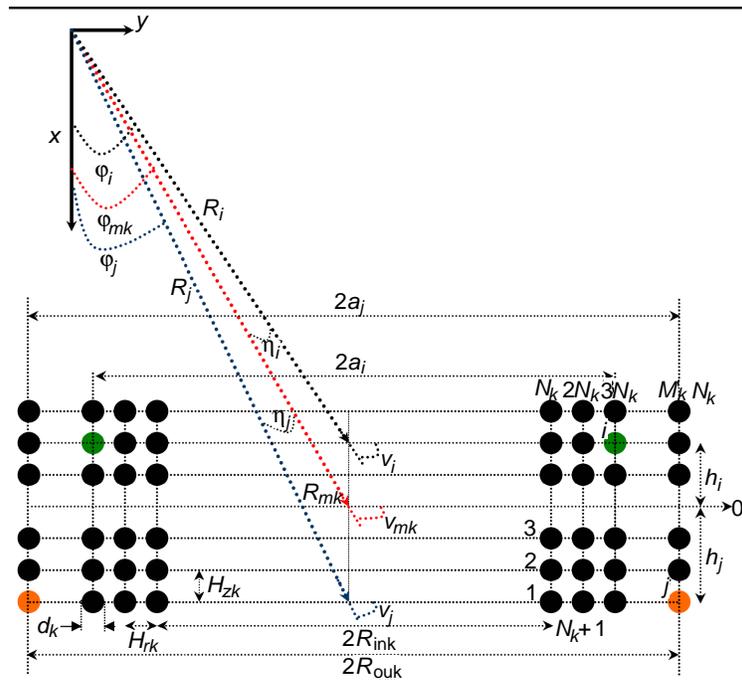


Figure 5: Latitudinal cross section of the k th sc.

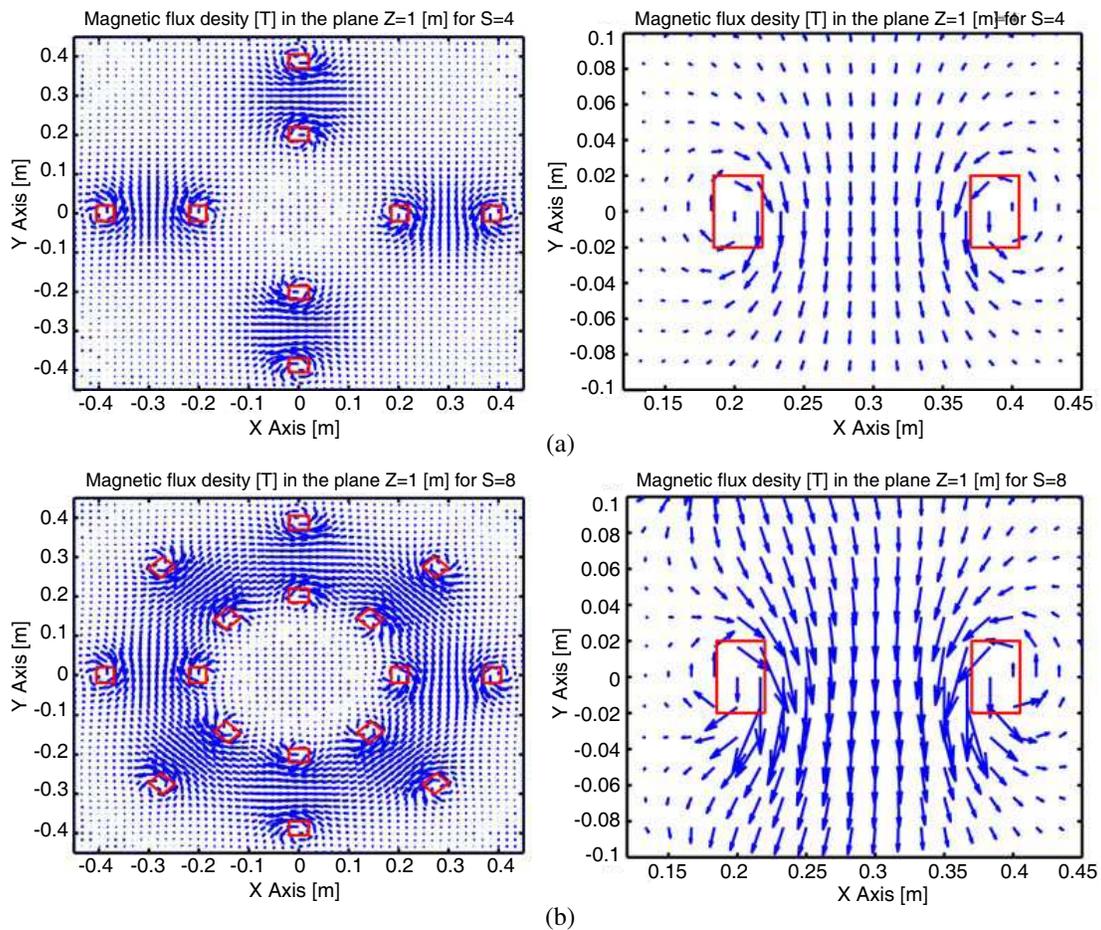


Figure 6: The magnetic flux density lines of the MTC composing of (a) 4 SC (b) 8 SC.

e.g., (x_q, y_q, z_q) . The numerical analysis of the integrations resulting from (16) is solved using the extended three-point Gaussian algorithm [4]. If the meshing step in the direction (x, y, z) is defined as $(\Delta x, \Delta y, \Delta z)$, then Equation (17) can be used to calculate (x_n, y_n, z_n) , the Cartesian coordinate of the vertex of the vector of the magnetic flux density corresponding to the point (x_q, y_q, z_q) . This procedure was used to calculate the Cartesian coordinate of the vertex of the vector of the magnetic flux density corresponding to all nodes. This algorithm was coded in MATLAB[®] m-file to calculate and draw the magnetic flux density lines of the MTC in order to study the leakage flux. In Fig. 6, the magnetic flux density lines of the symmetrical MTC composing of the 4 and 8 SCs, with the dimensional and geometrical parameters shown in Tables 2 and 3 of Ref. [5] are compared. As it is observed as the number of SCs is increased, the leakage flux decreases, and the behavior of the MTC approaches the ideal toroidal coil. It is added the in Fig. 6 the highlighted red rectangular show the geometrical loci of SCs. From the above results it is realized that although the presented algorithm is easy to implement, its results are of high quality and comparable with those of commercial software.

$$\vec{B} = B_x \vec{a}_x + B_y \vec{a}_y + B_z \vec{a}_z = \sum_{i=1}^A \left(\frac{\mu I}{4\pi} \right) \oint_{\rho=a_i} \frac{dl_i \times \vec{a}_{r_{qi}}}{|r_{qi}|^2} = \sum_{i=1}^A \left(\frac{\mu I}{4\pi} \right) \cdot \int_0^{2\pi} (h_x(\theta_i) \vec{a}_x + h_y(\theta_i) \vec{a}_y + h_z(\theta_i) \vec{a}_z) d\theta_i \quad (16)$$

where:

$$\begin{aligned} h_x(\theta_i) &= \frac{(z_q - z_i) dl_{iy} - (y_q - y_i) dl_{iz}}{[(x_q - x_i)^2 + (y_q - y_i)^2 + (z_q - z_i)^2]^{1.5}} \\ h_y(\theta_i) &= \frac{(x_q - x_i) dl_{iz} - (z_q - z_i) dl_{ix}}{[(x_q - x_i)^2 + (y_q - y_i)^2 + (z_q - z_i)^2]^{1.5}} \\ h_z(\theta_i) &= \frac{(y_q - y_i) dl_{ix} - (x_q - x_i) dl_{iy}}{[(x_q - x_i)^2 + (y_q - y_i)^2 + (z_q - z_i)^2]^{1.5}} \end{aligned} \quad (17)$$

$$\begin{aligned} x_n &= x_q + B_{x=x_q} \Delta x / \sqrt{B_{x=x_q}^2 + B_{y=y_q}^2 + B_{z=z_q}^2} \\ y_n &= y_q + B_{y=y_q} \Delta y / \sqrt{B_{x=x_q}^2 + B_{y=y_q}^2 + B_{z=z_q}^2} \\ z_n &= z_q + B_{z=z_q} \Delta z / \sqrt{B_{x=x_q}^2 + B_{y=y_q}^2 + B_{z=z_q}^2} \end{aligned}$$

4. CONCLUSION

The modular toroidal coil (MTC) is widely used in different areas including plasma research (e.g., the Tokamak reactors), the Superconductor Magnetic Energy Storage (SMES), and the nuclear fusion reactors. The theoretical studies in this field are still in development stage. Similar to that of helical toroidal coil, the calculation of the inductance of MTC can be employed to predict the transient state behavior, find the equivalent electrical circuit, and identify the electrical elements of the equivalent electrical circuit of the MTC. In this paper, equations for the calculation of the magnetic flux density of the MTC are presented and their numerical simulation is carried out in MATLAB[®]. It is also shown that the presented algorithm for the analysis of the leakage flux of the MTC using the Cubic Meshing approach due to its simplicity in analysis and ease of implementation of the non-symmetrical and three dimensional objects is advantageous to the commercial software.

REFERENCES

1. Alizadeh Pahlavani, M. R. and A. Shoulaie, "A novel approach for calculations of helical toroidal coils inductance usable in reactor plasmas," *IEEE Trans. on Plasma Science*, Vol. 37, No. 8, 1593–1603, Aug. 2009.
2. Romeo, A. and P. R. Garabedian, "Choice of coils for a fusion reactor," *Proceedings of the National Academy of Sciences*, Vol. 104, No. 30, 12250–12252. Jul. 24, 2007.
3. Mei, S. and Y. I. Ismail, "Modeling skin and proximity effects with reduced realizable RL circuits," *IEEE Trans. on Very Large Scale Integration (VLSI) Systems*, Vol. 12, No. 4, Apr. 2004.
4. Pennington, R. H., *Introductory Computer Methods and Numerical Analysis*, 4th Edition, New York, Macmillan, 1970.

5. Alizadeh Pahlavani, M. R., H. A. Mohammadpour, and A. Shoulaie, “Numerical and experimental analysis of modular toroidal coil inductance applicable to tokamak reactors,” *IEEE Trans. on Plasma Science*, Dec. 2009, accepted to be published.

Electromagnetic Force Distribution on Cylindrical Coils' Body

A. Shiri, M. R. Alizadeh Pahlavani, H. A. Mohammadpour, and A. Shoulaie

Department of Electrical Engineering, Iran University of Science and Technology, Tehran, Iran

Abstract— The paper focuses on distribution of the electromagnetic forces on different parts of the cylindrical coils. In design process of cylindrical coils in many magnetic devices, mechanical stresses exerted on different parts of these kinds of coils should be determined. Results obtained by analytical equations show that the outer turns of the coil in two axial ends are exposed to largest tension.

1. INTRODUCTION

Due to the extensive application of cylindrical coils in industry such as linear tubular motors, magnetic launchers and casting industries [1–3], the distribution of the mechanical stresses on different parts of the former is necessary. The calculation of the magnetic force distribution on current-carrying coils is closely related to the calculation of the magnetic force between them. To calculate the force between these coils, a variety of methods have been proposed in literature. In 1962 Grover presented some empirical equations and tables to calculate the force between coils with different shapes [4]. In other studies, the variation of the mutual inductance between two coils is used to calculate the force between them. In the latter method, first the mutual inductance between two coils is calculated and then used to calculate the force [5–7]. There are many articles on the calculation of the magnetic forces between magnetic coils. In references [8–10] magnetic force between spiral coils is calculated using a new and effective method. In other research, magnetic forces between cylindrical coils are evaluated [11, 12]. In this paper, using the method developed in [10], the magnetic force distribution on different parts of the cylindrical coils is calculated.

2. THE MAGNETIC FORCE BETWEEN TWO CONCENTRIC CIRCULAR FILAMENTS

Suppose a system of two current carrying rings as shown in Figure 1(a). To calculate the force between them, the concept of vector magnetic potential is employed. The vector magnetic potential of ring 1 in any point P on ring 2 is equal to [13]:

$$\vec{A} = \frac{\mu_0 I_1}{4\pi} \oint_{C_1} \frac{d\vec{l}'}{R_1} \quad (1)$$

where μ_0 is the vacuum permeability, I_1 is the current of ring 1, C_1 is the length of ring 1 and R_1 is the distance between the differential component of the source $d\vec{l}'$ at point p' and the field point p (Figure 1(b)).

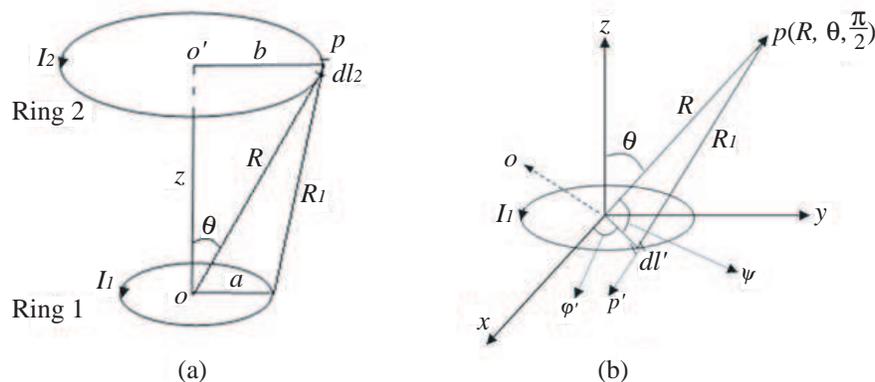


Figure 1: (a) Two current carrying concentric rings in z distance of each other, (b) determination of vector potential of a current carrying ring with radius a in any given point p .

By obtaining the vector magnetic potential, the flux density is calculated using the following equation [13]:

$$\vec{B} = \vec{\nabla} \times \vec{A} \quad (2)$$

To calculate the force of ring 1 exerted on ring 2, the following equation is employed [13]:

$$\vec{F}_{21} = I_2 \oint_{C_2} d\vec{l}_2 \times \vec{B} \quad (3)$$

where I_2 is the current of ring 2. Using Equations (1) and (2) in Equation (3) and doing some simple mathematical calculations, the following equation for the force is obtained:

$$\vec{F}_{21} = \vec{a}_z \left(\frac{\mu_0 I_1 I_2 z k}{2\sqrt{ab}(1-k^2)} \right) \left[(1-k^2) K(k) - \left(1 - \frac{1}{2}k^2 \right) E(k) \right] \quad (4)$$

In the above equation, a and b are radii of rings 1 and 2, respectively, z is axial distance of two rings and k ($0 < k < 1$) is a constant coefficient and is equal to:

$$k = \sqrt{\frac{4ab}{(a+b)^2 + z^2}} \quad (5)$$

and $K(k)$ and $E(k)$ are first and second order Elliptic Integrals respectively, which are defined as:

$$K(k) = \int_0^{\frac{\pi}{2}} \frac{d\theta}{(1-k^2 \sin^2 \theta)^{1/2}} \quad (6)$$

$$E(k) = \int_0^{\frac{\pi}{2}} (1-k^2 \sin^2 \theta)^{1/2} d\theta \quad (7)$$

3. THE MAGNETIC FORCE DISTRIBUTION ON DIFFERENT PARTS OF CYLINDRICAL COIL

In order to calculate the mechanical stresses exerted on different parts of the coil resulting from its current, the mesh-matrix method is employed. Suppose a coil with the turn number of N shown in Figure 2, where r_0 is the inner radius, b is the radial thickness and a is the height of the coil. As seen in this figure, the cross-section of the coil is divided into several segments. Here, the coil is divided into $n_r \times n_a$ cells. To calculate the distribution of the force on the different parts of the coil, the force between different filaments (in Figure 2, each filament is specified with two cells in both sides) of the coil is calculated and then added together.

The force between filaments j and l is calculated using Equation (4) as follows:

$$f_{jl} = a_z \left(\frac{\mu_0 i^2 z_{jl} k'}{2\sqrt{r_j r_l} (1-k'^2)} \right) \left[(1-k'^2) K(k') - \left(1 - \frac{1}{2}k'^2 \right) E(k') \right] \quad (8)$$

In the above equation, r_j and r_l are the radii of the filaments j and l , respectively, and z_{jl} is the center to center distance of the two filaments. The current of each filament is supposed to be concentrated on its center, and the current density of the entire coil is supposed to be uniform, and i which can be calculated using the following equation, is the current of each filament in the coil:

$$i = \frac{NI}{n_r \times n_a} \quad (9)$$

In the above equation, I is the current of the coil. Parameter k' in Equation (8) is a constant and is equal to:

$$k' = \sqrt{\frac{4r_j r_l}{(r_j + r_l)^2 + z_{jl}^2}} \quad (10)$$

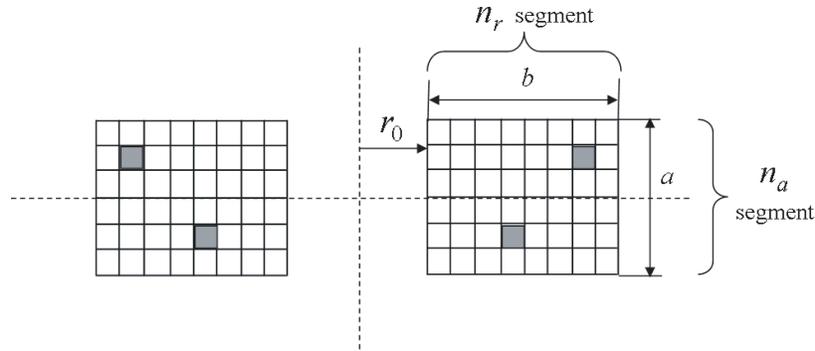


Figure 2: Division of the coil into different meshes to calculate the force distribution.

4. RESULTS

As mentioned before, the force between two filaments can be calculated using Equation (8). In order to calculate the force exerted on specific part of the coil, one can use the effect of all filaments on that part and sum them up. The characteristics of the coil used for calculations are shown in Table 1. The cross-section of the coil has been divided into 71×35 segments. The current of the coil is supposed to be 12 Amperes. Axially net magnetic force exerted on different parts of the coil is calculated. The force distribution on the outer part (turns located in outer part) of the coil in axial direction is illustrated in Figure 3(a). As seen in this figure, the magnitude of the force is symmetrical in respect to axial axis of the coil in such a way that there is no stress on the middle part of the coil. This is so because, the forces from both sides exerted on the middle part of the coil are equal in magnitude and opposite in direction and so cancel each other. Figure 3(b) shows the distribution of the force on radial direction. According to this figure, moving across radial direction, the force exerted on related parts increases and after reaching a maximum value, it starts to decrease at the outer parts. In order to compare the force distribution on all parts of the coil's body, the force profile among the axial and radial dimensions of the coil is illustrated in Figure 4. It is clear from this figure that the outer parts of the coil at two axial axis ends are the most critical portions exposed to the largest tension. For instance, a force of about 0.6 N is exerted on the outer parts of the coil located at about 9.5 cm in axial axis while the current following the coil is 12 amperes. The current can increase in fault conditions and give rise to deformation of the outer parts of the coil at two axial axis ends.

Table 1: Characteristics of the coil used in calculations.

Number of Turns	Inner radius, r_0 (cm)	Radial thickness, b (cm)	Height, a (cm)
2450	5	6	12

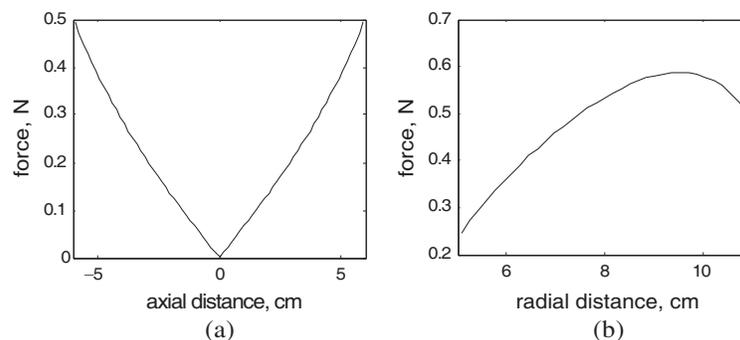


Figure 3: (a) Force distribution on the outer part of the coil in axial direction, (b) force distribution on lower part of the coil in radial direction.

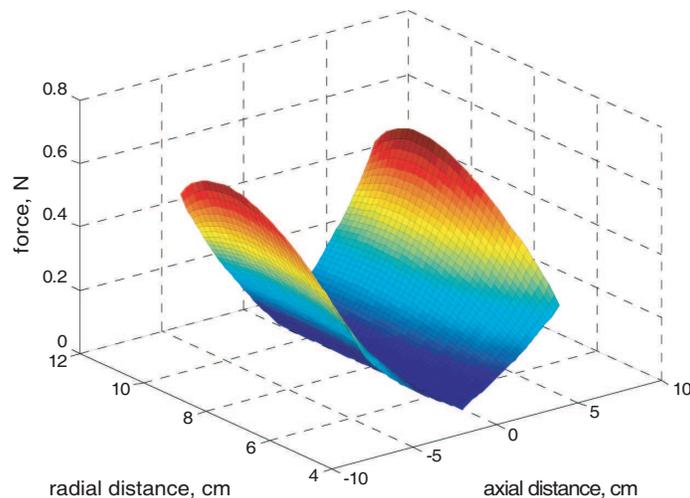


Figure 4: The force distribution on all parts of the coil's body.

5. CONCLUSION

In this paper the electromagnetic force distribution on cylindrical coils' body is calculated using exact analytical equations. According to the results obtained by MATLAB coding, the maximum stress is exerted on the outer parts of the coil at two axial axis ends. Thus, it is necessary to sufficiently support those parts of the coil in any application. On the other hand, considering the fact that the forces from both sides exerted on the middle parts of the coil cancel each other, these parts of the coil are exposed to low tension.

REFERENCES

1. Fujisaki, K., "Magnetohydrodynamic stability in pulse electromagnetic casting," *IEEE Trans. Ind. Appl.*, Vol. 39, No. 5, 1442–1447, September/October 2003.
2. Meriched, A., M. Feliachi, and H. Mohellebi, "Electromagnetic forming of thin metal sheets," *IEEE Trans. Magn.*, Vol. 36, No. 4, 1808–1811, July 2000.
3. Engel, T. G., D. Surls, and W. C. Nunnally, "Prediction and verification of electromagnetic forces in helical coil launchers," *IEEE Trans. Magn.*, Vol. 39, No. 1, 112–115, January 2003.
4. Grover, F. W., *Inductance Calculations, Working Formulas and Tables*, Dover Publications, USA, New York, 1962.
5. Babic, S., C. Akyel and S. J. Salon, "New procedures for calculating the mutual inductance of the system: Filamentary circular coil-massive circular solenoid," *IEEE Trans. Magn.*, Vol. 39, No. 3, 1131–1134, May 2003.
6. Kim, K. B., E. Levi, Z. Zabbar, and L. Birenbaum, "Mutual inductance of noncoaxial circular coils with constant current density," *IEEE Trans. Magn.*, Vol. 33, No. 5, 4303–4309, September 1997.
7. Akyel, C., S. Babic, and S. Kinicic, "New and fast procedures for calculating the mutual inductance of coaxial circular coils (circular coil — disk coil)," *IEEE Trans. Magn.*, Vol. 38, No. 5, 2367–2369, September 2002.
8. Shiri, A., "Calculation of magnetic forces between current carrying coils in two plates," M.Sc. Thesis, Iran University of Science & Technology, September 2006 (in Persian).
9. Shiri, A. and A. Shoulaie, "calculation of magnetic forces between spiral coils using mesh approach," *24th International Power System Conference*, Tehran, Iran, November 2009 (in Persian).
10. Shiri, A. and A. Shoulaie, "A new methodology for magnetic force calculations between planar spiral coils," *Progress In Electromagnetics Research*, PIER 95, 39–57, 2009.
11. Shiri, A. and A. Shoulaie, "New results in calculation of magnetic forces between cylindrical coils", *23th International Power System Conference*, Tehran, Iran, November 2008 (in Persian).
12. Shiri, A., M. R. Alizadeh Pahlavani, and A. Shoulaie, "A new and fast procedure for calculation of the magnetic forces between cylindrical coils," *International Review of Electrical Engineering (I.R.E.E.)*, Vol. 4, No. 5, Part B, October 2009.
13. Cheng, D. K., *Field and Wave Electromagnetics*, 2nd Edition, Addison Wesley, 1989.

Magnetic Flux Density Analysis of Helical Toroidal Coil Using Finite Element Approach

M. R. Alizadeh Pahlavani, A. Shiri, H. A. Mohammadpour, and A. Shoulaie

Department of Electrical Engineering, Iran University of Science and Technology (IUST), Tehran, Iran

Abstract— In this paper, analytical relations for magnetic flux density components of helical toroidal coils (HTC) are presented. The numerical integrations resulting from these equations are done using the extended three-point Gaussian algorithm. The obtained results using numerical simulations coincide with the results of particle-in-cell (PIC) method. The Finite Element (FE) approach is employed to present an algorithm to study the three dimensional leakage flux distribution pattern of the coil and to draw the magnetic flux density lines of the HTC. The presented algorithm, due to its simplicity in analysis and ease of implementation of the non-symmetrical and three dimensional objects, is advantageous to the commercial software such as ANSYS, MAXWELL, and FLUX. Finally, using the presented algorithm, magnetic flux density lines in several examples are drawn.

1. INTRODUCTION

Recent research work in Superconducting Magnetic Energy Storage (*SMES*) systems, nuclear fusion reactors, and plasma reactors such as the Tokamak has suggested the use of an advanced coil with a helical toroidal structure [1]. The main reason for these suggestions is the ability to implement special target functions for this coil in comparison with other structures such as the toroidal, the solenoidal, and the spherical coils [2]. The structure of this coil is shown in Fig. 1. In this coil, the ratio of the major radius to the minor radius ($A = R/a$), the number of turns in a ring (N), and the number of rings in a layer (v) are called aspect ratio, poloidal turns (or pitch number), and helical windings, respectively. For example, the coil in Fig. 1 is composed of five helical windings ($v = 5$) with nine poloidal turns ($N = 9$). The inductance formulas show that parameters a , R , and N of the helical toroidal coil can be used as design parameters to satisfy special target functions. With respect to the fact that each ring of the coil generates both toroidal and poloidal magnetic fields simultaneously, the coil can be regarded as a combination of coils with toroidal and solenoidal fields. Furthermore, the coil can be designed in a way to eliminate the magnetic force component in both the major and the minor radius directions. These are called Force Balanced Coils (*FBC*) and Stress Balanced Coils (*SBC*), respectively. Also the coils which utilize the virial theorem to balance these two force components are called Virial Limited Coils (*VLC*) [3]. In some applications, helical toroidal coils are used in a double layer manner with two different winding directions (respectively with two different or the same current directions in each layer) to reduce the poloidal leakage flux compared to the toroidal leakage flux or versa. In this work, the investigation is focused on the monolayer helical toroidal coil.

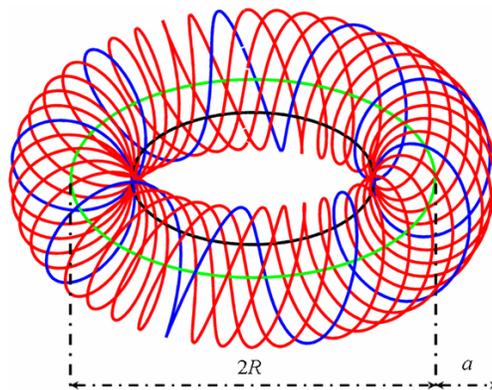


Figure 1: The structure of a monolayer helical toroidal coil with five rings of nine turns.

In general, any simulation program that simultaneously solves equations the particle position and its velocity can be called a Particle-in-Cell (PIC) simulation [4]. The name PIC comes from

the way of assigning macro-quantities (like density, current density, and so on) to the simulation particles. Inside the plasma community, PIC codes are usually associated with solving the equation of motion of particles with the Newton–Lorentz’s force. PIC codes are usually classified depending on the dimensionality of the code and base on the set of Maxwell’s equations used. The codes solving an entire set of Maxwell’s equations are called electromagnetic codes, while electrostatic ones solve just the Poisson equation. Some advanced codes are able to switch between different dimensional and coordinate systems, and use electrostatic, or electro-magnetic models. PIC simulation starts with an initialization and ends with the output of results. This part is similar to the input-output routines of any other numerical tool. Usually, the numerical methods based on the PIC simulation are obtained from the solution of partial differential-algebraic equations for example by the Fourth-order Runge-Kutta method. Considering the number of the particles which are of order of 10^{10} , the simulations based on PIC methods take long time to solve the above equations. Usually, in order to resolve the time issue in PIC methods, special computers may be employed. On the other hand, this paper used Biot-Savart equation and the mathematical equations of the current path in the conductor of the helical toroidal coil in order to obtain the magnetic flux density components. The numerical integrals resulting from the Biot-Savart equations are solved using the extended three point Gaussian algorithms. This method has the least error among all numerical integration methods. Also, the method used by the authors, contrary to the PIC method, does not need any special computers to solve the equations. In Section 2 of [5–8], the semi-toroidal coordinate system and the longitudinal components of a ring element are briefly discussed. In Sections 3 and 4 of [5–8], the formulas of the magnetic flux density components are presented, and the behavior of these components is simulated, respectively. The explanation of the particle-in-cell method (PIC) as well as the numerical simulation results of this method is presented in Section 2 of this paper. In Section 3 of the paper, using finite element approach, an algorithm is introduced in order to analyze the magnetic flux density of the helical toroidal coil, and the related simulation results are also presented.

2. ANALYSIS OF THE MAGNETIC FLUX DENSITY USING PARTICLE IN CELL METHOD

Magnetic flux density analysis using PIC method is based on solving the equation of particles’ position and their velocity, which are confined to the helical toroidal coil. The name PIC comes from the way of assigning macro-quantities (like density, current density, and so on) to the simulation particles. The position and velocity equations for N particles are as follows:

$$d\vec{X}_i/dt = \vec{V}_i, \quad i = 1, \dots, N \tag{1}$$

$$d\vec{V}_i/dt = \vec{F}_i(t, \vec{X}_i, \vec{V}_i, A), \quad A = L_1(B), \quad B = L_2(\vec{X}_1, \vec{V}_1, \vec{X}_2, \vec{V}_2, \vec{X}_3, \vec{V}_3, \dots, \vec{X}_N, \vec{V}_N) \tag{2}$$

Usually, any coding for simulation of macro-quantities, which are generated using Equation (2), is called PIC. In the above equation, A and B introduce macro-quantities, L_1 and L_2 are mathematical operations, and \vec{F}_i is the applied force to the i th particle. PIC analysis is widely utilized in many different fields like Plasma Physics, Quantum Physics, and semi-conductor. In plasma community, Equation (2) can be rewritten as Equation (3). Also, macro-quantities can be expressed as Maxwell equations of particles using Equations (4)–(6).

$$d\vec{V}_i/dt = e_i \left(\vec{E}(\vec{X}_i) + \vec{V}_i \times \vec{B}(\vec{X}_i) \right) / m_i \tag{3}$$

$$\nabla \vec{D} = \rho(\vec{r}, t), \quad \partial \vec{B} / \partial t = -\nabla \times \vec{E}, \quad \vec{D} = \epsilon \vec{E} \tag{4}$$

$$\nabla \vec{B} = 0, \quad \partial \vec{D} / \partial t = \nabla \times \vec{H} - \vec{J}(\vec{r}, t), \quad \vec{B} = \mu \vec{H} \tag{5}$$

$$\rho = \rho(\vec{X}_1, \vec{V}_1, \vec{X}_2, \vec{V}_2, \vec{X}_3, \vec{V}_3, \dots, \vec{X}_N, \vec{V}_N) \quad \vec{J} = \vec{J}(\vec{X}_1, \vec{V}_1, \vec{X}_2, \vec{V}_2, \vec{X}_3, \vec{V}_3, \dots, \vec{X}_N, \vec{V}_N) \tag{6}$$

In PIC analysis, these differential equations are solved using numerical methods such as the Fourth-order Runge-Kutta method. In this analysis, considering the number of the particles which are of the order of 10^{10} , the simulations based on PIC methods take long time to solve the above equations and need special computers. To solve this problem, the Section 4 of [5–8] studies the magnetic flux density without using these special computers with very shorter computational time. In Fig. 2, the magnetic flux density lines of toroidal components inside the helical toroidal coil for

$N = 2$, $v = 6$ are drawn [4]. This figure shows that this component is rotational as was shown in Section 4 of [5–8].

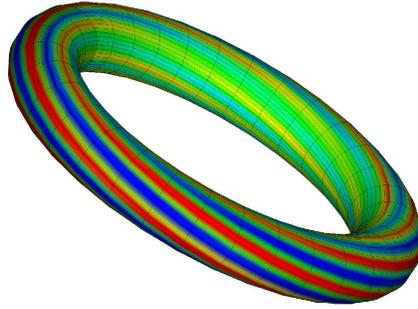


Figure 2: Magnetic flux density lines for toroidal component inside the helical toroidal coil using PIC method.

3. FIELD ANALYSIS OF HELICAL TOROIDAL COIL USING FEM APPROACH

In this section an algorithm for drawing the magnetic flux density lines of the HTC is presented in order to study the leakage fluxes. This algorithm can be easily implemented in MATLAB[®]. For this purpose, the analytical equations of the magnetic flux density components at any point in the three dimensional space (x_q, y_q, z_q) using the Biot-Savart equations (Equation (8) for v rings of N turn of [7]) are presented as Equation (7). Then, by cubic meshing of the three dimensional space for all different magnetic materials, the magnetic flux density components can be calculated for any node located on the cubes' vertex, e.g., (x_q, y_q, z_q) . If the meshing step in the direction of (x, y, z) is defined as $(\Delta x, \Delta y, \Delta z)$, then Equation (8) can be used to calculate (x_n, y_n, z_n) . It should be mention that (x_n, y_n, z_n) is the vertex of the vector of the magnetic flux density corresponding to the point (x_q, y_q, z_q) .

$$\vec{B} = B_{x=x_q} \vec{a}_x + B_{y=y_q} \vec{a}_y + B_{z=z_q} \vec{a}_z = \sum_{i=1}^v ({}^i B_{x=x_q} \vec{a}_x + {}^i B_{y=y_q} \vec{a}_y + {}^i B_{z=z_q} \vec{a}_z) \quad (7)$$

This procedure was used to calculate of the vertex of the vector of the magnetic flux density corresponding to all nodes in the Cartesian coordinate. This algorithm was coded in MATLAB[®] m-file to calculate and draw the magnetic flux density lines of the HTC in order to study the leakage flux.

$$\begin{aligned} x_n &= x_q + B_{x=x_q} \Delta x / \sqrt{B_{x=x_q}^2 + B_{y=y_q}^2 + B_{z=z_q}^2} \\ y_n &= y_q + B_{y=y_q} \Delta y / \sqrt{B_{x=x_q}^2 + B_{y=y_q}^2 + B_{z=z_q}^2} \\ z_n &= z_q + B_{z=z_q} \Delta z / \sqrt{B_{x=x_q}^2 + B_{y=y_q}^2 + B_{z=z_q}^2} \end{aligned} \quad (8)$$

Considering the explanations given above, for the entire nodes of cubic meshed three-dimensional space, the end point of the vector of the magnetic flux density corresponding to the same node can be calculated using Equations (7) and (8), and the corresponding vector can also be drawn. In Fig. 3, the magnetic flux density lines for two helical toroidal coils with $v = 3$ for $N = 2, 3$ in the toroidal plane of $\varphi' = 0, \pi/2$ using coding in MATLAB program and Equations (7) and (8) are compared with each other. As expected, by approaching the conductors geometrical loci, the magnitude of the magnetic flux density increases. Also, in each toroidal plane if N is even symmetry in the flux density lines with regard to Z axis is obviously seen while when N is odd, there is no symmetry. This figure like Figs. 3 to 5 shows of [7] that when the ring rotates, the magnetic flux density is also rotates. In these figures, the conductors are shown by colored circles, and each color represents one ring.

In Fig. 4, the magnetic flux density lines of the same two helical toroidal coils in the xy poloidal plane by using Equations (7) and (8) are drawn and compared with each other. This figure shows that when Nv increases, the toroidal leakage flux decreases. The obtained results show that the presented equations for field analyzing of the helical toroidal coil using FE approach considering

its simplicity in implementation and also not to need for modeling of asymmetry three dimensional geometrical objects has a substantial number of advantages in comparison with the available commercial software.

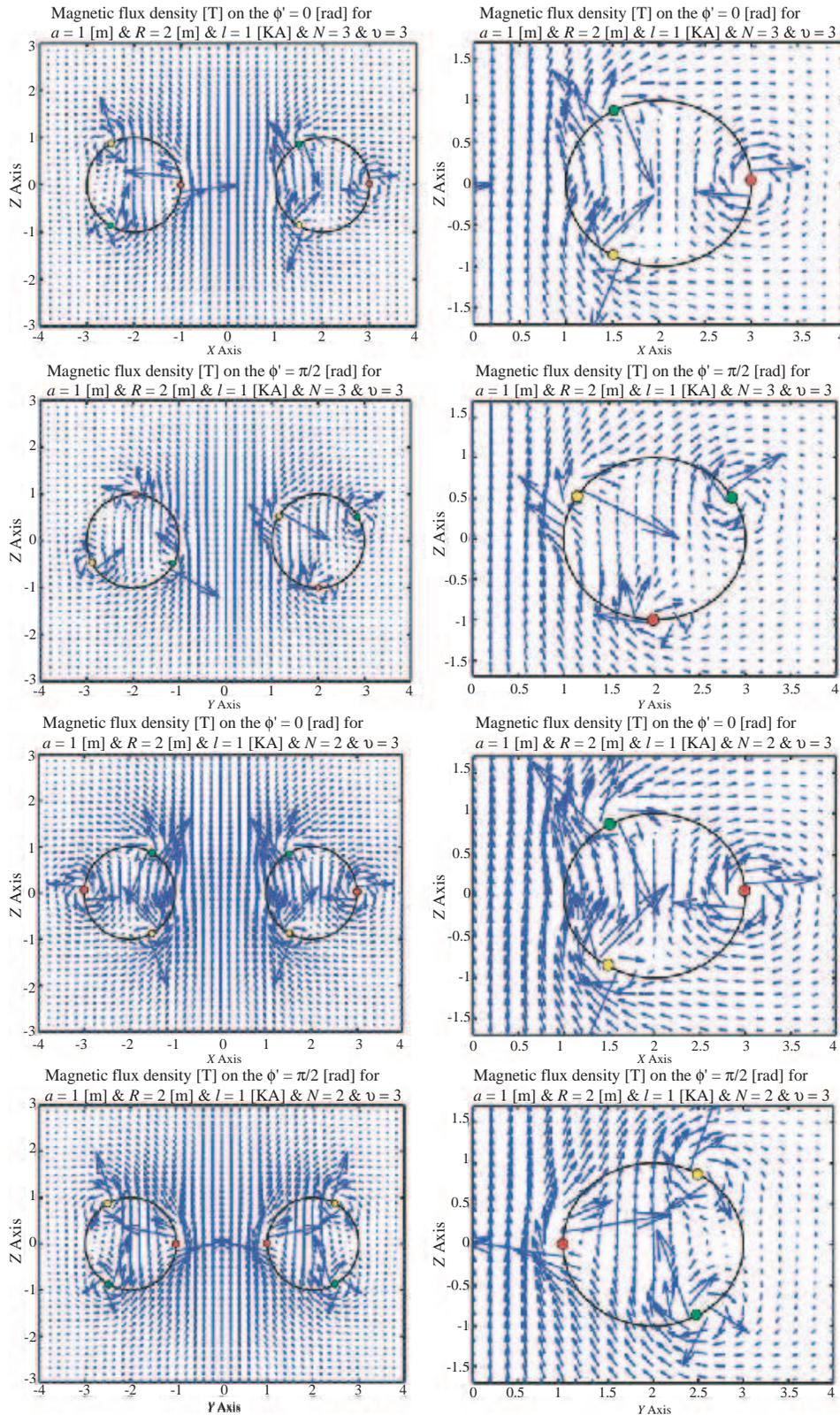


Figure 3: Magnetic flux density lines of helical toroidal coil.

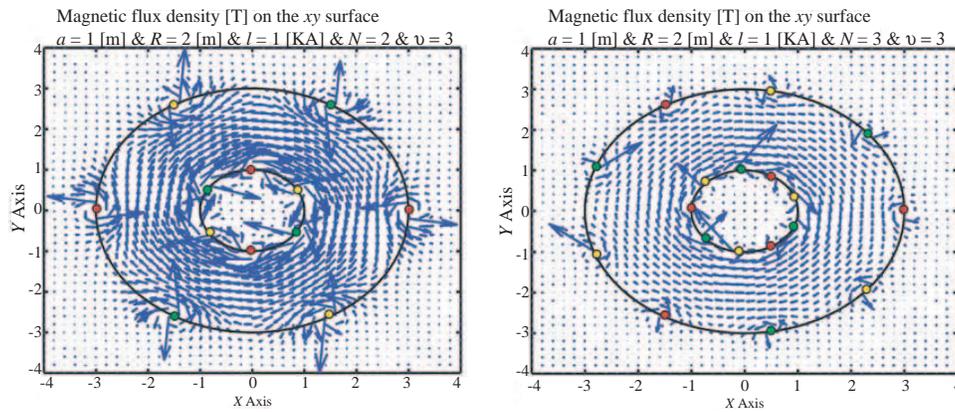


Figure 4: The magnetic flux density of helical toroidal coil.

4. CONCLUSION

Helical toroidal coils are superior to other coils and are extensively used in superconducting magnetic energy storage systems, nuclear fusion reactors, tokamak reactor, and plasma research work. Considering the complexity of the coils and the fact that not many investigations has been carried out in this field, this area of research is still open for much more academic work to come. Since the calculation of the inductance of the helical toroidal coil is based on the field analysis of the magnetic flux density, in this paper the magnetic flux density components of the helical toroidal coil are calculated by in two approach; Bio-Savart equations and PIC method, and the first approach is simulated in MATLAB program. Comparing the simulation results obtained by these two approach shows that the obtained results for both the Bio-Savart and PIC methods coincide with each other while the Bio-Savart method needs less time to be computed and does not need to special computers. In addition, an algorithm is presented in order to study the magnetic flux density. It is shown that this algorithm based on FE approach is much preferable than available commercial software because of its simplicity and not to need for modeling of asymmetry three dimensional geometrical object.

REFERENCES

1. Fabbri, M., D. Ajiki, F. Negrini, R. Shimada, H. I. Tsutsui, and F. Venturi, "Tilted toroidal coils for superconducting magnetic energy storage systems," *IEEE Trans. on Magnetic*, Vol. 39, No. 6, 3546–3550, Nov. 2003.
2. Nomura, S., N. Watanabe, C. Suzuki, H. Ajikawa, M. Uyama, S. Kajita, Y. Ohata, H. Tsutsui, and S. Tsuji-Iio, "Advanced configuration of superconducting magnetic energy storage," *Energy Elsevier*, Japan, 2004.
3. Tsutsui, H., S. Nomura, and R. Shimada, "Optimization of SMES coil by using virial theorem," *IEEE Trans. on Applied Superconductivity*, Vol. 12, No. 1, 800–803, Mar. 2002.
4. Tskhakaya, D., K. Matyash, R. Schneider, and F. Taccogna, "The particle-in-cell method," *Contrib. Plasma Phys.*, Vol. 47, No. 8–9, 563–594, 2007.
5. Alizadeh Pahlavani, M. R. and A. Shoulaie, "Elimination of compressive stress in helical toroidal coils advanced structure for reactor plasmas," *IEEE Trans. on Plasma Science*, Vol. 37, No. 11, 2166–2177, Nov. 2009.
6. Alizadeh Pahlavani, M. R. and A. Shoulaie, "Behavioral study of stress in toroidal, solenoidal, and helical toroidal coils with similar ring structures for reactor plasma," *International Review of Electrical Engineering*, Vol. 4, No. 1, 146–158, Jan.–Feb. 2009.
7. Alizadeh Pahlavani, M. R. and A. Shoulaie, "A novel approach for calculations of helical toroidal coils inductance usable in reactor plasmas," *IEEE Trans. on Plasma Science*, Vol. 37, No. 8, 1593–1603, Aug. 2009.
8. Alizadeh Pahlavani, M. R., H. A. Mohammadpour, and A. Shoulaie, "Numerical and experimental analysis of modular toroidal coil inductance applicable to tokamak reactors," *IEEE Trans. on Plasma Science*, accepted for publication, Dec. 2009.

Numerical Simulations and Analysis of Electromagnetic Scattering from a PEC Target below a Two-layered Dielectric Rough Surfaces: Vertical Polarization

A.-Q. Wang, L.-X. Guo, and C. Chai

School of Science, Xidian University, Xi'an, Shaanxi, China

Abstract— To analyze the properties of the electromagnetic (EM) scattering from a target located below a two-layered rough surfaces, the method of moments (MOM) is introduced in this paper. The incident wave is considered in the vertical polarization. The discrepancies of EM scattering from this composite model and two-layered rough surfaces only are shown firstly. Emphasis is put on discussing the influences of cylinder and surface parameters, such as the radius and altitude of the cylinder, the root mean square (rms) height and correlation length of the Gaussian rough surfaces on the bistatic scattering coefficient.

1. INTRODUCTION

The EM scattering from target and rough surfaces has large number of applications in the region of radar surveillance, target tracking, and oceanic remote sensing [1–4]. It is significant to study the interactions between target and multi-layered surfaces in the region of detecting the landmines, pipes, or other buried targets in the soil, and so on. The extended boundary condition method has been developed to solve the EM scattering from a cylinder buried in layered rough surfaces [3]. The MOM is applied to calculate the bistatic scattering from two-layered rough surfaces [4]. A composite model of two-dimensional PEC cylinder located below a stack of two one-dimensional Gaussian rough surfaces separating homogeneous medium is set up in this paper. The tapered wave suggested by Thorsos [5] is used to avoid the artificial edge diffraction resulting from the finite length of the Monte Carlo simulations. And the bi-conjugate gradient method is used to solve the matrix equations for the asymmetric impedance matrixes of this composite model.

2. THEORETICAL FORMULATIONS

Assuming a vertical polarized wave $\psi^{\text{inc}}(\mathbf{r})$ impinging upon a geometric model as shown in Fig. 1, where the height profile functions of the two Gaussian rough surfaces are $z = f_+(x)$ and $z = f_-(x)$ with parameters δ_{\pm} , l_{\pm} . θ_i and θ_s represent the incident and scattered angle, respectively. The height profile functions satisfy $\langle f_+(x) \rangle = 0$ and $\langle f_-(x) \rangle = -d$, where $d > 0$ is the average height between the two rough surfaces. Then the spaces are separated by the two rough surfaces to $\Omega_0(\varepsilon_0, \mu_0)$, $\Omega_1(\varepsilon_1, \mu_1)$ and $\Omega_2(\varepsilon_2, \mu_2)$. In this paper we assume $\mu_2 = \mu_1 = \mu_0$. Let the total fields in the three spaces $\{\Omega_j\}_{j=0,1,2}$ as ψ_j , and their normal derivatives as $\partial\psi_j/\partial n_{\pm}$ and $\partial\psi_j/\partial n_0$. The integral equations of EM scattering from PEC target located below two-layered dielectric rough

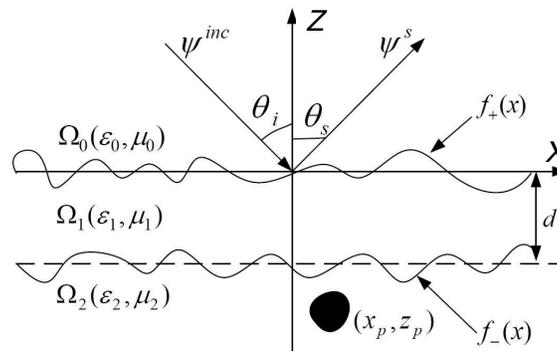


Figure 1: Geometric model of target located below two-layered Gaussian rough surfaces.

surfaces under vertical polarized incidence can be illustrated as

$$\frac{1}{2}\psi_0(\mathbf{r}) = \psi^{\text{inc}}(\mathbf{r}) + \int_{S_+} \left[\psi_0(\mathbf{r}') \frac{\partial G_0(\mathbf{r}, \mathbf{r}')}{\partial n_+} - G_0(\mathbf{r}, \mathbf{r}') \frac{\partial \psi_0(\mathbf{r}')}{\partial n_+} \right] dS' \quad \mathbf{r} \in \Omega_0 \quad (1a)$$

$$\begin{aligned} \frac{1}{2}\psi_1(\mathbf{r}) = & - \int_{S_+} \left[\psi_1(\mathbf{r}') \frac{\partial G_1(\mathbf{r}, \mathbf{r}')}{\partial n_+} - G_1(\mathbf{r}, \mathbf{r}') \frac{\partial \psi_1(\mathbf{r}')}{\partial n_+} \right] dS' \\ & + \int_{S_-} \left[\psi_1(\mathbf{r}') \frac{\partial G_1(\mathbf{r}, \mathbf{r}')}{\partial n_-} - G_1(\mathbf{r}, \mathbf{r}') \frac{\partial \psi_1(\mathbf{r}')}{\partial n_-} \right] dS' \quad \mathbf{r} \in \Omega_1 \end{aligned} \quad (1b)$$

$$\frac{1}{2}\psi_2(\mathbf{r}) = - \int_{S_-} \left[\psi_2(\mathbf{r}') \frac{\partial G_2(\mathbf{r}, \mathbf{r}')}{\partial n_-} - G_2(\mathbf{r}, \mathbf{r}') \frac{\partial \psi_2(\mathbf{r}')}{\partial n_-} \right] dS' + \int_{S_0} \psi_2(\mathbf{r}') \frac{\partial G_2(\mathbf{r}, \mathbf{r}')}{\partial n_0} dS' \quad \mathbf{r} \in \Omega_2 \quad (1c)$$

where, $\{G_j(\mathbf{r}, \mathbf{r}') = \frac{i}{4}H_0^{(1)}(k_j|\mathbf{r} - \mathbf{r}'|)\}_{j=0,1,2}$ is the two-dimensional Green's function in the space Ω_j . $H_0^{(1)}(\cdot)$ is the zeroth-order Hankel function of the first kind. And in this paper, the time dependence is $e^{-i\omega t}$ and the two dimensional position vector is $\mathbf{r} = x\hat{\mathbf{x}} + z\hat{\mathbf{z}}$.

When point \mathbf{r} is located on the two rough surfaces, the fields ψ_j and their normal derivatives $\partial\psi_j/\partial n_{\pm}$ satisfy the following boundary conditions [4]

$$\psi_0(\mathbf{r}) = \psi_1(\mathbf{r}) \quad \frac{\partial\psi_0(\mathbf{r})}{\partial n_+} = \frac{1}{\rho_{10}} \frac{\partial\psi_1(\mathbf{r})}{\partial n_+} \quad \mathbf{r} \in S_+ \quad (2a)$$

$$\psi_1(\mathbf{r}) = \psi_2(\mathbf{r}) \quad \frac{\partial\psi_1(\mathbf{r})}{\partial n_-} = \frac{1}{\rho_{21}} \frac{\partial\psi_2(\mathbf{r})}{\partial n_-} \quad \mathbf{r} \in S_- \quad (2b)$$

and, $\rho_{10} = \varepsilon_1/\varepsilon_0$, $\rho_{21} = \varepsilon_2/\varepsilon_1$.

If each rough surface is discretized into N segments according to their location in x axis, the target is discretized into N_0 segments, and the point matching and pulse basis functions [4, 6] are applied to the integral equations, we will obtain the matrix equation as

$$\begin{bmatrix} \bar{\bar{\mathbf{A}}} & \bar{\bar{\mathbf{B}}} & \bar{\bar{\mathbf{0}}} & \bar{\bar{\mathbf{0}}} & \bar{\bar{\mathbf{0}}} \\ \bar{\bar{\mathbf{C}}} & -\rho_{10}\bar{\bar{\mathbf{D}}} & \bar{\bar{\mathbf{E}}} & \bar{\bar{\mathbf{F}}} & \bar{\bar{\mathbf{0}}} \\ \bar{\bar{\mathbf{G}}} & -\rho_{10}\bar{\bar{\mathbf{H}}} & \bar{\bar{\mathbf{I}}} & \bar{\bar{\mathbf{J}}} & \bar{\bar{\mathbf{0}}} \\ \bar{\bar{\mathbf{0}}} & \bar{\bar{\mathbf{0}}} & \bar{\bar{\mathbf{K}}} & -\rho_{21}\bar{\bar{\mathbf{L}}} & \bar{\bar{\mathbf{M}}} \\ \bar{\bar{\mathbf{0}}} & \bar{\bar{\mathbf{0}}} & \bar{\bar{\mathbf{N}}} & -\rho_{21}\bar{\bar{\mathbf{P}}} & \bar{\bar{\mathbf{Q}}} \end{bmatrix}_{(N_0+4N) \times (N_0+4N)} \begin{bmatrix} \bar{\bar{\mathbf{V}}}_1 \\ \bar{\bar{\mathbf{V}}}_2 \\ \bar{\bar{\mathbf{V}}}_3 \\ \bar{\bar{\mathbf{V}}}_4 \\ \bar{\bar{\mathbf{V}}}_5 \end{bmatrix}_{(N_0+4N) \times 1} = \begin{bmatrix} \bar{\bar{\psi}}^{\text{inc}} \\ \bar{\bar{\mathbf{0}}} \\ \bar{\bar{\mathbf{0}}} \\ \bar{\bar{\mathbf{0}}} \\ \bar{\bar{\mathbf{0}}} \end{bmatrix}_{(N_0+4N) \times 1} \quad (3)$$

and, $V_1(x) = \psi_0(\mathbf{r})$ ($\mathbf{r} \in S_+$), $V_2(x) = \frac{\partial\psi_0(\mathbf{r})}{\partial n_+}$ ($\mathbf{r} \in S_+$), $V_3(x) = \psi_1(\mathbf{r})$ ($\mathbf{r} \in S_-$), $V_4(x) = \frac{\partial\psi_1(\mathbf{r})}{\partial n_-}$ ($\mathbf{r} \in S_-$), $V_5(x) = \psi_2(\mathbf{r})$ ($\mathbf{r} \in S_0$). Using the bi-conjugate gradient method to solve the matrix equation, one can obtain the unknowns $\bar{\bar{\mathbf{V}}}_1$, $\bar{\bar{\mathbf{V}}}_2$, $\bar{\bar{\mathbf{V}}}_3$, $\bar{\bar{\mathbf{V}}}_4$ and $\bar{\bar{\mathbf{V}}}_5$. The scattered field in the space Ω_0 and the definition of the bistatic scattering coefficient σ are the same as single-layered rough surface [6].

The matrixes $\bar{\bar{\mathbf{A}}}$, $\bar{\bar{\mathbf{B}}}$, $\bar{\bar{\mathbf{C}}}$, $\bar{\bar{\mathbf{D}}}$, $\bar{\bar{\mathbf{E}}}$, $\bar{\bar{\mathbf{F}}}$, $\bar{\bar{\mathbf{G}}}$, $\bar{\bar{\mathbf{H}}}$, and $\bar{\bar{\mathbf{I}}}$, $\bar{\bar{\mathbf{J}}}$, $\bar{\bar{\mathbf{K}}}$, $\bar{\bar{\mathbf{L}}}$ are provided in the appendix of Ref. [4]. The matrix $\bar{\bar{\mathbf{Q}}}$, which depends only on the target can be expressed as

$$Q_{st} = \begin{cases} -\Delta l_t \frac{ik_2}{4} (\hat{\mathbf{n}}_0 \cdot \mathbf{R}_0) H_1^{(1)}(k_2|\mathbf{r}_s - \mathbf{r}_t|) & s \neq t \\ \frac{1}{2} - \frac{Z_0''(x_t)\Delta x}{4\pi[1 + Z_0'^2(x_t)]} & s = t \end{cases} \quad (4)$$

here, $\mathbf{r}_s, \mathbf{r}_t$ are the field point and source point on the target, respectively, $Z_0(x)$ is the height function of target, $Z_0'(x)$ and $Z_0''(x)$ are the first- and second-order differential of $Z_0(x)$, Δx_t is the interval of the target, $\Delta l_t = \Delta x_t \sqrt{1 + [Z_0'(x_t)]^2}$, $\hat{\mathbf{n}}_0 = \frac{-Z_0'(x_t)\hat{\mathbf{x}} + \hat{\mathbf{z}}}{\sqrt{1 + [Z_0'(x_t)]^2}}$, $\mathbf{R}_0 = \frac{\mathbf{r}_s - \mathbf{r}_t}{|\mathbf{r}_s - \mathbf{r}_t|}$.

The matrixes $\bar{\bar{\mathbf{M}}}$, $\bar{\bar{\mathbf{N}}}$, $\bar{\bar{\mathbf{P}}}$, which indicate the interactions between the target and the lower rough

surface, can be exhibited as

$$M_{pt} = -\Delta l_t \frac{ik_2}{4} (\hat{\mathbf{n}}_0 \cdot \mathbf{R}_1) H_1^{(1)}(k_2 |\mathbf{r}_p - \mathbf{r}_t|) \quad (5a)$$

$$N_{sq} = \Delta l_q \frac{ik_2}{4} (\hat{\mathbf{n}}_- \cdot \mathbf{R}_2) H_1^{(1)}(k_2 |\mathbf{r}_s - \mathbf{r}_q|) \quad (5b)$$

$$P_{sq} = \Delta l_q \frac{i}{4} H_0^{(1)}(k_2 |\mathbf{r}_s - \mathbf{r}_q|) \quad (5c)$$

$\mathbf{r}_p, \mathbf{r}_q$ are the field points and source point located on the lower rough surface, $f'_-(x)$ and $f''_-(x)$ are the first- and second-order differential of $f_-(x)$, $\mathbf{R}_1 = \frac{\mathbf{r}_p - \mathbf{r}_t}{|\mathbf{r}_p - \mathbf{r}_t|}$, $\mathbf{R}_2 = \frac{\mathbf{r}_s - \mathbf{r}_q}{|\mathbf{r}_s - \mathbf{r}_q|}$, $\Delta l_q = \Delta x \sqrt{1 + [f'_-(x_q)]^2}$, $\hat{\mathbf{n}}_- = \frac{-f'_-(x_q)\hat{\mathbf{x}} + \hat{\mathbf{z}}}{\sqrt{1 + [f'_-(x_q)]^2}}$.

3. NUMERICAL RESULTS AND DISCUSSIONS

In Fig. 2, the bistatic scattering from a cylinder located below two-layered rough surfaces with different simulated length ($L = 51.2\lambda$, $L = 102.4\lambda$ and $L = 204.8\lambda$) are shown. The results show that the scattering pattern of $L = 102.4\lambda$ is in good agreement with that of $L = 204.8\lambda$.

In the following numerical results, the simulated length of the two-layered Gaussian rough surfaces is $L = 102.4\lambda$ with interval $\Delta x = \lambda/10$, $\lambda = 0.3$. The cylinder is located at $x_p = 0$ with segments $N_0 = 100$. The incident angle is $\theta_i = 45^\circ$ with $g = L/4$. All the numerical results are computed by averaging 50 Monte Carlo realizations. The average height between the two-layered rough surfaces is $d = 5\lambda$.

The discrepancies of the bistatic scattering from single-layered and two-layered rough surfaces, and a cylinder located below a two-layered rough surfaces are depicted in Fig. 3. The relative permittivity of the upper and lower medium are $\varepsilon_1 = 2.5$ and $\varepsilon_2 = 9.8$, and $\varepsilon = 2.5$ for single-layered model. From Fig. 3, one can notice that the bistatic scattering coefficient of the two-layered rough surfaces is larger than that of the single-layered rough surface, and is slightly smaller than that of a cylinder below two-layered rough surfaces.

The influences of the cylinder parameters, the radius r , the altitude z_p on the bistatic scattering are illustrated in Fig. 4. It is readily seen that σ increases with the increasing of the radius, and decreases as the altitude increasing. That is because the composite scattered field is related to the composite area. The larger composite area, the larger is σ .

The properties of surface parameters, the rms height δ , the correlation length l on the bistatic scattering are depicted in Fig. 5. There is an obvious peak in the specular direction for smaller rms height, and the enhanced backscattering is evident for larger rms height in Fig. 5(a). We attribute this behavior to the fact that σ depends on the roughness of rough surface. The smaller rms height, the flatter is the rough surface, leading to the obvious peak in the specular direction. From Fig. 5(b), one can find that σ increases with the decreasing of the correlation length far

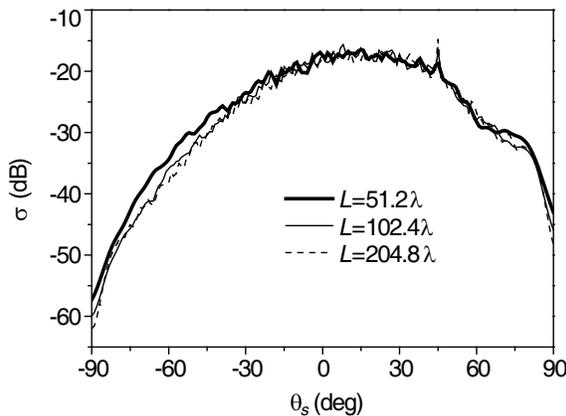


Figure 2: The bistatic scattering from cylinder below two-layered rough surface with different simulated length.

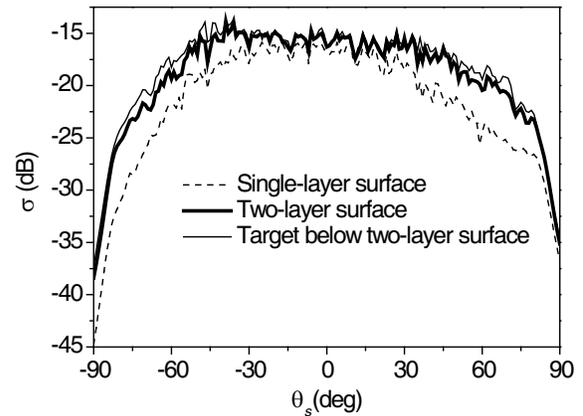


Figure 3: The bistatic scattering from single-layered and two-layered rough surfaces with and without a target.

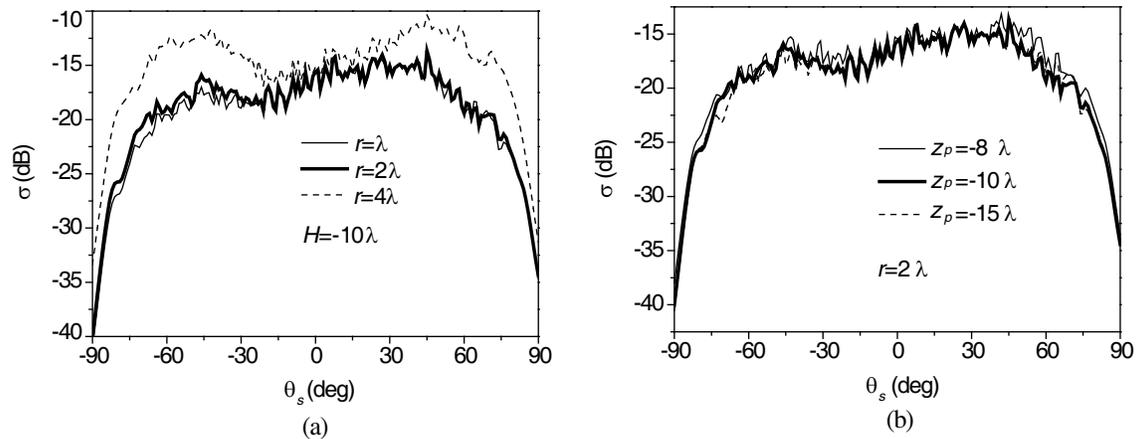


Figure 4: The bistatic scattering from cylinder below two-layered rough surface with different radius or altitude.

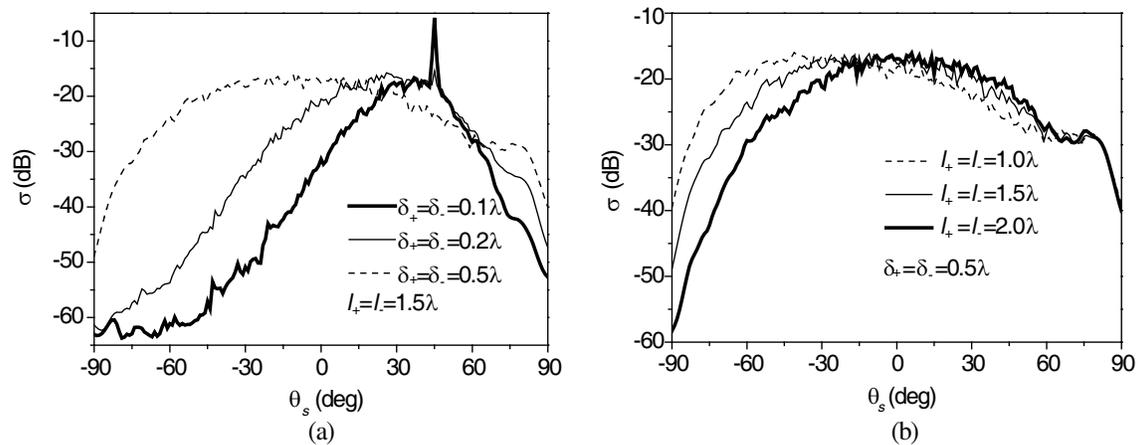


Figure 5: The bistatic scattering from cylinder below two-layered rough surface with different rms height or correlation length.

from the specular direction. The smaller correlation length l , the height fluctuation of Gaussian rough surface becomes more heavily, which leads to a more incoherent scattering contribution to the bistatic scattering in the non-specular direction.

4. CONCLUSIONS

In this paper, the EM scattering from PEC target located below the two-layered dielectric rough surfaces for vertical incidence is investigated through the MOM. The influences of the radius and altitude of cylinder, the rms height and correlation length of Gaussian rough surface on the bistatic scattering are exhibited. The properties of EM scattering of this composite models under the horizontal incidence will be studied in the next step.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China (Grant No. 60971067) and by the Specialized Research Fund for the Doctoral Program of Higher Education, China (Grant No. 20070701010). The authors would like to thank the reviewers for their constructive suggestions.

REFERENCES

1. Guo, L. X., A. Q. Wang, and J. Ma, "Study on EM scattering from 2-D target above 1-D large scale rough surface with low grazing incidence by parallel MOM based on PC clusters," *Progress In Electromagnetics Research*, PIER 89, 149–166, 2009.
2. Lawrence, D. E. and K. Sarabandi, "Electromagnetic scattering from a dielectric cylinder buried beneath a slightly rough surface," *IEEE Trans. Antennas Propag.*, Vol. 50, No. 10, 1368–1376, 2002.

3. Kuo, C. H. and M. Moghaddam, “Electromagnetic scattering from a buried cylinder in layered media with rough interfaces,” *IEEE Trans. Antennas and Propag.*, Vol. 54, No. 8, 2392–2401, 2006.
4. Thorsos, E. I., “The validity of the kirchhoff approximation for rough surface scattering using a gaussian roughness spectrum,” *Journal of Acoustical Society of American*, Vol. 83, No. 1, 78–92, 1988.
5. Déchamps, N., N. De Beaucoudrey, C. Bourlier, and S. Toutain, “Fast numerical method for electromagnetic scattering by rough layered interfaces: Propagation-inside-layer expansion method,” *J. Opt. Soc. Am. A*, Vol. 23, No. 2, 359–369, 2006.
6. Tsang, L., J. A. Kong, K. H. Ding, and C. O. Ao, *Scattering of Electromagnetic Waves: Numerical Simulations*, John Wiley & Sons, 2001.

Design and Development of a Ground-based Microwave Radiometer System

Yu Zhang^{1,2}, Jieying He^{1,2}, and Shengwei Zhang¹

¹Center for Space Science and Applied Research, Chinese Academy of Sciences, Beijing 100190, China

²Graduate University of Chinese Academy of Science, Beijing 100190, China

Abstract— This paper presents the design and development of a ground-based microwave radiometer system (MWR) for the offshore oil platform. It will be used to provide long time series of geophysical parameters such as the sea-surface temperature, the near-surface wind speed and the sea ice type. The ground-based MWR is a five-frequency, dual-polarized microwave radiometer which is a total power type microwave radiometer based on a heterodyne receiver. It operates at 6.8 GHz, 10.7 GHz, 18.7 GHz, 23.8 GHz and 37 GHz, respectively.

The scanning mode adopts circular scan in a vertical plane. During each period, two-point calibration is performed to calibrate the receiver gain and noise. The antenna beam directs different angles and the blackbody used in calibration to obtain measurement data and calibration data.

The ground-based MWR consists of three units: antenna and receiver unit, electronic unit and power supply unit. Antenna and receiver unit collects emission from the atmosphere. It is constructed with several RF components, which include voltage-controlled oscillator, high power amplifier, directional coupler, RF switch, band-pass filter, isolators, detectors and mixers. The received signal is down-converted by the double side band mixer to Intermediate Frequency, and then the IF signals are down-converted to Low Frequency by the detector and integrator. The electronic unit digitizes the LF signal, controls the scanning mechanism and measures physical temperature of the hot target for calibration, and takes charge of communication with a remote computer. The power supply unit performs DC/DC conversion, distributes the DC lines to various sub assemblies and switches.

In paper, the basic principle and technical specifications of microwave radiometer will be briefly introduced, and the block diagram and detail design parameters of the proposed system will be described. The system followed radiometer electronics design and achieved exacting stability and high accuracy requirements.

1. INTRODUCTION

In the past few decades, microwave remote sensing has become an important tool for monitoring the atmosphere and the earth surface. Microwave radiometers as an important part of microwave remote sensing, can be used to retrieve the profile of the atmospheric and sea-surface temperature, the near-surface wind speed, water vapor density, the liquid-water content of clouds and rain and the sea ice type. It can also be used to monitor the snow-water content and the spatial distribution of soil-moisture content, which are important factors for hydrology, agriculture and meteorology.

The main purpose of this ground-based MWR is for ocean research and atmosphere research and it enables high absolutely accuracy. It includes five frequencies, 6.8 GHz, 10.7 GHz, 18.7 GHz, 23.8 GHz and 37 GHz, and each frequency band is monitored by vertical polarized and horizontal polarized wave. The ground-based MWR scans the land surface by mechanically rotating the antenna and acquires radiance data of the land surface. It allows observation of everyday changes of earth's environment and it is expected to contribute to the study of the earth's environmental system.

The basic principle of microwave remote sensing by MWR and the technical specifications of the ground-based MWR are described, the block diagram, general system design and detail design are also presented in paper.

2. SYSTEM CONTENT AND TECHNICAL SPECIFICATION

As the moisture in the atmosphere could attenuate the microwave radiation from the earth surface, the brightness temperature retrieved from the observation data represents the information of atmospheric temperature and humidity etc. The principle of atmospheric microwave remote sensing is to measure atmospheric molecular rot-vib absorption spectrum which is broadened by atmospheric pressure effect, and therefore complete the retrieval of temperature and humidity profiles. The

ground-based MWR operates at 6.8 GHz, 10.7 GHz, 18.7 GHz, 23.8 GHz and 37 GHz to retrieve the profile of the atmospheric and sea-surface temperature, the near-surface wind speed, water vapor density and so on.

2.1. General System Design

The ground-based WMR is a total power type microwave radiometer based on a heterodyne receiver. The system consists of three units: antenna and receiver unit, electronic unit and power supply unit. The functional block diagram of the ground-based MWR with five receivers is shown in Fig. 1.

The system performs two-point calibration periodically to calibrate the receiver gain and noise by means of noise injection. And the input of the receiver switches automatically between the antenna unit and the calibration unit. The noise source is a microwave absorber inside the MWR and the physical temperature of which could be real-time measured accurately by temperature transducer. The noise source couples with the matched microwave load in the cases of power on and off to make the calibration unit obtain different brightness temperature. So the accuracy of the ground-based MWR is dependent upon the accuracy with the quality of the noise source and other components in the noise injection circuitry.

The received signal is focused upon the antenna feed, and then it is down-converted by the double side band mixer to Intermediate Frequency signal. It is sent to the electronic unit after processing by the amplifiers, low pass filters and directors. The five receivers are installed on the side of one cylinder which could rotate automatically; the observed angle of the ground-based MWR could be changed by changing the rotation angle of the cylinder.

2.2. Technical Specification

The technical specifications are introduced below.

Operating frequency:	6.8 GHz, 10.7 GHz, 18.7 GHz, 23.8 GHz, 37 GHz
Receiver type:	Total power receiver
Polarization mode:	<i>V, H</i>
Bandwidth:	100 MHz ~ 500 MHz
Integration time:	200 ms, 500 ms, 1000 ms (Optional)
Brightness temperature range:	90 K ~ 400 K
Brightness temperature resolution:	0.2 K
Brightness temperature accuracy:	0.5 K

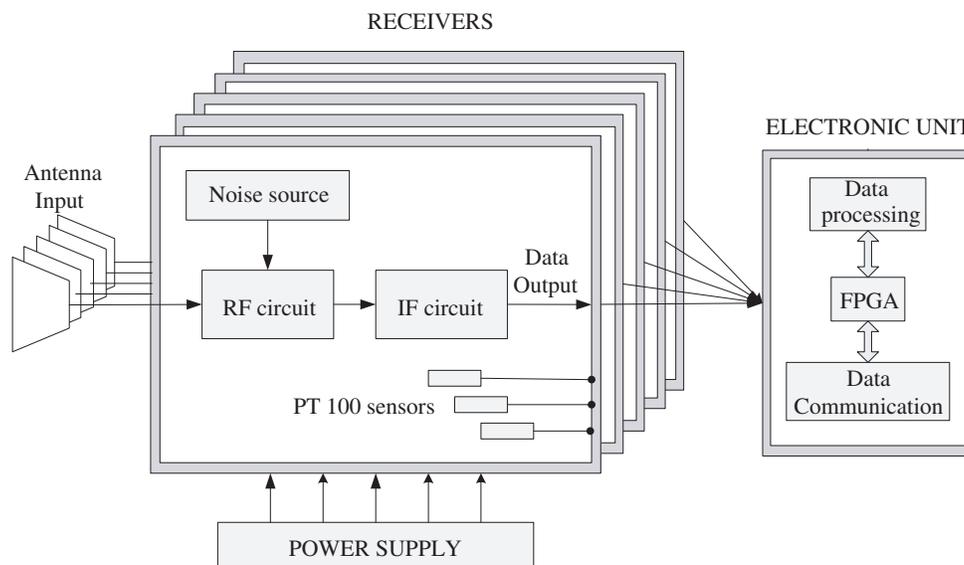


Figure 1: Functional block diagram of the proposed system.

3. DETAILED SYSTEM DESIGN

The ground-based MWR consists of three units: antenna and receiver unit, electronic unit and power supply unit. The receivers could be grouped into two separate units, the low-frequency unit and the high-frequency unit, both of which collect data from up to ten radiometer data channels. They are only identified by their frequency and channel names. The electronic unit is also able to measure up to 25 temperatures of the PT100 sensors to regulate the physical temperature inside the receivers.

3.1. Antenna and Receiver Unit

The construction of the antenna and receiver unit is illustrated in Fig. 2. Five conical corrugated horns are used in the antenna unit as feeds, they can offer a nearly circularly symmetric radiation pattern in a quite wide bandwidth and very low cross-polarization. Considering the actual measurement of sea-surface, the antenna apertures are sealed by dielectric lens to shorten the length of the antennas and avoid the corrosion of salt spray.

The receivers of the proposed radiometer system are used in pairs: H and V polarization. Hence, it is practical to construct them as dual receiver units: a pairs of receivers for one frequency. The orthogonal mode transducer (OMT), input calibration switches, couplers and noise sources are also presented in the design.

The low-frequency receivers consist of three receivers operating at 6.8, 10.7, 18.7 GHz; and the high-frequency receivers operating at 23.8 and 37 GHz adopted waveguide technology, all of which are super-heterodyne receivers with separate horizontal and vertical channels, giving a total of ten channels. The received signal is down-converted by the single side band mixer to Intermediate Frequency, and then the IF signals are sent to the detector. The analog signal needs to be integrated by the integrator at a certain time interval. The maximum signal output level from the integrator is in the order of scores of mill volts, and a sensitive low frequency amplifier is used before the signal can be properly digitized.

The performance of the receivers are strongly dependent on their physical temperature, in order to achieve the desired radiometric sensitivity, the measurement of the inside temperatures are very important for the calibration. Each receiver has five temperature sensors placed at different locations, and the measurement system is implemented using platinum resistive thermometer Pt100 connected to five high-accuracy instrumentation amplifiers, the analog signals output from the Instrumentation amplifiers and the low frequency amplifiers are all sent to the electronic unit.

3.2. Electronic Unit

The system controlling and the real time data processing are performed by this unit. The construction of the electronic unit is illustrated in Fig. 3.

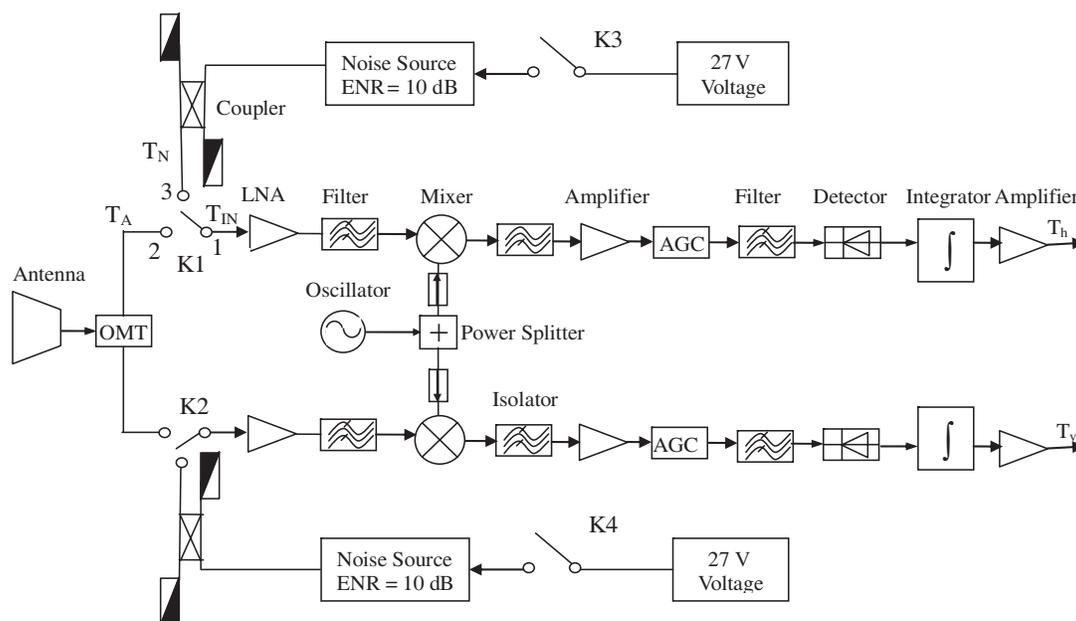


Figure 2: The block diagram of antenna and receiver unit.

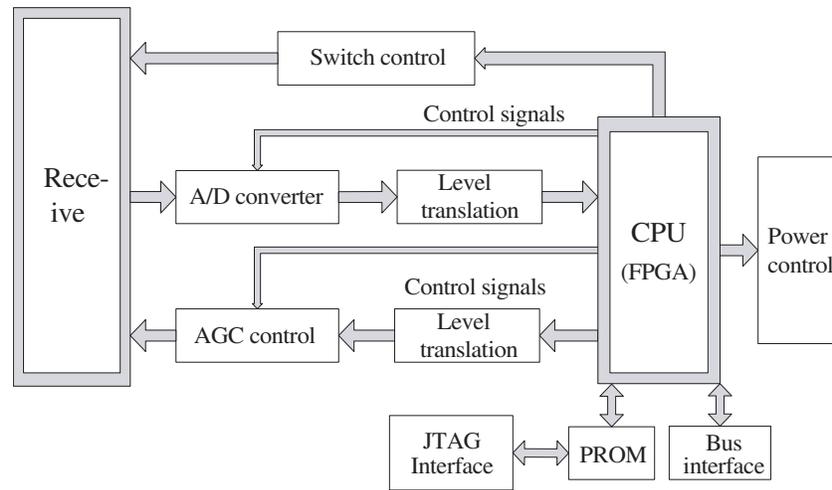


Figure 3: The block diagram of electronic unit.

The analog signal output from the receiver unit is digitized by AD converter, and then sent to Xilinx Virtex-IV FPGA. The electronic unit will also feed back the AGC voltage to the receiver unit by analyzing the received data. At the same time, the electronic unit contains the scanning mechanism control and the power control of the electronic unit itself and the radiometer receivers, which are implemented by six 24 volts DC power input relays.

Interfacing the radiometer to the remote computer is accomplished by the protocol of RS485 interface. Ten-channel observation data and twenty-five-channel temperature data will be transmitted to the remote computer after each period.

3.3. Power Supply Unit

The power supply is composed of DC/DC converter of the secondary power supply and the power distributor, which will distribute the DC lines to various sub assemblies and switches. The power supply includes +5 V and +12 V for receivers, +12 V and -12 V for the Detectors and Integrators, +24 V for the relays and +27 V for the noise sources.

4. CONCLUSION

The design of a dual-polarized, multi-frequency ground-based microwave radiometer system is presented, and the practical performance is well accordance with the technical requirement. This radiometer system can be used as a tool for continuously monitoring the atmosphere and the earth surface, and plays an important role in the observation of earth environment.

ACKNOWLEDGMENT

This project is funded by Chinese National Programs for High Technology and Development (863 plans), No: 2008AA09A403.

REFERENCES

1. Skou, N. and D. Le Vine, *Microwave Radiometer Systems Design and Analysis*, 2nd Edition, Artech House, Boston, 2006.
2. Pihlflyckt, J., "Control and measurement system for multi-channel microwave radiometer," Helsinki University of Technology, Russia, February 12, 2007, <http://lib.tkk.fi/Dipl/2007/urn007760.pdf>.
3. Jiang, J.-S., Z.-F. Zheng, H.-G. Liu, X.-Z. Wang, Y.-H. Zhang, and J. Huang, "China advanced microwave remote sensor," *Space Technol.*, Vol. 18, No. 1/2, 27–63, 1998.
4. Chan, Y. K., C. Y. Ang, V. C. Koo, and C. S. Gan, "Design and development of a FMCW ground based imaging radar system," *PIERS Online*, Vol. 5, No. 3, 265–268, 2009.
5. Ware, R., D. Cimini, P. Herzegh, F. Marzano, J. Vivekanandan, and E. Westwater, "Ground-based microwave radiometer measurements during precipitation," *8th Specialist Meeting on Microwave Radiometry*, Rome, Italy, February 2004.

6. Westwater, E. R. and Y. Han, “A multi-frequency scanning microwave radiometer for temperature profiling,” *Ninth ARM Science Team Meeting Proceedings*, San Antonio, Texas, March 1999.
7. Fonte, A., D. Zito, and F. Alimenti, “CMOS microwave radiometer: Experiments on down-conversion and direct detections,” *15th IEEE International Conference on Electronics, Circuits and Systems*, 1273–1276, Malta, August 2008.
8. Chan, P. W., “Performance and application of a multi-wavelength, ground-based microwave radiometer in intense convective weather,” *Meteorologische Zeitschrift*, Vol. 18, No. 3, 253–265, 2009.
9. Filatov, A., N. Karataeva, and A. Loshchilov, “A microwave radiometer for measuring signals with a low effective noise temperature,” *Instruments and Experimental Techniques*, Vol. 49, No. 6, 817–823, 2006.
10. Turk, J., J. Vivekanandan, F. S. Marzano, R. E. Hood, R. W. Spencer, and F. J. LaFontaine, “Active and passive microwave remote sensing of precipitating storms during CaPE. Part I: Advanced microwave precipitation radiometer and polarimetric radar measurements and models,” *Meteorology and Atmospheric Physics*, Vol. 54, No. 1–4, 1994.
11. Brown, S., C. Ruf, S. Keihm, and A. Kitiyakara, “Jason microwave radiometer performance and on-orbit calibration,” *Marine Geodesy*, Vol. 27, No. 1–2, 199–220, 2004.

Relationship between Lightning Discharges and Rapid Changes in Cross Polarization Discrimination of the Ka-band Satellite Radio Signal

Yasuyuki Maekawa

Osaka Electro-Communication University, Japan

Abstract— Rapid changes in cross-polarization discrimination (XPD) of the satellite radio wave signal (19.45 GHz) are observed at 1 sec interval in the thunderstorm events. About one third of the rapid changes are found to coincide with the cloud-to-ground lightning strokes which occurred on the south side of our earth station at the distance up to 20 km. The distribution of the lightning location indicates that the center of thunder clouds primarily exists to the westward of the radio wave propagation path. At the moment of the rapid changes, more than half the observed data indicate the decrease in XPD. Also, more than two thirds of them show the changing direction of cross-polar phases toward -90 deg, which means the decrease in canting angles of ice crystals possibly causing the depolarization changes. These features seem to be enhanced in the middle stage of each lightning event. The decrease in XPD may be related to cancellation effects of depolarization due to the difference in canting angles between the ice crystals near the lightning and those in other places far from the lightning. Thus, the measurements of XPD and cross-polar phase are shown to be important to reveal the electrification process of thunder clouds and to predict lightning discharges.

1. INTRODUCTION

In satellite communication links, degradation of cross-polarization discrimination (XPD) of the radio wave is caused by raindrops and ice crystals along the propagation paths. In thunderstorm events, moreover, rapid changes of XPD and cross-polar phase relative to co-polar phase are frequently observed at time intervals of less than 1 sec [1, 2]. These peculiar phenomena are considered to be related to the change of electric field due to lightning discharges as well as the aerodynamic forces of convective air flows [3]. Up to now, however, detailed mechanisms of the rapid changes of XPD and cross-polar phase are not understood very well.

In this study, statistical distributions of the changes of XPD and cross-polar phase are examined using the Ka-band beacon signal observations of Japan's domestic communication satellites [4] which have been conducted for the past 17 years at 1 sec (partly 0.1 sec) interval in Osaka Electro-Communication University. The thunderstorm events are observed more than 70 times from 1990 to 2006, and more than 2500 examples of rapid XPD changes are obtained during these years in our university. The occurrence time and the amplitude of these changes are then compared with the lightning location and time provided by the Kansai Electric Power Company in Osaka area. Also, the decrease and increase in XPD and cross-polar phase are discussed in relation to the alignment of ice crystals due to electrostatic force and convective air flow in thunder clouds.

2. EXAMPLE OF THUNDERSTORM EVENTS

First, an example of the thunderstorm event observed on August 12, 1990 is presented. Fig. 1 shows the variations of (a) XPD and (b) cross-polar phase relative to co-polar phase, respectively. On this day, no rainfall was recorded at the station, and the attenuation did not exceed 2 dB during the thunderstorm event. However, large depolarization with the XPD values down to nearly 25 dB, and rapid changes in (a) XPD and (b) cross-polar phase can be detected almost every minute during 15:10–15:40 LT. Dashed lines indicate the time of cloud-to-ground (CG) lightning strokes near the station recorded by the Lightning Location and Protection System (LLPS) of the Kansai Electric Power Company. Several samples of the rapid changes are thus found to actually coincide with the moments of lightning strokes near the station, while the other considerable samples are seen to occur without CG lightning strokes. It should be noted that these characteristic changes in depolarization are primarily caused by ice crystals above the rain height, since no appreciable rain attenuation occurred in this event. Hence, these rapid changes without CG lightning discharges rather seem to be related to inter-cloud (IC) lightning discharges that occur in higher part of the thunder clouds.

Next, Fig. 2 depicts the distribution of mean canting angles of ice crystals inferred from the observed cross-polar phase just before (thin line) and after (thick line) the rapid changes, respectively, for the thunderstorm events observed from 1990 to 1998. Here, positive angles are defined as clockwise rotation seen from the satellite. Thus, the ice crystals are basically canting around 10–30 deg in clockwise even after the rapid changes as shown by the thick line. Moreover, the thin line suggests that the mean canting angles are increased up to 30–50 deg in the “same” clockwise direction before the changes, and then they return to around 10–30 deg after the changes associated with CG or possible IC lightning discharges.

Figure 3 illustrates horizontal distribution of lightning locations which affected the rapid changes of XPD and cross-polar phase observed in thunderstorm events during 1990–1998. The location of our earth station is at the origin of the diagram. The size of symbols is classified according to the amplitude of the phase changes such as more than 20 or 40 deg. It is seen from Fig. 3 that the data points primarily exist on the south side of the station at the distance up to about 20 km. Note that the large data points with conspicuous phase changes are mainly found in southward and westward directions, and that they are centered around a south-west point (–4 km, –6 km) from the station with a radius of approximately 4–12 km. The distribution of these data points seem to indicate the area where the cloud base producing CG strokes stays in each thunderstorm, so the center of the clouds primarily exists to the westward of the propagation path shown by a thick dashed line.

As the elevation angle of the satellite is about 50 deg, the beacon signal radio wave may pass through comparatively higher part of the thunderclouds mostly composed of ice crystals. So, the location of the cloud center being westward from the radio wave path in Fig. 3 may explain the afore-mentioned average clockwise rotation of ice crystals shown in Fig. 2 above the rain height as follows. In an “aerodynamic-gravitational” point of view, the small clockwise inclination of ice crystals that still exists after the rapid changes suggests the effects of westerly (eastward) wind shear on ice crystals possibly due to the convection of thunder clouds together with the jet stream that is high-altitude westward winds common in mid latitudes. In an “electrostatic” point of view,

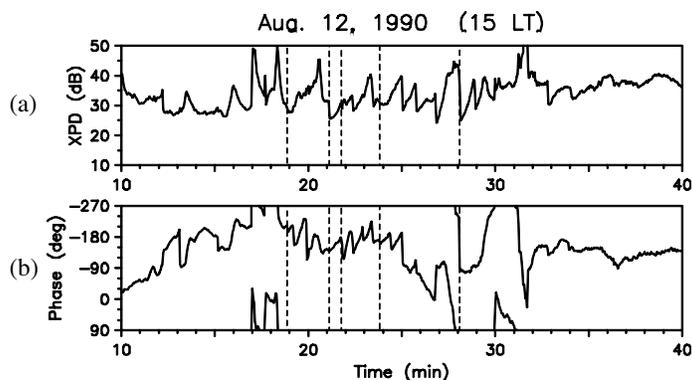


Figure 1: Example of (a) XPD and (b) cross-polar phase observed in the thunderstorm event on August 12, 1990. Dashed lines indicate the time of cloud-to-ground (CG) lightning strokes recorded near the station.

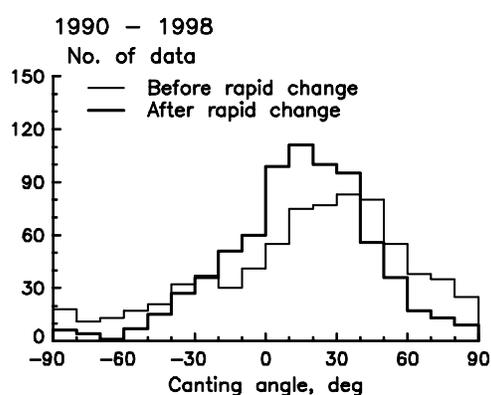


Figure 2: Distribution of mean canting angle of ice crystals before and after the rapid changes.

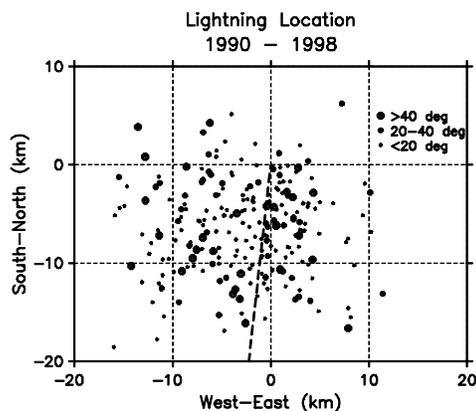


Figure 3: Horizontal distribution of lightning location around the earth station.

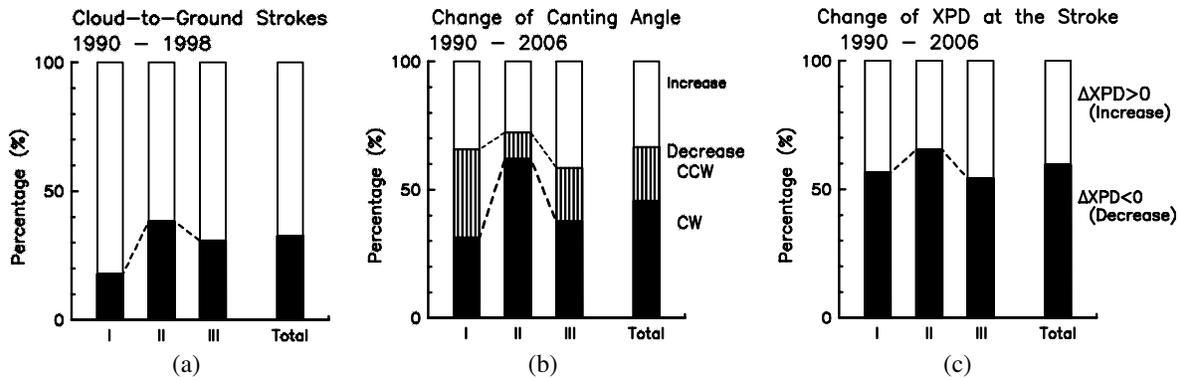


Figure 4: Distribution of (a) coincident rates with CG lightning strokes, (b) clockwise and counter-clockwise rotations for the canting angles of ice crystals, and (c) increase and decrease in the rapid changes of XPD. The results are shown for the total period and three periods of each thunderstorm event equally divided by every one third of the duration time of each event.

on the other hand, the large clockwise inclination of the canting angles just before the rapid changes is certainly caused by a strong vertical electric field, which is expected in the height of ice crystals at the moment of lightning discharges [5].

3. LONG-TERM STATISTICS

In this section, the statistical nature of XPD and cross-polar phase variations is further investigated for the entire observational period from 1990 to 2006. Fig. 4 shows the distribution of (a) coincident rates with cloud-to-ground (CG) lightning strokes, (b) clockwise (CW) and counter-clockwise (CCW) rotations for the canting angles of ice crystals, and (c) increase and decrease in XPD at the moment of lightning strokes, respectively. All the statistics are calculated for the entire observational period and three periods of each thunderstorm event which are equally divided by every one third of the duration time of each event and approximately correspond to developing, mature, and dissipating stages of thunder clouds [5]. As for the comparison with CG lightning strokes shown in Fig. 4(a), however, the statistics are limited up to 1998, since the LLPS data are only available for the duration from 1990 to 1998 near the station.

It is seen from Fig. 4(a) that about one third of the rapid changes of the depolarization coincide with CG lightning discharges near the station. Their percentages are found to increase in the middle stage of thunderstorms, when CG lightning strokes frequently occur. On the other hand, the percentages extremely decrease at the first period when thunder clouds are still developing, since inter-cloud (IC) discharges are, in general, active in this early stage [5].

As was previously shown in Fig. 2, the canting angles of ice crystals inferred from the cross-polar phase variations tend to rotate largely in the clockwise direction seen from the satellite before the lightning discharge and then return to nearly horizontal directions after the discharge. Fig. 4(b) statistically indicates that more than two thirds of the canting angles decrease at the moment of lightning discharges. Also, more than two thirds of the decreased canting angles originally rotate in the clockwise direction, and this tendency is more conspicuous in the middle stage of thunderstorms when their convective air motion is usually most active. Thus, these features of the canting angles are possibly caused by the afore-mentioned combination of both “aerodynamic-gravitational” and “electrostatic” effects.

Finally, Fig. 4(c) shows that there are two different cases in which the XPD values increase (depolarization decreases) and decrease (depolarization increases) at the moment of rapid changes, with decreasing cases slightly larger than decreasing cases. These results are distinct from those of other previous XPD observations using INTELSAT in low elevation angles, which primarily show the decrease in XPD at the moment of rapid changes in thunderstorms [6]. In the present observations, moreover, the XPD values tend to increase when the average XPD is comparatively high such as at the beginning or end of thunderstorms, while they tend to decrease when the average XPD is generally low such as in the middle of thunderstorms, as shown by Fig. 4(c).

Note that if the size of thunder clouds is small such as at the beginning or end of thunderstorms, XPD values will be only affected by the alignment of ice crystals near the discharging point before lightning. In this case, when the electrostatic force is released after a lightning discharge, the XPD

values for circular polarizations can be increased by random orientation of ice crystals. On the other hand, if the size of thunder clouds is large such as in the middle of thunderstorms, XPD values may be rather decreased before lightning by the cancellation effects between ice crystals in other places far from the lightning and those near the lightning. It should be noted that the mean canting angle of ice crystals is extremely increased near the lightning location before a discharge, while it does not seem to be increased so much in other places far from the lightning since “electrostatic” forces can not reach so far. Thus, the difference of their mean canting angles seems to yield depolarization cancellation as the charge is being accumulated. After the lightning occurs, the XPD value then decreases rapidly because their mean canting angles approach and become nearly horizontal direction.

4. CONCLUSIONS

The characteristic rapid changes in XPD and cross-polar phase are presented for the thunderstorm events obtained from the Ka-band beacon signal observations of the Japan’s domestic communication satellites for the past 17 years. About one third of the rapid changes are found to coincide with the cloud-to-ground (CG) lightning strokes, which primarily occurred on the south side of our earth station at the distance up to 20 km. The distribution of the lightning location indicates that the center of thunder clouds that produced these lightning strokes primarily exists to the westward of the radio wave propagation path. At the moment of rapid changes, two thirds of the observed cross-polar phases tend to move toward -90 deg, which means the decrease in canting angles of ice crystals inferred from the cross-polar phase. Also, more than half the observed XPD data indicate the decrease in XPD, which is equivalent to the increase in depolarization at their rapid changes. These features seem to be enhanced in the middle stage of thunderstorm events, when the thunder clouds become most developed and their convective motion and electrification are very active.

The average canting angles of ice crystals are considered to become very large in the clockwise direction seen from the satellite before lightning discharges, while they seem to return to nearly horizontal directions. These features of the canting angles are found to be well explained by the combination of “aerodynamic-gravitational” forces due to the eastward wind shear at the high altitude and strong vertical “electrostatic” forces before the lightning discharges. On the other hand, the present observations indicate a number of cases in which the XPD increases at the moment of rapid changes, in contrast to the past low elevation-angle observations using INTELSAT that showed the decrease in XPD for almost all cases. In our high-elevation observations, however, the increase in XPD after the lightning seems to be directly related to the anisotropic effects of ice crystal alignment caused by the decrease in the standard deviation of ice crystal’s canting angles due to electrostatic forces just before the lightning. To contrast, the decrease in XPD after the lightning may be related to cancellation effects of depolarization before the lightning due to the difference in canting angles between the ice crystals near the lightning and those in other places far from the lightning.

ACKNOWLEDGMENT

The LLPS data were supplied by the courtesy of the Kansai Electric Power Company.

REFERENCES

1. Haworth, D. P., P. A. Watson, and N. J. McEwan, “Model for the effect of electric fields on satellite-earth microwave radio propagation,” *Electron. Lett.*, Vol. 13, 562–564, 1977.
2. Cox, D. C. and H. W. Arnold, “Observation of rapid changes in the orientation and degree of alignment of ice particles along an earth-space radio propagation path,” *J. Geophys. Res.*, Vol. 84, 5003–5010, 1979.
3. Bostian, C. W. and J. E. Allnutt, “Ice depolarization on satellite-earth microwave radio path,” *Proc. IEE*, Vol. 126, 951–960, 1979.
4. Karasawa, Y. and Y. Maekawa, “Ka-band earth-space propagation research in Japan,” *Proc. IEEE*, Vol. 85, No. 6, 821–842, 1997.
5. Takahashi, T., “Thunderstorm electrification — A numerical study,” *J. Atmos. Sci.*, Vol. 41, 2541–2558, 1984.
6. Furuta, O., H. Yuki, and M. Yamada, “Abrupt changes in cross polarizations observed during thunder,” *IEEE Trans. Antennas and Propagat.*, Vol. 33, 625–632, 1985.

Calibration and Temperature Retrieval of Improved Ground-based Atmospheric Microwave Sounder

Jie Ying He^{1,2}, Yu Zhang^{1,2}, and Sheng Wei Zhang¹

¹Center for Space Science and Applied Research, Chinese Academy of Sciences, Beijing 100190, China

²Graduate University of Chinese Academy of Sciences, Beijing 100190, China

Abstract— Calibration and retrieval are two important and critical techniques for ground-based atmospheric microwave sounder. Adopting proper calibration methods will ensure a high sounding resolution. The paper uses two calibrated methods like LN2 (22–31 GHz and 51–59 GHz) calibration and sky tipping (22–31 GHz) to realize periodic calibration and using noise injection and in-built blackbody to realize internal calibration. Assuming the atmospheric temperature profile is known to 2 K (state-of-the-art with present remote sensors) and instrument noise is 1–2 K brightness temperature (achievable with current technology). The paper extracts several clear-air datasets randomly in the available radiosonde datasets. It derives atmospheric absorption spectrum from MPM93 model, and simulates brightness temperature using radiative transfer equation. In order to estimate atmospheric temperature profiles from the radiometer data, the algorithm of back-propagation neural network has been used. The retrievals yield good results in the temperature profiles from the surface to nearly 400 hPa. So, the prototype of improved ground-based atmospheric microwave sounding meets the requirements and using proper algorithm the retrieval results can be used in many fields.

1. INTRODUCTION

Microwave radiometer is one of primary instruments for earth and ocean observation because of its ability of working on all-weather and all-day. Ground-based atmospheric microwave sounder has the advantages of higher resolution at the bottom of troposphere, lower maintenance cost, easy to match and operate and so on.

2. DESCRIPTION OF SOUNDING PRINCIPLE

The principle of atmospheric temperature and humidity sounding is to measure the atmospheric (oxygen and water vapor) molecular rotating absorbing spectrum and its wings; all of them are pressure broadened and therefore complete the retrieval of temperature and humidity profiles. Where, the oxygen molecular absorbing spectrum at about 50–60 GHz can be used for retrieving temperature profiles and water molecular absorbing spectrum at about 20–30 GHz for retrieving humidity profiles.

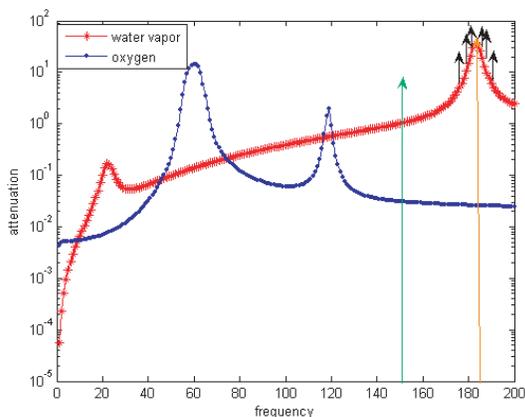


Figure 1: Opacity of microwave transmission.

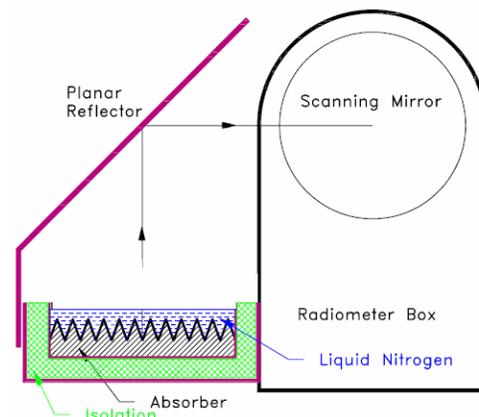


Figure 2: External cold load attached to the radiometer box.

Figure 1 shows that the high opacity of atmosphere microwave transmission [1] (oxygen and water vapor). In the band of 0–200 GHz, there are two lines of oxygen (detecting atmospheric temperature profiles) and water vapor (detecting atmospheric humidity profiles), respectively. Among

them, the energy spectrum enhances gradually as the frequency becomes higher and higher. The oxygen lines are at 50–60 GHz (usually measures space-borne vertical distribution of atmospheric oxygen) and 118 GHz. The water vapor lines are at 22.235 GHz and 183.31 GHz [2].

Having the same principle and referring to MP3000A [3], the ground-based atmospheric microwave radiometer operates at k-band (22~31 GHz) and V-band (51~59 GHz), and retrieves the atmospheric temperature profiles and humidity profiles and obtains parameters like liquid water vapor, flux and delay and so on.

3. CALIBRATION

3.1. LN2 calibration

One of absolute calibration standards is the liquid nitrogen cooled target [4] that is attached externally to the radiometer box. This standard — together with the internal ambient load — is used for the absolute calibration procedure.

The boiling temperature of the liquid nitrogen and thus the physical temperature of the cold load depend on the barometric pressure p . The radiometer's pressure sensor is read during absolute calibration to determine the corrected boiling temperature according to the equation:

$$T_c = T_0 - 0.00825 * (1013.25 - p) \quad (1)$$

where, $T_0 = 77.25$ K is the boiling temperature at 1013.25 hPa. The calibration error due to microwave reflections at the LN/air interface is automatically corrected by the calibration software (embedded PC). It is recommended to wrap a plastic foil around the load + radiometer (wind protection) during absolute calibration to avoid the formation of condensed water above the liquid surface (caused by wind etc.).

3.2. Sky Tipping Calibration

Sky tipping [5] (tip curve) is a calibration procedure suitable for those frequencies where the earth's atmosphere opacity is low (i.e., high transparency) which means that the observed sky brightness temperature is influenced by the cosmic background radiation temperature of 2.7 K. High opacity channels like all temperature profiler channels > 53 GHz are saturated in the atmosphere and must be calibrated by other methods. Sky tipping assumes a homogeneous, stratified atmosphere without clouds or variations in the water vapor distribution. If these requirements are fulfilled, the following method is applicable: The radiometer scans the atmosphere from zenith to around 20° in elevation and records the corresponding detector readings for each frequency. The path length for a given elevation angle α is $1/\sin(\alpha)$ times the zenith path length (often referred to as "air mass"), thus the corresponding optical thickness should also be multiplied by this factor (if the atmosphere is stratified).

The optical thickness is related to the brightness temperature by the equation:

$$\tau(\infty) = -\ln\left(\frac{T_{mr} - T_i}{T_{mr} - T_{B0}}\right) \sec(\theta) \quad (2)$$

T_{mr} is a mean atmospheric temperature in the direction θ , T_{B0} is the 2.7 K background radiation temperature and T_i is the brightness temperature of frequency channel i .

$$T_{mr} = \frac{\int_0^{\infty} T(z) e^{-\tau(z)} \sigma_a dz}{1 - e^{-\tau(\infty)}} \quad (3)$$

T_{mr} is a function of frequency and is usually derived from radiosonde data. A sufficiently accurate method is to relate T_{mr} with a quadratic equation of the surface temperature measured directly by the radiometer.

The sky tipping calibrates the system noise temperature and the gain factor for each frequency without using a liquid nitrogen cooled target. The disadvantage of this method is that the assumption of a stratified atmosphere is often questionable even under clear sky conditions due to invisible inhomogeneous water vapor distributions.

So, LN2 method is mainly be used in v-band operating at 51–59 GHz, and TIP is mainly used in k-band operating at 22–30 GHz.

3.3. Internal Calibration

External calibration is to realize validation and calibrating periodically (every one year or half). Internal calibration [6] is to realize quasi-real-time calibration, and calibrate gain and noise of receiver, and avoid the affects of noise fluctuations and noise drifts.

The internal calibration period depends on short-term stability and can be 10–20 minutes with temperature-controller in relative stable environment. Internal calibration unit consists of noise injection block and in-built calibration blackbody, and both provide stable reference. Noise injection block consists of noise source (noise diode) which provides noise to be calibrated, switch which turn on and off the noise signal and directional coupler which is used for noise injection. We can used microwave switch or supply power to turn on and off the noise signal. Directional coupler is used for feeding into noise signal which temperature is 100–200 K.

Internal calibration blackbody provides standard brightness temperature (\sim ambient temperature). To ensure the stability, there are many pt-resistances to measure the temperature gradients. In order to reduce the gradient, it is optimal to use foam material which has performance of insulation as blackbody calibration layer and DC mini-fan to drive the airflow.

Internal calibration unit has another function. It can correct nonlinear error by Noise injection method. Although ideal radiometer receiver is a linear system, nonlinear error caused by nonlinearity of detector diode and it is not negligible and mostly reaches to 1 K order.

The nonlinearity of detector can be as follows:

$$U = GP^\alpha \quad 0 < \alpha < 1 \quad (4)$$

where, U is the detector voltage, G is the detect-coefficient, α is a nonlinearity factor and P is the total noise power that is proportional to the radiometric brightness temperature T_r according to Planck radiation law. The relationship between them is:

$$U_2 = G'(T_{REC} + T_{COLD} + T_A)^\alpha \quad (5)$$

G' (system gain), α (nonlinear factor) and T_{sys} (noise temperature of receiver) are three unknown quantities. To correct nonlinearity of system, we generate four temperature points by additional noise injection of temperature T_n . So there four independent equations with four unknowns (G , α , T_{sys} and T_n) and then fit the nonlinear curve showing in Figure 2.

By injecting known noise-temperature into receiver, we can calibrate multi-points through nonlinear response. In the design, by observing low temperature and ambient temperature blackbody, there're two calibration points $T_c \sim U_1$ and $T_h \sim U_2$, and other two points $T_c + T_n \sim U_3$ and $T_h + T_n \sim U_4$ after injecting noise. Then 3 calibration parameters and noise injection and then decide the nonlinear factor α can be derived.

Noise temperature and gain of receiver change when ambient temperature changes. Internal calibration is a method which calibrates noise temperature and gain of receiver periodically through two known reference targets, then obtaining a real-time calibration equation to measure the real-time target brightness temperature.

4. REREIEVAL PRINCIPLE

In recent years, artificial neural network (ANN) has been used widely [7]. The structure of ANN shows in Figure 3. The layers 1, 2, and 3 represent the input layer, the hidden layer, and the output layer, respectively. The neurons of the input layer are represented by vector X_i ($i = 1, 2, 3, \dots, n$, where n is the number of the input neurons).

For the i th node in the hidden layer, this can be expressed as

$$Y_j = S \left(\sum_{i=1}^L w_{ij} x_i + b_j \right) \quad (6)$$

where, S denotes the sigmoid function,

$$S(\alpha) = \frac{1}{1 + \exp(-\alpha)} \quad (7)$$

where, w_{ij} is the weighting of the connection between the j th input neuron and the i th hidden neuron. b_j denotes the bias between calculated and measured values. Between hidden layers and output layers it uses purelin function.

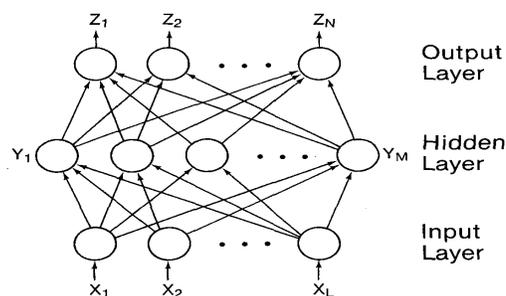


Figure 3: Schematic diagram of a three layer back-propagation neural network.

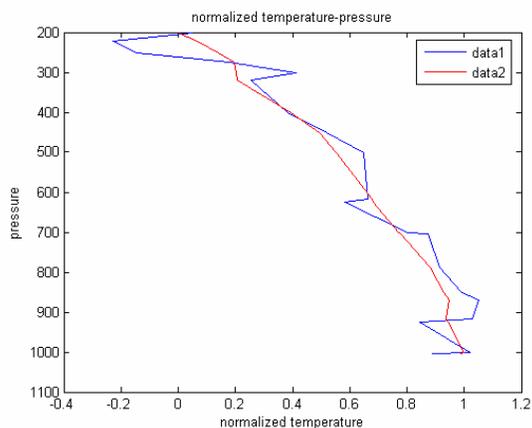


Figure 4: Normalized temperature.

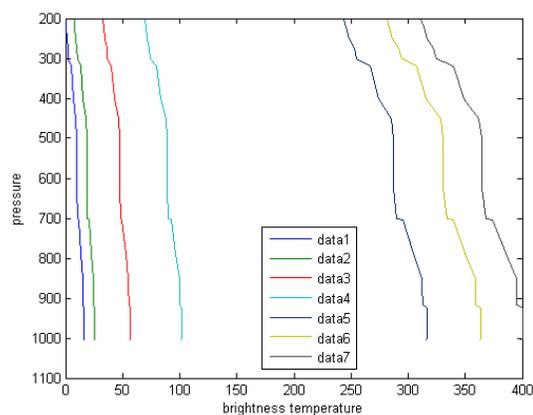


Figure 5: Brightness temperature.

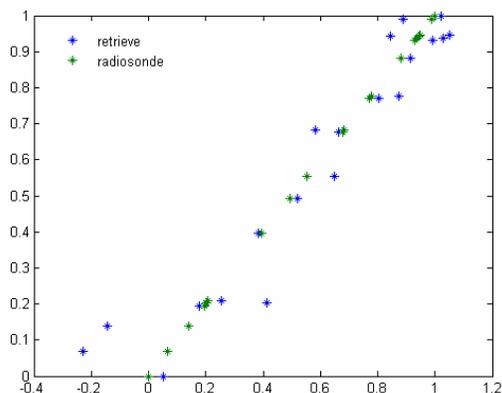


Figure 6: Normalized retrieve temperature.

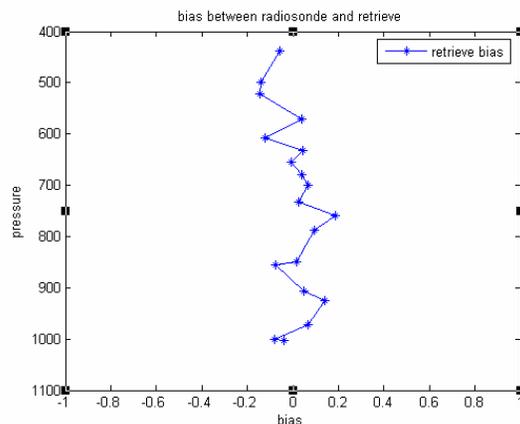


Figure 7: Retrieval bias.

5. SIMULATION RESULTS

The paper uses 225 radiosonde data. Choose 20% randomly as train sets and choose 20% randomly as tests. The input has 10 elements include 7 brightness temperatures (7 channels in v-band), troposphere temperature, troposphere pressure, troposphere height. The output includes 20 temperature values in 20 values of pressure from 1000 mba to 300 mbar or height from 0 to 10 km [8].

6. CONCLUSIONS

According to the radiative transfer equation, we retrieve the temperature from advanced prototype of ground-based microwave sounder. The neural network technique can derive the temperature weighting function by train the radiosonde datasets, and using other radiosonde datasets to test the model. Different temperature profiles of different regions can easily be retrieved using this method; also it can be used for different microwave sounders of ground-based and satellite platforms.

REFERENCES

1. Liebe, H. and D. Lyton, "MM-wave properties of the atmosphere: Laboratory studies and propagation modeling," NTIA report 87-224, 80, Oct. 1987.
2. Ulaby, F. T., R. K. Moore, and A. K. Fung, *Microwave Remote Sensing (I)*, Addison-Wesley Publishing Company, New York, 1982.
3. MP-Series Microwave Profilers, Radiometrics Corporation, http://www.radiometrics.com/MP_Specifications_7-3-07.pdf.
4. RPG's Atmospheric Remote Sensing Radiometers, <http://www.radiometer-physics.com>.
5. RPG-150-90/RPG-DP150-90 High Sensitivity LWP Radiometers, <http://www.radiometer-physics.com>.
6. Rose, T., R. Zimmermann, and R. Zimmermann, "A precision auto calibrating 7 channel radiometer for environmental research applications," *Journal of Remote Sensing Society of Japan*, 1999.
7. Shi, L., "Retrieval of atmospheric temperature profiles from AMSU-A measurement using neural network approach," *Journal of Atmospheric and Oceanic Technology*, Mar. 2001.
8. Churnside, J. H., T. A. Stermitz, and J. A. Schroeder, "Temperature profiling with neural network inversion of microwave radiometer data," *Journal of Atmospheric and Oceanic Technology*, 1993.

Investigation of GPS-measured Ionospheric Total Electron Content Variations Generated by HF-heating at Mid-latitudes

V. E. Kunitsyn¹, A. M. Padokhin¹, A. E. Vasiliev¹,
G. A. Kurbatov¹, V. L. Frolov², and G. P. Komrakov²

¹M. V. Lomonosov Moscow State University, Russia

²Radiophysical Research Institute, N. Novgorod, Russia

Abstract— The aim of this work is to present experimental results of the influence of electron density perturbations caused by HF-heating of ionospheric F2 layer on GNSS navigational signals at mid-latitudes. The experiments were carried out on the Sura heater near N. Novgorod, Russia with radiating frequency 4.3 MHz (during all series of experiments the plasma frequency of F2 layer was greater than the heating frequency) and effective power 40 MW and 80 MW. During the heating series with different time modulation of radiative power, the ionospheric penetration points for several navigational satellites and GNSS-receiver installed near the Sura facility crossed the heated area. For these satellites we investigated variations of total electron content (TEC), proportional to the differential carrier phase of navigational signals. It is shown that for square heating pulses, perturbations with main heating period appeared in the spectrum of TEC variations. The amplitudes of such perturbations were of the order of 0.07 TECU. These TEC oscillations support the presence of electron density irregularities caused by such modes of HF-heating in the heated region.

1. INTRODUCTION

Numerous investigations of the characteristics of artificial turbulence in ionospheric F2 region due to powerful O-mode HF radiation showed that near the heights of radio wave reflection, ponderomotive parametric, thermal (resonant) parametric and self-focusing instabilities develop, see articles in special issues of journals [1–7] and review articles [8–13] and references therein. This causes particularly the significant heating of plasma in this area and generation of artificial irregularities in electron density distribution with the scales from meters to kilometers. These irregularities significantly influence the properties of HF radiowaves, so one of the ways to study this phenomena is to analyze the properties (differential carrier phase) of highly stable L-band navigational radio signals, in particular GPS-signals propagating through the heated area in the ionosphere [14, 15]. In this work we present the results confirming the presence of perturbations in GPS signals caused by the modulation of HF-heating of mid-latitude ionosphere in daytime conditions, obtained on the Sura heater near N. Novgorod, Russia.

2. EXPERIMENTAL SETUP

Series of experiments were carried on the Sura heater, with coordinates $\phi = 56.15N$; $\lambda = 46.1E$, inclination of the geomagnetic field for the site 71° , for the period of low and medium geomagnetic activity ($K_p = 0-3$) in March 12–19, 2009. Plasma frequencies f_{oF2} of the ionospheric F2 layer during this campaign were, as a rule, not greater than 5 MHz even at noon. So the use of minimal for the Sura facility heating wave frequency 4.3 MHz determined the possible time for the experiments: From 9:00 till 16:00–17:00 local time. Hence the experiments were carried out only in sunlit ionosphere in the morning, daytime, and early evening when the efficiency of artificial ionospheric turbulence excitation is significantly lower than in night conditions, when for example experiments discussed in [14] were carried out.

Except two cases, when only two transmitters of the Sura heater were used in the experiments and the effective radiative power (ERP) was $ERP = 40$ MW, the measurements were carried out with three synchronously working transmitters of the Sura heater, which provided $ERP = 80$ MW. The level of absorption of the heating wave energy was not controlled in the experiments but was expected to be of the order 3–5 dB in daytime conditions, which caused the decrease not less than 2–3 times in transported to the F2-region ERP.

To increase the efficiency of artificial ionospheric turbulence excitation we used the magnetic zenith effect [11, 16] and tilted the beam 12° southward in the plane of geomagnetic meridian.

In such conditions at the height 200 km, corresponding to the height of heating wave reflection, the HF-heated spot corresponding to 0.1ERP level was ~ 70 km in diameter centered at $\phi =$

55.79N; $\lambda = 46.03E$. The measurements were conducted when plasma frequency of F2 layer f_{OF2} , controlled by the ionosonde located near Sura heater, was greater than 4.3 MHz and the ionospheric penetration points (IPP) for some satellites of GPS constellation and GNSS receiver installed near Sura facility crossed the heated area.

We used different time modulations of ERP: Square pulses of maximum ERP with the widths 30 sec, 3 minutes, 5 minutes with 30 sec, 3 and 5 minutes gaps between the pulses correspondingly in O-mode, and the regime with 10-minutes-wide square pulses with 5 minutes gaps between pulses. Note that IPPs for high-orbital navigational satellites, such as GPS or GLONASS, move rather slowly and remain within the heated area for 35–50 minutes, which allows to investigate the time variations in the electron density field in the heated region using the signals of these systems.

3. DATA PROCESSING AND RESULTS

The experimental data were collected using the dual-frequency GPS/GLONASS receiver JAVAD LEXON installed near the Sura heater. For analysis we used raw data of observed phases L_1 and L_2 on both GPS-system frequencies $f_1 = 1575.42$ MHz and $f_2 = 1227.6$ MHz with 10 Hz sampling.

These measurements were converted to slant Total Electron Content (slTEC), proportional to the differential carrier phase as follows:

$$\text{slTEC} = \left(\frac{L_1}{f_1} - \frac{L_2}{f_2} \right) \frac{f_1^2 f_2^2}{f_1^2 - f_2^2} \frac{c}{K} + \text{const} \quad (1)$$

where $K = 40.308 \frac{m^3}{s^2}$, c — speed of sound in vacuum, and const is an additive constant in TEC estimates due to the unknown initial phase.

In order to study the effects of the HF-heating with the modulation of ERP discussed in Section 2, which are expected to be rather small, we removed the trend from slTEC time series using adjacent averaging with 20-minutes time window. Thus we obtain TEC variations during HF-heating, which can be analyzed using, for example, the wavelet transform to estimate the local spectrum of TEC variations in order to compare it with the time modulation of ERP.

For example, on March 16, 2009 the heating experiment was carried out from 6:15 UT till 8:26 UT in undisturbed geomagnetic conditions ($K_p = 1$ during the experiment). The heating wave excited rather intensive F_{spread} on ionograms. Its reflection height was ~ 205 km. At the first part of the experiment till 07:36 UT, we used 30 sec-wide square pulses of ERP = 40 MW with 30 sec gaps between pulses. Since 07:36 UT the modulation was the same but with ERP = 80 MW, and than at 7:51 UT we switched to the regime of 5 minutes square pulses with 5 minutes gap between pulses. During this heating experiment starting from 7:23 UT till 8:11 UT, the IPP for GPS prn22 satellite crossed the heated area. In Figure 1, one can see the trajectory of the IPP as it crossed the heated area.

TEC variation for the period 7:35 UT–7:50 UT is shown in Figure 2. Variations in TEC with period $T = 1$ min are clearly seen in this figure and also revealed in the local spectrum of TEC variations presented in Figure 3. This period in TEC variations corresponds to time modulation of ERP of Sura heater. Note that $T = 1$ min in TEC variations was clearly registered only since 07:36 UT, when the Sura heater started to work at ERP = 80 MW, and disappeared at 07:51 UT, when the Sura heater was switched to another modulation.

Maximum TEC variations of the order of 0.04 TECU were observed at 07:40 UT practically in the magnetic zenith region. This clearly evidences the artificial nature of the observed TEC variations and also points out that in those conditions present during the experiment, the ERP = 40 MW with 30 sec modulation was insufficient to generate irregularities in electron density field detectable by GPS measurements. Nevertheless, if we estimate the electron density perturbation at the reflection height of the heating wave using simple relation:

$$\frac{\delta N}{N} \sim \frac{\delta \text{TEC}}{\text{TEC}} \frac{H_{\text{ion}}}{l} \quad (2)$$

where N is electron density at the reflection height of heating wave, H_{ion} is the thickness of ionosphere, δTEC is TEC variation caused by HF-heating; and assuming that the thickness of the disturbed region l does not exceed 5–15 km, we obtain the perturbations about 10%, which is rather significant.

For the period March 12–19, 2009 we obtained 14 registrations of TEC variations during HF-heating with different modulations of ERP. In all of these registrations we observed the spectral

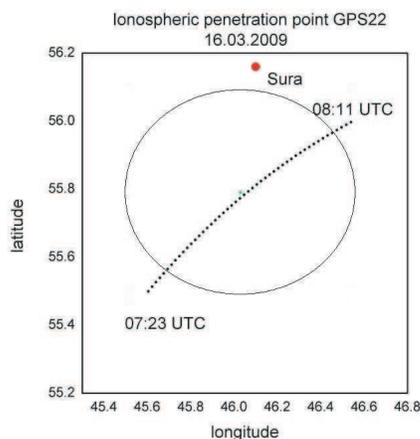


Figure 1: The trajectory of IPP for GPS prn22 and LEXON receiver located near the Sura heater (red dot), March 16, 2009. The oval on the figure outlines the contours of the heater beam limited by 0.1 ERP at height 200 km.

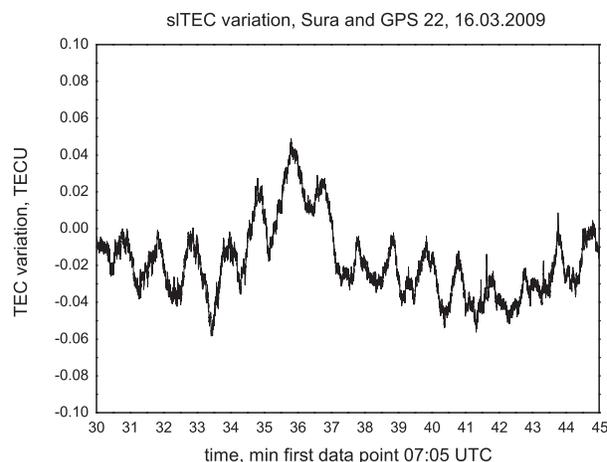


Figure 2: Variation of slant TEC for GPS prn22 when IPP crossed the heated area as shown in Figure 1.

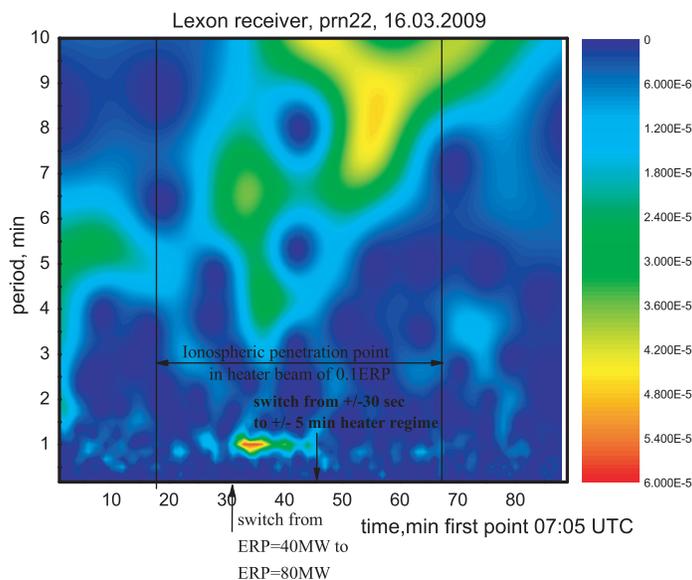


Figure 3: Local spectrum of TEC variations for GPS prn22 when IPP crossed the heated area as shown in Figure 1. Borders of heated area as well as the moments when modulation and ERP were changed, are indicated by arrows.

components in TEC variations corresponding to ERP modulation of HF-heating, in particularly one registration with 1 min period, two registrations with 6 min period, and six registrations with 10 min period. For the modes with 10 min heating and 5 min gap between the pulses we observed 15 min and 7.5 min periods in TEC variations. The observed amplitudes of TEC variations during the heating were of the order of 0.04–0.07 TECU, which corresponds to 10–15% perturbations of electron density.

4. CONCLUSION

Our experiments carried out in the mid-latitude dayside ionosphere show that in GPS-measured TEC variations the spectral components appear corresponding to the modulation of HF-heating, when the ionospheric penetration point of the tracked GPS-satellite crosses the heated area. Amplitudes of such variations were about 0.04–0.07 TECU, which corresponds to 10–15% perturbations in electron density, and did not show significant dependence on time of observation within the period from 9:00 till 16:00–17:00 local time. Sometimes together with main heating modulation frequency

we observed multiple frequencies in TEC variations with smaller amplitudes of variations and its maximum shifted in time compared to the main heating modulation frequency.

ACKNOWLEDGMENT

The work was supported by the Russian Foundation for Basic Research (RFBR), grants No. 08-05-00676, 07-05-01120, 08-02-00171.

REFERENCES

1. *J. Atmos. Terr. Phys.*, Vol. 44, No. 12, 1982.
2. *J. Atmos. Terr. Phys.*, Vol. 47, No. 12, 1985.
3. *J. Atmos. Terr. Phys.*, Vol. 59, No. 18, 1997.
4. *Radiophysics and Quantum Electronics*, Vol. 37, No. 5, 1994.
5. *Radiophysics and Quantum Electronics*, Vol. 42, No. 7–8, 1999.
6. *Radiophysics and Quantum Electronics*, Vol. 48, No. 9, 2005.
7. *Radiophysics and Quantum Electronics*, Vol. 51, No. 11, 2008.
8. Erukhimov, L. M., S. A. Metelev, E. N. Myasnikov, N. A. Mityakov, and V. L. Frolov, “Artificial ionospheric turbulence (review),” *Radiophysics and Quantum Electronics*, Vol. 30, No. 2, 156–171, 1987.
9. Leyser, T. B., “Stimulated electromagnetic emissions by high-frequency electromagnetic pumping of the ionospheric plasma,” *Space Sci. Rev.*, Vol. 98, No. 3–4, 223–328, 2001.
10. Stubbe, P. and T. Hagfors, “The earth’s ionosphere: A wall-less plasma laboratory,” *Surveys in Geophysics*, Vol. 18, 57, 1997.
11. Gurevich, A. V., “Nonlinear effects in the ionosphere,” *Phys. Usp.*, Vol. 50, 1091, 2007.
12. Frolov, V. L., N. V. Bakhmet’eva, V. V. Belikovich, G. G. Vertogradov, V. G. Vertogradov, G. P. Komrakov, D. S. Kotik, N. A. Mityakov, S. V. Polyakov, V. O. Rapoport, E. N. Sergeev, E. D. Tereshchenko, A. V. Tolmacheva, V. P. Uryadov, and B. Z. Khudukon, “Modification of the earth’s ionosphere by high-power high-frequency radio waves,” *Phys. Usp.*, Vol. 50, 315, 2007.
13. Belikovich, V. V., S. M. Grach, A. N. Karashtin, D. S. Kotik, and Y. V. Tokarev, “The “Sura” facility: Study of the atmosphere and space (a review),” *Radiophysics and Quantum Electronics*, Vol. 50, No. 7, 497–526, 2007.
14. Tereshchenko, E. D., A. N. Milichenko, V. L. Frolov, and R. Y. Yurik, “Observations of the magnetic-zenith effect using GPS/GLONASS satellite signals,” *Radiophysics and Quantum Electronics*, Vol. 51, No. 1, 842–846, 2008.
15. Milikh G., A. Gurevich, K. Zybin, and J. Secan, “Perturbations of GPS signals by the ionospheric irregularities generated due to HF-heating at triple of electron gyrofrequency,” *Geophys. Res. Lett.*, Vol. 35, L22102, 2008.
16. Tereshchenko, E. D., B. Z. Khudukon, A. V. Gurevich, K. P. Zybin, V. L. Frolov, E. N. Myasnikov, N. V. Muravieva, and H. C. Carlson, “Radio tomography and scintillation studies of ionospheric electron density modification caused by a powerful HF-wave and magnetic zenith effect at mid-latitudes,” *Physics Letters A*, Vol. 325, 381–388, 2004.

Fluctuation of Electromagnetic Field Parameters Propagating in Magnetized Plasma with Random Variation of Electron Density and Magnetic Field

G. V. Jandieri¹, A. Ishimaru³, V. G. Jandieri⁴, I. B. Shirokov¹,
Yu. B. Gimpilevich¹, A. G. Khantadze⁵, and N. N. Zhukova²

¹Physics Department, Georgian Technical University, Tbilisi, Georgia

²Institute of Cybernetics, Tbilisi, Georgia

³Department of Electrical Engineering, University of Washington, USA

⁴Department of Electrical and Computer Engineering, Kumamoto University, Japan

⁵Physics Department, Tbilisi State University, Tbilisi, Georgia

Abstract— Statistical characteristics of scattered and reflected electromagnetic waves by turbulent magnetized plasma layer with electron density and magnetic field fluctuations at arbitrary orientation of an external magnetic field are considered using the perturbation method. The obtained results are valid for near and far zones. Analytical expressions for scattered electric field components are obtained using the boundary conditions. Correlation functions of these statistical characteristics are analyzed analytically and numerically for the ordinary waves at different orientation of external magnetic field using satellite and remote sensing data.

1. INTRODUCTION

Statistical characteristics of the angular power spectrum (broadening and displacement of its maximum) of scattered electromagnetic waves by turbulent anisotropic collision magnetized ionospheric plasma was investigated in the complex geometrical optics approximation [1] taking into account both electron density and magnetic field fluctuations. However, it should be noted that the magnetic field is fluctuating only on size, but not in direction. This approximation is applicable only for small-length paths, so called in the near zone. Scintillation level of scattered radiation was calculated for *F*-region irregularities of the ionosphere using both power-law and anisotropic Gaussian [2, 3] correlation functions of electron density fluctuations. Ionospheric effects due to irregularities on the performance of space-based radar system and synthetic aperture radar system were considered in [4–6]. Statistical characteristics (correlation function of the phase and relative amplitude fluctuations) of scattered electromagnetic radiation propagating through magnetized plasma layer are considered in this paper using the perturbation method. Analytical expression for scattered electric field along the external magnetic field is obtained satisfying boundary conditions and taking into account both fluctuations of electron density and magnetic field fluctuations. Numerical calculations of the correlation function of scattered radiation towards the direction of an external magnetic field were carried out for anisotropic Gaussian correlation function.

2. FORMULATION OF THE PROBLEM

Let's ions frequency satisfies the condition $\omega \gg \Omega_i = eH_0/Mc$, where Ω_i is the ion gyrofrequency, H_0 is external magnetic field, M is mass of an ion. If $\omega \gg \nu_{eff}$, where ν_{eff} electron effective collision frequency with neutral particles and ions, conduction current should be neglected and total current in a magnetized plasma is equal to the displacement current $\mathbf{j} = -eN\mathbf{w}$, \mathbf{w} is the velocity of electrons. The set of equations describing propagation of electromagnetic waves in the collisionless magnetized plasma has the following form:

$$\text{grad div } \mathbf{E} - \Delta \mathbf{E} - k_0^2 \mathbf{E} = -\frac{\tilde{v} k_0^2}{1 - \tilde{u}} \left\{ \mathbf{E} - \frac{ie}{mc\omega} [\mathbf{E} \cdot \mathbf{H}_0] - \left(\frac{e}{mc\omega} \right)^2 (\mathbf{E} \cdot \mathbf{H}_0) \mathbf{H}_0 \right\}, \quad (1)$$

where $\tilde{u} = (eH_0/mc\omega)^2$, $\tilde{v} = \omega_p^2/\omega^2$, $\omega_p = (4\pi N e^2/m)^{1/2}$ plasma frequency of electrons, $k_0 = \omega/c$.

Electric induction vector $\mathbf{D} = \mathbf{E} - 4\pi ieN_0\mathbf{w}/\omega$ is equal to:

$$\mathbf{D} = \left(1 - \frac{\tilde{v}}{1 - \tilde{u}} \right) \mathbf{E} + i \frac{\tilde{v}}{1 - \tilde{u}} \left[\mathbf{E} \cdot \frac{\boldsymbol{\Omega}_H}{\omega} \right] + \frac{\tilde{v}}{1 - \tilde{u}} \left(\mathbf{E} \cdot \frac{\boldsymbol{\Omega}_H}{\omega} \right) \frac{\boldsymbol{\Omega}_H}{\omega}, \quad (2)$$

where $\Omega_H = eH_0/mc$ gyrofrequency of electrons. Each quantity in Equation (1) we present as sum of the mean and fluctuating terms $\mathbf{E} = \langle \mathbf{E} \rangle + \mathbf{e}$, $\mathbf{H} = \langle \mathbf{H} \rangle + \mathbf{h}$, $\mathbf{H}_0 = \langle \mathbf{H}_0 \rangle + \mathbf{h}_0$, $N = \langle N \rangle + n$. The mean values of both electron density and magnetic field not depend on the spatial coordinates. If fluctuations are small, we can use the perturbation method and for fluctuating electric field we obtain the differential equation:

$$\left(\frac{\partial^2}{\partial x_i \partial x_j} - \Delta \delta_{ij} - k_0^2 \varepsilon_{ij} \right) e_j = j_i, \quad (3)$$

where

$$\mathbf{j} = \frac{k_0^2 v}{1-u} \left\{ i\sqrt{u} [\langle \mathbf{E} \rangle \mathbf{h}'_0] + u(\langle \mathbf{E} \rangle \boldsymbol{\tau}) \mathbf{h}'_0 + u(\langle \mathbf{E} \rangle \mathbf{h}'_0) \boldsymbol{\tau} \right\} \\ - \frac{k_0^2 v}{1-u} \cdot \left[n' + \frac{2u}{(1-u)^2} (\boldsymbol{\tau} \cdot \mathbf{h}'_0) \right] \left\{ \langle \mathbf{E} \rangle - i\sqrt{u} [\langle \mathbf{E} \rangle \boldsymbol{\tau}] - u (\langle \mathbf{E} \rangle \boldsymbol{\tau}) \boldsymbol{\tau} \right\}$$

is the current density, $v = \omega_{p0}^2/\omega^2$, $\omega_{p0}^2 = 4\pi e^2 \langle N \rangle / m\omega^2$, $u = (e\langle H_0 \rangle / mc\omega)^2$, $n' = n/\langle N \rangle$, $\mathbf{h}'_0 = \mathbf{h}_0/\langle \mathbf{H}_0 \rangle$, $\boldsymbol{\tau} = \langle \mathbf{H}_0 \rangle / H_0$. Let unperturbed EM field propagates along the z axis and vector $\boldsymbol{\tau}$ ($\mathbf{k} \parallel z$, $\langle \mathbf{H}_0 \rangle \in yoz$) lies in the $yo z$ plane (principal plane). Components of the second rank tensor ε_{ij} for collisionless plasma are [7]: $\varepsilon_{xx} = 1 - v \cdot (1-u)^{-1}$, $\varepsilon_{yy} = 1 - v(1-u \sin^2 \alpha) \cdot (1-u)^{-1}$, $\varepsilon_{zz} = 1 - v(1-u \cos^2 \alpha) \cdot (1-u)^{-1}$, $\varepsilon_{xy} = -\varepsilon_{yx} = iv\sqrt{u} \cos \alpha \cdot (1-u)^{-1}$, $\varepsilon_{yz} = \varepsilon_{zy} = uv \sin \alpha \cos \alpha \cdot (1-u)^{-1}$, $\varepsilon_{xz} = -\varepsilon_{zx} = -iv\sqrt{u} \sin \alpha \cdot (1-u)^{-1}$, where α is the angle between \mathbf{k} and \mathbf{H}_0 vectors. From Equation (3) we obtain set of differential equations for perturbed electric field components:

$$\begin{aligned} \tilde{e}_x'' + a_1 \tilde{e}_z' + b_1 \tilde{e}_x + c_1 \tilde{e}_y + d_1 \tilde{e}_z &= f_1, \\ \tilde{e}_y'' + a_2 \tilde{e}_z' + b_2 \tilde{e}_x + c_2 \tilde{e}_y + d_2 \tilde{e}_z &= f_2, \\ \tilde{e}_x' + a_3 \tilde{e}_y' + b_3 e_x + c_3 e_y + d_3 e_z &= f_3 \end{aligned} \quad (4)$$

where: $a_1 = -ik_x$, $b_1 = k_0^2 \varepsilon_{xx} - k_y^2$, $c_1 = k_0^2 \varepsilon_{xy} + k_x k_y$, $d_1 = k_0^2 \varepsilon_{xz}$, $a_2 = -ik_y$, $b_2 = k_0^2 \varepsilon_{yx} + k_x k_y$, $c_2 = k_0^2 \varepsilon_{yy} - k_x^2$, $d_2 = k_0^2 \varepsilon_{yz}$, $a_3 = k_y/k_x$, $b_3 = ik_0^2 \varepsilon_{zx}/k_x$, $c_3 = ik_0^2 \varepsilon_{zy}/k_x$, $f_1 = -k_0^2 g_x$, $f_2 = -k_0^2 g_y$, $f_3 = -ik_0^2 g_z/k_x$, $d_3 = -i(k_x^2 + k_y^2 - k_0^2 \varepsilon_{zz})/k_x$, \mathbf{g} is the 2D Fourier transformation of the current density \mathbf{j} . Boundary conditions for Equations (4) are as follows: waves propagating in negative direction should be absent at $z \geq L$ (L is a thickness of inhomogeneous plasma layer); waves propagating in positive direction should be absent at $z \leq 0$. Since all the functions inside the layer are finite functions in the interval $0 \leq z \leq L$ ($k_0 l_{N,M} \gg 1$, $k_0 L \gg 1$, $L \gg l_{N,M}$, l_N and l_M are characteristic linear scales of electron density of magnetic field fluctuations, respectively), we could use the spectral method. Applying the residue theory, particularly, for z component of scattered electric field satisfying the boundary conditions we obtain:

$$\begin{aligned} \tilde{e}_z(\mathbf{x}, z) = & -\frac{1}{\pi \varepsilon_{zz}} \left(\frac{1}{\delta_1} \left\{ B_1(\gamma_x + i\gamma_y) d'_0 \int_0^L dz'_0 [d'_1 n'_1(\mathbf{x}, z') + d'_2 h'_z(\mathbf{x}, z')] \cdot \cos[(l-z')k_0 x_1] \right. \right. \\ & \left. \left. + i [B_3 + B_4(\gamma_x^2 + \gamma_y^2)] \cdot d'_3 \int_0^L dz' [h'_x(\mathbf{x}, z') + ih'_y(\mathbf{x}, z')] \cdot \sin[(l-z')k_0 x_1] \right\} \right. \\ & \left. + \frac{1}{\delta_2} \left\{ B_2(\gamma_x + i\gamma_y) d'_0 \int_0^L dz' [d'_1 n'_1(\mathbf{x}, z') + d'_2 h'_z(\mathbf{x}, z')] \cdot \cos[(l-z')k_0 x_2] \right. \right. \\ & \left. \left. + i [B_5 + B_6(\gamma_x^2 + \gamma_y^2)] \cdot d'_3 \int_0^L dz' [h'_x(\mathbf{x}, z') + ih'_y(\mathbf{x}, z')] \cdot \sin[(l-z')k_0 x_2] \right\} \right). \end{aligned} \quad (5)$$

where: $\gamma_x = k_x/k_0$, $\gamma_y = k_y/k_0$, $d'_0 = -v\langle E_x \rangle / (1-u)$, $d'_1 = 1 + \sqrt{u}$, $d'_2 = \sqrt{u} + 2u(1 + \sqrt{u}) / (1-u)$, $\delta_1 = 8\varepsilon_{xx} \tilde{\varepsilon}_{xy} c_1$, $\delta_2 = -8\varepsilon_{xx} \tilde{\varepsilon}_{xy} c_2$, $c_1 = \varepsilon_{xx} + \tilde{\varepsilon}_{xy}$, $c_2 = \varepsilon_{xx} - \tilde{\varepsilon}_{xy}$, $x_1 = c_1 - d_1(\gamma_x^2 + \gamma_y^2)$, $x_2 = c_2 - d_2(\gamma_x^2 + \gamma_y^2)$, $d_1 = (c_1 + \varepsilon_{zz})/2\varepsilon_{zz}$, $d_2 = (c_2 + \varepsilon_{zz})/2\varepsilon_{zz}$, $\mathbf{x} = \{k_x, k_y\}$, B_j ($j = 1 \dots 6$) are complex functions of plasma parameters.

3. STATISTICAL CHARACTERISTICS OF SCATTERED RADIATION

We are investigating correlation function of scattered EM plane normal (ordinary and extraordinary) waves propagating through the plasma layer with random electron density and magnetic field fluctuations. Each component of the mean field $\langle E_j \rangle = |E_j| \exp(i q_{0j} z)$ is a slowly-varying function of z coordinate; the mean field is dumping owing to transformation of the mean field energy to the fluctuating one. We assume that the thickness of a slab is substantially smaller than the dumping length and $|E_j|$ is constant. In the perturbation theory fluctuation of the phase $\delta\varphi$ and the relative fluctuations of the amplitude δA of scattered electric field e_i are determined by the formulae: $\delta\varphi = \text{Im}(e_i/\langle E \rangle)$, $\delta A = \text{Re}(e_i/\langle E \rangle)$. Hence, correlation functions of the phase and amplitude fluctuations are connected through the correlation function of scattered field. According to [7], we have:

$$N_j^2 = \frac{q_{0j}}{k_0} = 1 - \frac{2v(1-v)}{2(1-v) - u \sin^2 \alpha \mp \sqrt{u^2 \sin^4 \alpha + 4u(1-v)^2 \cos^2 \alpha}} \quad (6)$$

N_j^2 is square of refractive index corresponding two normal waves propagating in a homogeneous magnetized plasma (without taking into account thermal motion); upper sign in the first formula is devoted to the extraordinary wave ($j = 1$), lower sign — to the ordinary wave ($j = 2$). Both waves in collisionless magnetized plasma, in general, are elliptically polarized. At the angle $\alpha = 0^\circ$ for the ordinary wave $\langle E_y \rangle = i\langle E_x \rangle$, $\langle E_z \rangle = 0$; for extraordinary wave $\langle E_y \rangle = -i\langle E_x \rangle$, $\langle E_z \rangle = 0$. In this case, $\varepsilon_{xx} = \varepsilon_{yy} = 1 - v \cdot (1 - u)^{-1}$, $\varepsilon_{xy} = -\varepsilon_{yx} = iv\sqrt{u} \cos \alpha \cdot (1 - u)^{-1} \equiv i\tilde{\varepsilon}_{xy}$, $\varepsilon_{xz} = \varepsilon_{zx} = \varepsilon_{yz} = \varepsilon_{zy} = 0$, $\varepsilon_{zz} = 1 - v$.

2D Fourier-transformation of the current density fluctuations are determined by the expressions:

$$\begin{aligned} g_x &= -\frac{v}{1-u} \langle E_y \rangle \left[(1 + \sqrt{u}) n' + \left(\sqrt{u} + 2u \frac{1 + \sqrt{u}}{1 - u} \right) h'_{0z} \right], \\ g_y &= i g_x, \\ g_z &= v \frac{1 + \sqrt{u}}{1 - u} \langle E_y \rangle (h'_{0x} + i h'_{0y}). \end{aligned}$$

This means that fluctuations of the current density in the xoy plane are caused by both electron density and magnetic field fluctuations, which are statistically independent; z component contains only magnetic field fluctuations (external magnetic field is directed along the z axis). Square of the refractive index for the ordinary and extraordinary waves in this case are equal to $N_2^2 = 1 - v(1 + \sqrt{u})^{-1}$, $N_1^2 = 1 - v(1 - \sqrt{u})^{-1}$, respectively. In the integrands of Equation (5) we keep only first terms of the expansion into a series due to the inequality $k_x, k_y \sim 2\pi/l \ll k_0$, while in the exponent — three terms, as the second and third terms can give the substantial contribution at great distances of z . For anisotropic Gaussian correlation function of electron density fluctuations [8]:

$$W_N(k_x, k_y, \rho_z) = \langle N_1^2 \rangle \frac{l_\perp \bar{l}}{4\pi} \exp\left(-\frac{k_x^2 \bar{l}^2}{4} - \frac{k_y^2 l_\perp^2}{4}\right) \exp\left(-\chi^2 \frac{\bar{l}^2}{l_\parallel^4} \rho_z^2\right) \cdot \exp\left[i(\chi^2 - 1) \frac{\bar{l}^2}{l_\parallel^2} \sin \gamma_0 \cos \gamma_0 k_x \rho_z\right] \quad (7)$$

where: $\bar{l}^2 = l_\parallel^2 / (\sin^2 \gamma_0 + \chi^2 \cos^2 \gamma_0)$, $\chi = l_\parallel / l_\perp$ is the coefficient of anisotropy, γ_0 is the angle of inclination of prolate irregularities with respect to the external magnetic field, on the lower boundary ($z = L$) z component of the correlation function of scattered electric field could be written as

$$\begin{aligned} & \left[\frac{\text{Re}\langle e_z(x + \rho_x, y + \rho_y, L) e_z(x, y, L) \rangle}{\text{Im}\langle e_z(x + \rho_x, y + \rho_y, L) e_z(x, y, L) \rangle} \right]_N = \frac{\sigma_N^2}{\sqrt{\pi}} \cdot \frac{L}{l_\parallel} \cdot \frac{v^2(1 + \sqrt{u})^2}{\chi^2 \varepsilon_{zz}^2 (1 - u)^2} \cdot \langle E_x \rangle^2 \cdot \begin{bmatrix} \cos(2\varphi) \\ \sin(2\varphi) \end{bmatrix} \\ & \cdot \left\{ \frac{B_1^2}{\delta_1^2} \left(\frac{1}{2\xi \frac{L}{l_\parallel} c_1} \cdot \begin{bmatrix} \text{Re}\mathbf{K}_1^{(1)} \\ \text{Im}\mathbf{K}_1^{(1)} \end{bmatrix} + 2 \exp\left(-\frac{\xi^2 c_1^2}{4m^2}\right) \cdot \begin{bmatrix} \text{Re}\mathbf{K}_2^{(1)} \\ \text{Im}\mathbf{K}_2^{(1)} \end{bmatrix} \right) + \frac{B_2^2}{\delta_2^2} \cdot \left(\frac{1}{2\xi \frac{L}{l_\parallel} c_2} \cdot \begin{bmatrix} \text{Re}\mathbf{K}_1^{(2)} \\ \text{Im}\mathbf{K}_1^{(2)} \end{bmatrix} \right. \right. \\ & \left. \left. + 2 \exp\left(-\frac{\xi^2 c_1^2}{4m^2}\right) \cdot \begin{bmatrix} \text{Re}\mathbf{K}_2^{(2)} \\ \text{Im}\mathbf{K}_2^{(2)} \end{bmatrix} \right) + \frac{4B_1 B_2}{\delta_1 \delta_2} \cdot \left[\exp\left(-\frac{t_1^2 \xi^2}{16m^2}\right) \cdot \begin{bmatrix} \text{Re}\mathbf{K}_3^{(1)} \\ \text{Im}\mathbf{K}_3^{(1)} \end{bmatrix} \right. \right. \\ & \left. \left. + \exp\left(-\frac{y_1^2 \xi^2}{16m^2}\right) \cdot \begin{bmatrix} \text{Re}\mathbf{K}_3^{(2)} \\ \text{Im}\mathbf{K}_3^{(2)} \end{bmatrix} \right] \right. \end{aligned} \quad (8)$$

where $\mathbf{K}_i^{(j)}$ are complex functions of plasma parameters, $\xi = k_0 l_{\parallel}$, ρ_y and ρ_x are the distances between the observation points in principle and perpendicular planes, respectively.

4. NUMERICAL CALCULATIONS

Numerical results illustrate the dependence of the correlation function of a scattered electromagnetic field towards z direction versus polar angle φ of the wave vector in the plane perpendicular to the magnetic field at moderate (Figure 1) and strong (Figure 2) external magnetic fields. Inhomogeneous area having thickness 50 km is located at a height 300 km. Plasma parameters: $u = 0.0012$, $v = 0.0133$, basic frequency of an incident radiation is $\nu = 40$ MHz. Anisotropy coefficient $\chi = 10$, angle of inclination of prolate irregularities with respect to the external magnetic field $\gamma_0 = 0^\circ$, longitudinal characteristic linear scale $l_{\parallel} \approx 12$ km, the variance of electron density fluctuation is equal to $\sigma_N^2 = 10^{-4}$. Numerical calculations show that the intensity of scattered radiation oscillates

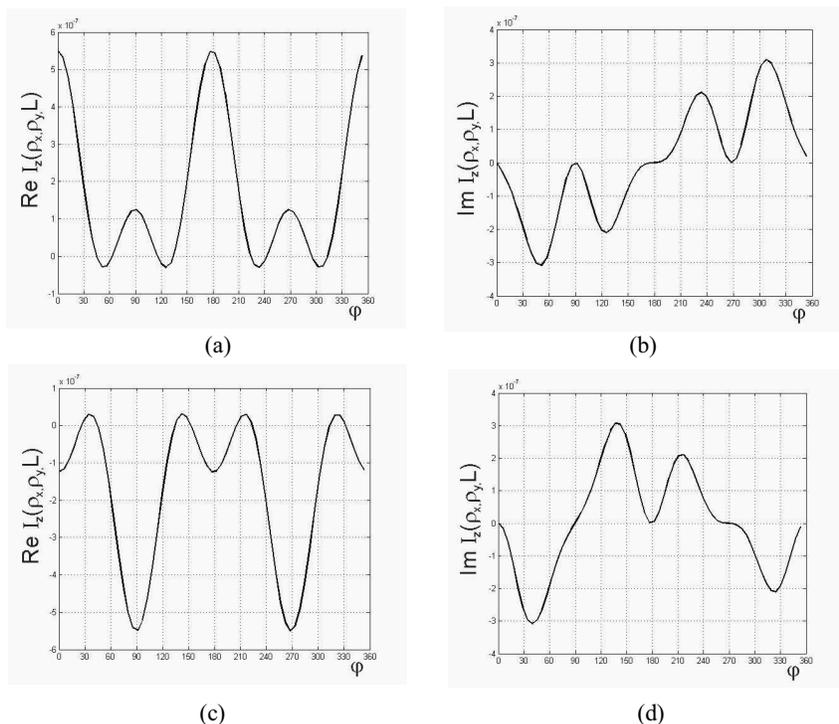


Figure 1: Illustrates the dependence of the correlation function of a scattered electromagnetic field versus polar angle φ in the direction perpendicular of an external magnetic field ($u = 0.0012$) in the perpendicular ($\rho_x = 1, \rho_y = 0$) and principle ($\rho_x = 0, \rho_y = 1$) planes, Figures 1(a)–1(d), respectively. Angle $\gamma_0 = 0$, anisotropy coefficient $\chi = 10$.

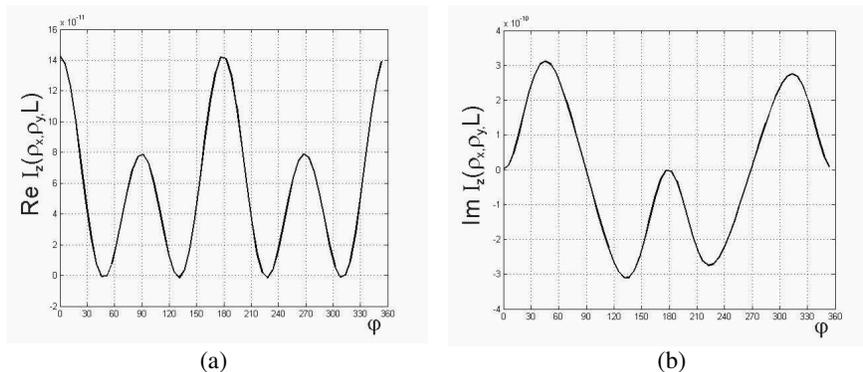


Figure 2: Illustrates the dependence of the correlation function of a scattered electromagnetic field versus polar angle in the direction of strong external magnetic field ($u = 20$) in the principle ($\rho_x = 0, \rho_y = 1$) plane; angle $\gamma_0 = 0, \chi = 10$.

and gaps arise with increasing parameter of anisotropy, which is in agreement with [9]. Oscillations disappear in isotropic case ($\chi = 1$) and the curves have the Gaussian form. The behaviour of the correlation function of scattered radiation is different in the principle yoz and perpendicular planes. The amplitudes of oscillations are increased in proportion to the intensity of an external magnetic field in both observation planes.

REFERENCES

1. Jandieri, G. V., A. Ishimaru, V. G. Jandieri, A. G. Khantadze, and Z. M. Diasamidze, "Model computations of angular power spectra for anisotropic absorptive turbulent magnetized plasma," *Progress In Electromagnetics Research*, PIER 70, 307–328, 2007.
2. Gavrilenko, V. G., A. V. Sarokin, G. V. Jandieri, and V. G. Jandieri, "Some properties of the angular power distribution of electromagnetic waves multiply scattered in a collisional magnetized turbulent plasma," *Plasma Physics Report*, Vol. 31, No. 7, 604–615, 2005.
3. Jandieri, G. V., A. Ishimaru, K. Yasumoto, A. G. Khantadze, and V. G. Jandieri, "Angle-of-arrival of radio waves scattered by turbulent collisional magnetized plasma layer," *International Journal of Microwave and Optical Technology (IJMOT)*, Vol. 4, No. 3, 160–169, 2009.
4. Xu, Z.-W., J. Wu, and Z. S. Wu, "A survey of ionospheric effects on space-based radar," *Waves in Random media*, Vol. 14, S189–S273, 2004.
5. Li, G., B. Ning, W. Wan, and B. Zhao, "Observations of GPS ionospheric scintillations over Wuhan during geomagnetic storms," *Ann. Geophys.*, Vol. 24, 1581–1590, 2006.
6. Wu, Z.-S., H.-Y. Wei, R.-K. Yang, and L.-X. Guo, "Study of scintillation considering inner and outer-scales for laser beam propagation on the slant path through the atmospheric turbulence," *Progress In Electromagnetics Research*, PIER 80, 277–293, 2008.
7. Ginzburg, V. L., *Propagation of Electromagnetic Waves in Plasma*, Gordon and Beach, New York, 1961.
8. Jandieri, G. V., A. Ishimaru, V. G. Jandieri, K. V. Kotetishvili, and T. N. Bzhalava, "A radio scintillation in the collision magnetized plasma," *The 2007 World Congress in Computer Science, Computer Engineering, and Applied Computing WORLDCOMP'07, Proceedings of The International Conference on Scientific Computing*, Vol. CSC2007, 189–195, 2007.
9. Jandieri, G. V., V. G. Gavrilenko, A. Ishimaru, and V. G. Jandieri, "Peculiarities of spatial spectrum of scattered electromagnetic waves in anisotropic inhomogeneous medium," *Progress In Electromagnetics Research B*, Vol. 7, 191–208, 2008.

Radars Imaging of Target above the Gaussian Random Rough Surface Using the Accelerated MOM/PO Hybrid Method

Si-Yuan He, Fang-Shun Deng, Jing-Jing Yao, and Guo-Qiang Zhu
School of Electronic Information, Wuhan University, China

Abstract— In this paper, the range profile analysis for a 3D perfectly electric conductor (PEC) target above a 2D PEC Gaussian random rough surface was successfully implemented using the MOM/PO hybrid method accelerated by the matrix decomposition technique. The numerical scattering model of 3D target/rough surface is constructed by considering the target as the MOM region and the rough surface as the PO region where the RWG vector basis function is employed to expand the induced surface currents. Based on the EM numerical simulation, the wide-band backscattered field of the combined model is obtained using a stepped frequency waveform (SFW), which is produced by linearly sampling the desired bandwidth B at specific frequencies. By performing an inverse discrete Fourier transform (IDFT) on the field data in the frequency domain, the time domain response is provided. Accordingly, the 1-D high resolution range profiles (HRRPs) of the target above the random rough surface is achieved, since the time delay of the signal is associated with the position of the scattering center projected on the line of the sight. The range profiles provide information not only about the position but also the scattering strength of the scattering centers along the range direction.

1. INTRODUCTION

Electromagnetic scattering from a target above or on a rough surface has been an attractive subject in electromagnetic field because it has many applications on radar surveillance, remote sensing, target identification, etc. [1–4]. In the past, several numerical methods have been developed for 2-D target/rough surface scattering, e.g., the forward backward method with spectral acceleration algorithm (GFBM/SAA) [2], the accelerated MOM/PO hybrid method [3]. Recently, the range profile analysis of the 2-D target above a 1-D rough surface has been accomplished in our previous research [4] based on the EM numerical simulation. However, investigation of scattering from 3-D target above the 2-D random rough surface is much more complicated than the 2-D case. Very few reports have been found for the numerical scattering model of 3-D target and rough surface. In this paper, the range profile characteristics and imaging mechanisms of a 3-D PEC target above a 2-D rough surface is investigated based on the EM numerical simulation. The MOM/PO hybrid method [5] accelerated by the UV method [6] is adopted as the fast forward algorithm to obtain the wide-band scattered field. By performing an inverse discrete Fourier transform (IDFT) on the backscattered field data in the frequency domain, the 1-D range profiles of the target above the rough surface is obtained. The range profile characteristics and the imaging mechanisms could be understood and analyzed based on the ray theory. Compared to the target above a planar surface, the interaction scatterings based on the strong reflection are decreased greatly in the rough surface case because of the rough surface diffuse scattering. Simulated results show that the interaction scattering between the target and the bottom surface leads to a series of equivalent range profiles, especially when the bottom surface is smooth. Range locations of the equivalent range profiles made a good agreement with the predictions made from the ray theory. The analysis process could be performed by referring to the case of 2-D target above the 1-D rough surface in our previous research.

2. TARGET AND ROUGH SURFACE SCATTERING MODEL

Figure 1 illustrates the combined model considered in this paper. Here we consider the 3-D cube with side length $s = 2$ m and centered at a height of $H = 2$ m. The bottom surface is of size $L_x \times L_y$. The frequency response of the scattering model is simulated using a SFW, the parameters of which are appropriately designed considering the size of the composite model. The bandwidth is B and the frequency step is Δf , so the range resolution ΔR and the unambiguous range R_u could be computed as $\Delta R = c/2B$; $R_u = c/2\Delta f$. In this paper, a UV technique [6] accelerated MOM/PO hybrid method [5] is adopted as the forward algorithm for the modeling of the induced surface currents on the target (taken as the MOM region) and the rough surface (taken as the PO region). By performing an IDFT on the wide-band backscattered field, the range profiles of

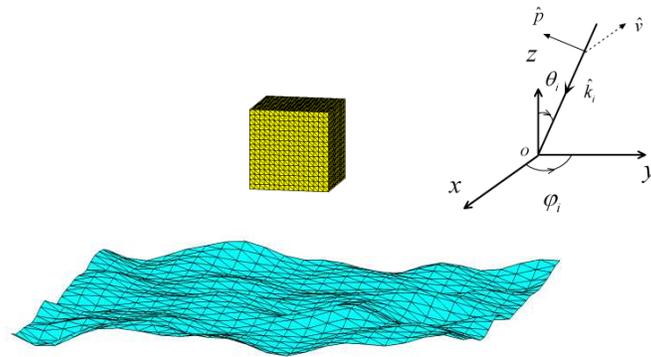


Figure 1: The target/rough surface scattering model.

the composite system are obtained. The total scattered field could be modeled as a combination of the four scattering components contributed by the four current components, J_{t0} (the induced currents on the target in free space), J_{s0} (the induced currents on the rough surface in free space), the interacting currents J_{td} and J_{sd} .

3. NUMERICAL RESULTS

Firstly, the case of the cube above a finite planar surface excited by a plane wave is studied. The parameters are considered as follows, $B = 600$ MHz, $\Delta R = c/2B = 0.25$ m, $\Delta f = 3$ MHz, $R_u = c/2\Delta f = 50$ m, $L_x = L_y = 8$ m. To highlight the target's scattering, the direct scattering of the bottom surface excited by the incident wave is subtracted and thus we get the difference scattering of the composite model, which contains the target direct scattering and the interaction scattering between the target and the bottom surface. In Fig. 2, the scattering centers of the target, shown as 1° , 2° , are followed by a series of equivalent range profiles caused by the interaction scatterings between the target and the rough surface. Peaks $1'$, $2'$, $3'$, $4'$ are the equivalent range profiles contributed by the interaction scatterings of different orders. The range profile characteristics and the imaging mechanisms could be analyzed based on the ray theory. The analysis process could be performed by referring to the case of 2-D target above the 1-D rough surface in our previous research. The peak $1'$ is the first order range profile caused by the double bounce between the left face of the cube and the planar surface. The amplitude of $1'$ is extremely large while the higher order range profiles are observed with much lower intensity. The locations of the range profiles match well with the predictions based on the ray theory. As shown in Fig. 2, both the current components J_{td} and J_{sd} contribute greatly to the first order range profile $1'$.

Next the tapered wave is employed to avoid the edge scattering effects of the bottom surface and thus the simulation of the same cube above the infinite rough surface is accomplished. The results of the same cube above the Gaussian random rough surface of different roughness are considered and compared with the planar surface case (viz. cube above the planar surface). Here the parameters

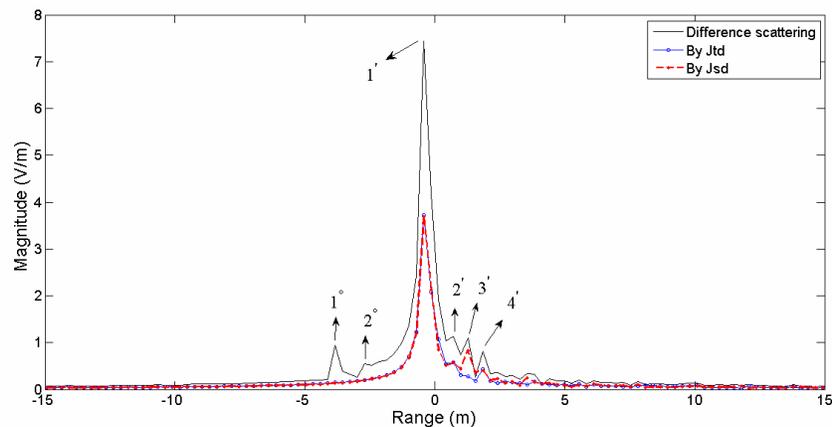


Figure 2: Range profiles of the cube above the planar surface (Current contributions).

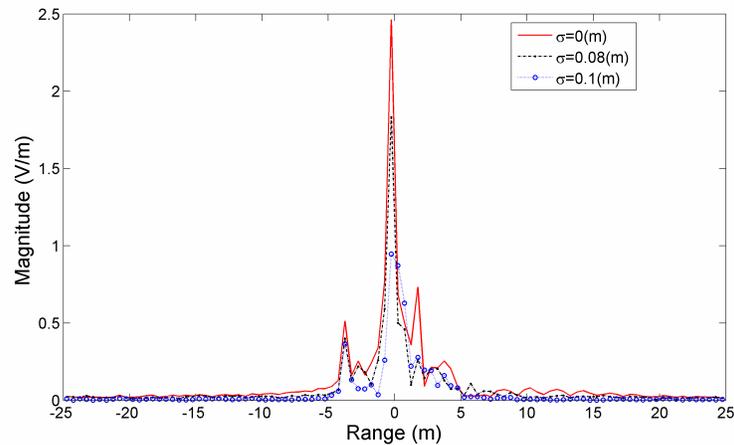


Figure 3: Range profiles of the cube above the infinite rough surface.

are considered as, $B = 300$ MHz, $\Delta R = c/2B = 0.5$ m, $\Delta f = 3$ MHz, $R_u = c/2\Delta f = 50$ m and $L_x = L_y = 24$ m. The correlation length of the rough surface is set as $cL = 1$ m and the RMS height is chosen as $\sigma = 0.1$ m and $\sigma = 0.08$ m respectively in this case. In Fig. 3, it could be observed that as the roughness increases, the intensity of the equivalent range profiles is reduced accordingly because of the rough surface scattering. The decrease in the amplitude of the first order range profile is especially significant due to the great reduce of the specula scattering effect.

4. CONCLUSIONS

The range profile analysis for a 3D target above a 2D Gaussian random rough surface was successfully implemented using the MOM/PO hybrid method accelerated by the matrix decomposition technique. The proposed hybrid method enables possible the analysis of the range profiles for the 3-D target and rough surface composite model. The numerical results could be understood and analyzed based on the ray theory.

ACKNOWLEDGMENT

This work is supported by the National Natural Science Foundation of China under Grant No. 60671040 and the Chinese National High Technology Research Plan (863 Plan) under Grant No. 2007AA12Z172.

REFERENCES

1. Johnson, J. T., "A numerical study of scattering from an object above a rough surface," *IEEE Trans. Antennas Propagat.*, Vol. 50, No. 10, 1361–1367, 2002.
2. Pino, M. R., R. J. Burkholder, and F. Obelleiro, "Spectral acceleration of the generalized forward-backward method," *IEEE Trans. Antennas Propagat.*, Vol. 50, No. 6, 785–797, 2002.
3. He, S. Y. and G. Q. Zhu, "A hybrid MM-PO method combining UV technique for scattering from two dimensional target above a rough surface," *Microwave Opt. Technol. Lett.*, Vol. 49, No. 12, 2957–2960, 2007.
4. He, S. Y., F. S. Deng, H. T. Chen, W. D. Hu, W. X. Yu, and G. Q. Zhu, "Range profile analysis of the 2-D target above a rough surface based on the electromagnetic numerical simulation," *IEEE Trans. Antennas Propagat.*, Vol. 57, No. 10, 3258–3263, 2009.
5. Jakobus, U. and F. M. Landstorfer, "Improved PO-MM hybrid formulation for scattering from three-dimensional perfectly conducting bodies of arbitrary shape," *IEEE Trans. Antennas Propagat.*, Vol. 43, No. 2, 162–169, 1995.
6. Deng, F. S., S. Y. He, H. T. Chen, W. D. Hu, W. X. Yu, and G. Q. Zhu, "Numerical simulation of vector wave scattering from the target and rough surface composite model with 3-D multilevel UV method," *IEEE Trans. Antennas Propagat.*, revised.

Channel Capacity Enhancement by Applying 3-D Space-polarization Diversity to MIMO Systems

L. Hai and Y.-R. Zhang

Country College of Electronic Science and Engineering
Nanjing University of Posts and Telecommunications, China

Abstract— Multi-input multi-output (MIMO) systems have two main styles of space-diversity and polarization-diversity. In order to simulate MIMO systems with arbitrary diversity, the electromagnetic scattering effects are abstracted into a set of statistical parameters, based on which a 3-D stochastic physical MIMO channel model, which can be elastically applied to different conditions, is established. Research work on systems with various diversity is carried out. Simulation results show that systems with both space-diversity and polarization-diversity may gain some advantage when compared with single-diversity systems, and suggest that with applying space-polarization diversity to MIMO systems, we may attain best system performance by choosing most suitable antenna diversity styles for various environments, which is especially meaningful for systems with restrict antenna spaces.

1. INTRODUCTION

Multi-input multi-output (MIMO) systems have two main styles of space-diversity [1] and polarization-diversity [2], in which the space-diversity systems may attain greater system capacity, yet require adequate spaces for large antenna array and are obviously affected by azimuth angles of antenna array, while the polarization-diversity systems require less spaces, yet have limited ability in improving system capacity because of the conspicuous path-loss of horizontal polarization wave. How to make full use of space and polarization diversity is required to be studied.

Channels of MIMO systems change with multi-antenna's positions and polarizations. In order that we can freely study MIMO systems with any antenna diversities, a common MIMO channel model is required. For some traditional simplified MIMO channel models [3, 4], depolarization and path loss together are regarded as path loss, which results in that polarization-diversity is not able to be modeled, while for some MIMO channel models with electromagnetic scattering calculation [5, 6], depolarization is modeled well but large amount of calculation is required.

In this paper, a 3-D space-polarization diversity MIMO channel model, in which the electromagnetic scattering effects are abstracted into a set of statistical parameters, which results in that depolarization and path loss can be easily simulated, is established in Section 2. Then several MIMO systems with different diversities and environments are given, and their performances are discussed with the support of simulation results in Section 3, which is followed by a short conclusion in Section 4.

2. A NEW 3-D SPACE-POLARIZATON DIVERSITY MIMO CHANNEL MODEL

In this section, we propose a scattering matrix to describe scattering process

$$A_l = \begin{pmatrix} A_{xx}^{(l)} & A_{xy}^{(l)} & A_{xz}^{(l)} \\ A_{yx}^{(l)} & A_{yy}^{(l)} & A_{yz}^{(l)} \\ A_{zx}^{(l)} & A_{zy}^{(l)} & A_{zz}^{(l)} \end{pmatrix} \quad (1)$$

where each element may be properly designed to simulate various scatters, then different environments. This simulating process requires much less calculation than electromagnetic scattering calculation.

Consider a mix-diversity system consists of M transmit antennas and N receive antennas of arbitrary positions and polarizations. Suppose the signal on m th transmit antenna, which has a pattern as G_{mT} , is $x_m(t)$, which has a polarization defined as $(\alpha_{mT}, \beta_{mT})$, where α_{mT} is the angle between polarization orientation of m th transmit antenna and z -axis, β_{mT} is the angle between projection of polarization orientation of m th transmit antenna on xoy plane and x -axis. For convenient analysis, we divide $x_m(t)$ into

$$\begin{pmatrix} T_{mx} \\ T_{my} \\ T_{mz} \end{pmatrix} = \begin{pmatrix} \sin \alpha_{mT} \cos \beta_{mT} \\ \sin \alpha_{mT} \sin \beta_{mT} \\ \cos \alpha_{mT} \end{pmatrix} x_m(t) \quad (2)$$

Suppose there are L scatters between transmitter and receiver. Suppose the polarization of n th receive antenna, which has a pattern as G_{nR} , is defined as $(\alpha_{nR}, \beta_{nR})$. The signal received on n th receive antenna is

$$y_n(t) = \sum_{l=1}^L G_{nR}(\vec{r}_{nl}) G_{mT}(\vec{r}_{ml}) g_{nl} g_{ml} \left[\frac{-j2\pi(|\vec{r}_{ml}| + |\vec{r}_{nl}|)}{\lambda} \right] \begin{pmatrix} \sin \alpha_{nR} \cos \beta_{nR} \\ \sin \alpha_{nR} \sin \beta_{nR} \\ \cos \alpha_{nR} \end{pmatrix}^T A_l \begin{pmatrix} T_{mx} \\ T_{my} \\ T_{mz} \end{pmatrix} \quad (3)$$

where g_{ml} is path loss from m th transmit antenna to l th scatter, $\frac{2\pi|\vec{r}_{ml}|}{\lambda}$ is phase offset from m th transmit antenna to l th scatter, and g_{nl} is the path loss from l th scatter to n th receive antenna, $\frac{2\pi|\vec{r}_{nl}|}{\lambda}$ is the relevant phase offset. If the distances between diversity antennas are much smaller than the distance between transmitter and receiver, then path loss of l th path may be

considered as $g_{ml}g_{nl} = g_l$, which is irrelevant to m and n . Let $\begin{pmatrix} r_1^n \\ r_2^n \\ r_3^n \end{pmatrix} = \begin{pmatrix} \sin \alpha_{nR} \cos \beta_{nR} \\ \sin \alpha_{nR} \sin \beta_{nR} \\ \cos \alpha_{nR} \end{pmatrix}$,

$\begin{pmatrix} t_1^m \\ t_2^m \\ t_3^m \end{pmatrix} = \begin{pmatrix} \sin \alpha_{mT} \cos \beta_{mT} \\ \sin \alpha_{mT} \sin \beta_{mT} \\ \cos \alpha_{mT} \end{pmatrix}$, the channel response from m th transmit antenna to n th receive antenna may be written as

$$h_{mn} = \sum_{l=1}^L G_{mT}(\vec{r}_{ml}) G_{nR}(\vec{r}_{nl}) g_l \exp \left[\frac{-j2\pi(|\vec{r}_{ml}| + |\vec{r}_{nl}|)}{\lambda} \right] \sum_{i=1, j=1}^3 r_i^n t_j^m a_{ij}^{(l)} \quad (4)$$

Thus the channel response model under 3-D conditions is established by geometrical means.

3. SYSTEM MODELS AND SIMULATION RESULTS

For simplicity, we consider a 2×2 MIMO system model which has an airview like Fig. 1, where D is the distance between transmit end and receive end, γ_t and γ_r are azimuth angles of transmit antenna array and receive antenna array. Consider there are L scatters on the circle which has a radius of R and a centre on the receive end. Let the transmit end take coordinate of $(0, 0)$, and the direction from transmit end to receive end parallels y -axis, and the vertical parallels z -axis. Scatters and antennas may take different altitudes on z -axis.

Firstly, we consider a situation where the same antenna diversity is applied in both transmit end and receive end. For space-diversity and mix-diversity systems, T_{x1} and T_{x2} take coordinates of $(0, \lambda)$ and $(0, -\lambda)$, while R_{x1} and R_{x2} take coordinates of (D, λ) and $(D, -\lambda)$, for polarization-diversity and mix-diversity systems, polarization of T_{x1} and R_{x1} is $(0, 0)$, while polarization of T_{x2} and R_{x2} is $(\pi/2, 0)$. Performances of different diversity systems in a uniform depolarization environment where electric field has the same chance of being scattered to all directions, and a rather actual room environment where component of electric field on propagation direction attenuates conspicuously, are shown in Fig. 2 and Fig. 3.

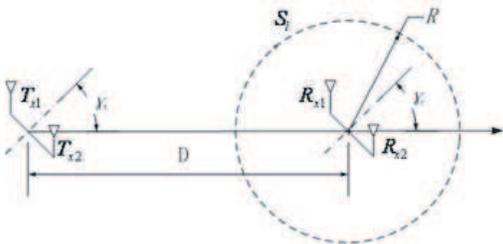


Figure 1: Channel environment and geometry of the model.

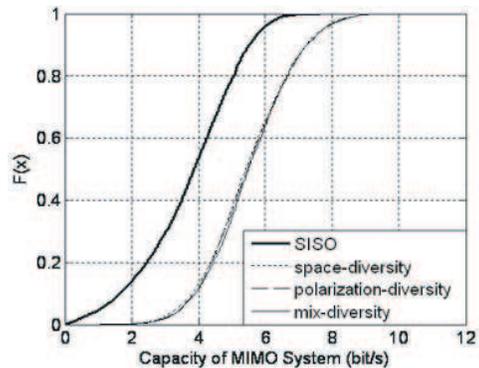


Figure 2: Space diversity on x -axis, performances of different systems in uniform depolarization environment.

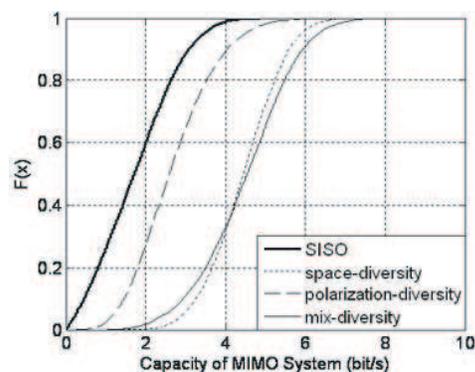


Figure 3: Space diversity on x -axis, performances of different systems in room environment.

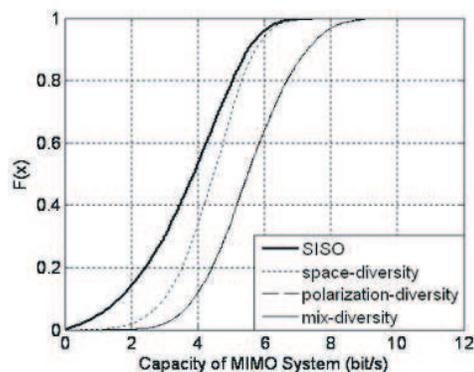


Figure 4: Space diversity on y -axis, performances of different systems in uniform depolarization environment.

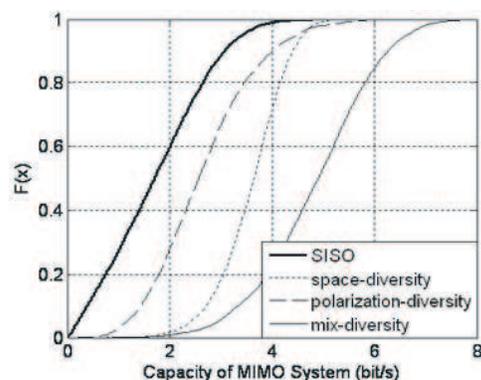


Figure 5: Space diversity on y -axis, performances of different systems in room environment.

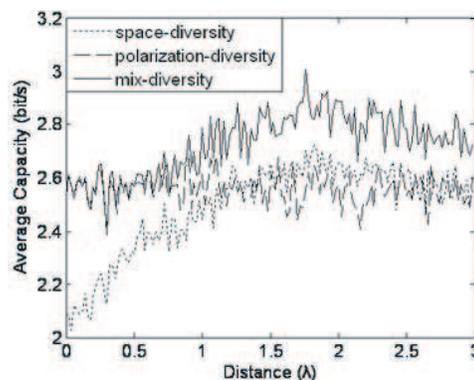


Figure 6: Average capacities of different diversity on transmit end in room environment.

From Fig. 2, we can see that in a uniform depolarization environment where path-loss of horizontal polarization wave is not remarkable, which results in that polarization-diversity and mix-diversity systems gain nearly the same performance as space-diversity system. And from Fig. 3, we know that because of path-loss of horizontal polarization wave, polarization-diversity system gains less performance than space-diversity system, while mix-diversity system works even better.

Secondly, we consider T_{x1} and T_{x2} take coordinates of $(\lambda, 0)$ and $(-\lambda, 0)$, while R_{x1} and R_{x2} take coordinates of $(D + \lambda, 0)$. Performances of different diversity systems are shown in Fig. 4 and Fig. 5.

From Fig. 2, we can see that when space diversity appears on y -axis, space-diversity system has a poor performance because that correlation coefficients between sub-channels is not effectively reduced. And from Fig. 3, we know that performance of polarization-diversity is also reduced because of path-loss, while mix-diversity system gains the best performance.

Thirdly, we consider a system which has random azimuth angles in both ends, and calculate its average capacities in room environment. Let the receive end take a mix-diversity antenna array, performance of different diversity antenna arrays in transmit end are shown in Fig. 6.

From Fig. 6, we can see that average capacity of space-diversity system is seriously affected by distance between T_{x1} and T_{x2} , and average capacity of polarization-diversity system is a little less than space-diversity system when distance between T_{x1} and T_{x2} is far enough, while the mix-diversity system has a much better and stable average capacity than single-diversity systems.

4. CONCLUSION

In this paper, both space-diversity and polarization-diversity are applied to MIMO systems. Simulation results show that for different environments, single-diversity systems achieve different and unstable performance, while performance of mix-diversity system remains well.

There should be a most suitable diversity style for a certain channel environment. Further work

should be carried out in order that we can give a convenient way to find a most efficient diversity style for various environments.

ACKNOWLEDGMENT

This work was supported by the National Basic Research Program of China under Grant 2007CB310-603.

REFERENCES

1. Byers, G. J. and F. Takawira, "Spatially and temporally correlated MIMO channels: Modeling and capacity analysis," *IEEE Transactions on Vehicular Technology*, Vol. 53, No. 3, 634–643, 2004.
2. Alatossava, M., L. Hentilä, V.-M. Holappa, and J. Meinilä, "Comparison of outdoor-to-indoor and indoor-to-outdoor MIMO propagation characteristics at 5.25 GHz," *IEEE 65th Vehicular Technology Conference*, 445–449, Dublin, Ireland, Apr. 22–25, 2007.
3. Abdi, A. and M. Kaveh, "A space-time correlation model for multielement antenna systems in mobile fading channels," *IEEE J. Select. Areas Commun.*, Vol. 20, 550–560, 2002.
4. Pätzold, M. and B. O. Hogstad, "A space-time channel simulator for MIMO channels based on the geometrical one-ring scattering model," *Wireless Communications and Mobile Computing, Special Issue on Multiple-Input Multiple-Output (MIMO) Communications*, Vol. 4, No. 7, 727–737, Nov. 2004.
5. Svantesson, T., "A double-bounce channel model for multi-polarized MIMO systems," *Proc. IEEE Vehicular Technology. Conf.*, Vol. 2, 691–695, 2002.
6. Wang, X.-D. and J.-D. Li, "A generic channel model for MIMO systems," *International Conference on Communication Technology*, 1–4, Nov. 27–30, 2006.

Convergecast of Multi-destinations in Zigbee Tree-based Wireless Sensor Network

P. Juleang¹, S. Mitatha¹, and P. P. Yupapin²

¹Hybrid Computing Research Laboratory, Faculty of Engineering
King Mongkut's Institute of Technology Ladkrabang, Bangkok 10520, Thailand

²Advanced Research Center for Photonics, Faculty of Science
King Mongkut's Institute of Technology Ladkrabang, Bangkok 10520, Thailand

Abstract— Multi-destinations in wireless sensor network have extremely benefited for monitoring system. Then, convergecast is a fundamental operation in wireless sensor network. Existing convergecast solution focused on latency and energy consumption. However, a good design should be complaint to standard. This paper proposes convergecast solution of multi-destinations in Zigbee tree-based wireless sensor network and proves that this solution is simulation. Our solution is optimal convergecast of multi-destination in wireless sensor network and using accompany with network's resource for all destinations. Then, our solution is compliant with the low-power design of IEEE 802.15.4. Simulation results show that the proposed solution can indeed achieve convergecast of multi-destinations.

1. INTRODUCTION

Nowadays, usability of wireless sensor network (WSN) used widely. Especially, WSN used by Zigbee standard. For example, Smart home system [1] which increase capability more than regular home. Zigbee car parking management system [2] is using Zigbee network for spots status report such as having car in spot. Location system [3–5] is using Zigbee for identify position of person or thing such as for identify position of elderly and disabled people. Then, usability of Zigbee in monitoring system sense all environment such as bridge condition, air pollution, precision viticulture, home care system [6–9]. Including, Zigbee wireless sensor network have benefit for logistic system [10].

If usability of Zigbee wireless sensor network in present have system's design for Zigbee multi-destination wireless sensor network will have benefited for monitoring system. For example, mostly person in Bangpakong to rear snapper. If usability of wireless sensor network was happen for water and water pollution system will help fisherman for plan to rear snapper. Which, fishermen want monitoring system's data form Bangpakong river. This case, Multi-destinations in wireless sensor network have extremely benefited for monitoring system.

2. OVERVIEW OF ZIGBEE

Zigbee is a standard that defines a set of communication protocols for low-data-rate wireless network. Zigbee-based wireless devices operate in 868 MHz, 915 MHz and 2.4 GHz frequency bands. Then, the maximum data rate is 250 k bits per second. Zigbee is targeted mainly for battery-powered applications where low data rate, low cost and long battery life are main requirements. The Zigbee standard is developed by the Zigbee Alliance, which has hundreds of number companies, from semiconductor industry and software developers to original equipment manufacturer and installers.

There are two methods for channel access: contention based or contention free. In contention-based channel access, all the devices that want to transmit in the same frequency channel use the CDMA-CA, and the first one that finds the channel clear starts transmitting. In the contention-free method, the PAN coordinator dedicates a specific time slot to a particular device. This called a guaranteed time slot (GTS). To provide a GTS, the PAN coordinator needs to ensure that all the devices in the network are synchronized. Beacon is a message with specific format that is used to synchronize the clocks of nodes in network. A coordinator has the option to transmit beacon signals to synchronize the devices attached to it. This called a beacon-enabled PAN.

One of the advantages of a beacon-enabled network is the availability of guaranteed time slots (GTSs). The beacon are MAC frames that contain beacon information such as the time interval between the beacons and number of GTSs. In beacon-enable operation, it is possible to use a superframe structure. A superframe, shown in Figure 1, is bounded by two beacon frame. The use of super frame structure is optional in the IEEE 802.15.4 standard. There can be up to three

types of periods in a superframe: the contention access period (CAP), the contention-free period (CFP), and the inactive period. This paper proposes convergecast solution which is management other slot called slot assignment.

3. DECIDE PROBLEM

This section formally defines the convergecast problem in ZigBee networks [11]. Given a ZigBee network, we model it by a graph $G = \{V, E\}$ where V contains all routers and the coordinator and E contains all symmetric communication links between nodes in V . The coordinator also serves as the sink of the network. End devices can only associate with routers, but are not included in V . From G , we can construct an interference graph $G_I = \{V, E_I\}$ where edge $(i, j) \in E_I$ if there are direct/indirect interferences between i and j . We denote by $k = 2^{(BO-SO)}$ the number of active portions (or slots) per beacon interval [11].

The beacon scheduling problem is to find a slot assignment $s(i)$ for each router $i \in V$, where $s(i)$ is an integer and $s(i) \in [0, k - 1]$, such that router i 's active portion is in slot $s(i)$ and $s(i) \neq s(j)$ if $(i, j) \in E_I$. Here, the slot assignment means the position of the outgoing superframe of each router (the position of the incoming superframe, as clarified earlier, is determined by the parent of the router).

Given a slot assignment for G , the report latency from node i to node j [11], where $(i, j) \in E$, is the number of slots, denoted by d_{ij} , that node i has to wait to relay its collected sensory data to node j .

$$d_{ij} = (s(j) - s(i)) \bmod k. \tag{1}$$

Then, report latency of Zigbee beacon-enabled wireless sensor network $L(G)$ shown in Equation (2). Good design of Zigbee wireless sensor network designing is minimum report latency $L(G)$.

$$L(G) = \sum (d_{ij}) \text{ on longest part of network } (G). \tag{2}$$

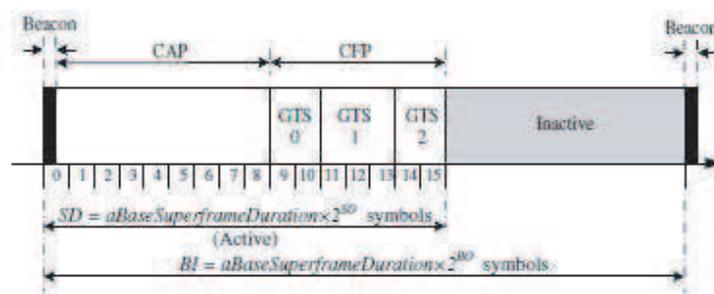


Figure 1: IEEE 802.15.4 superframe structure.

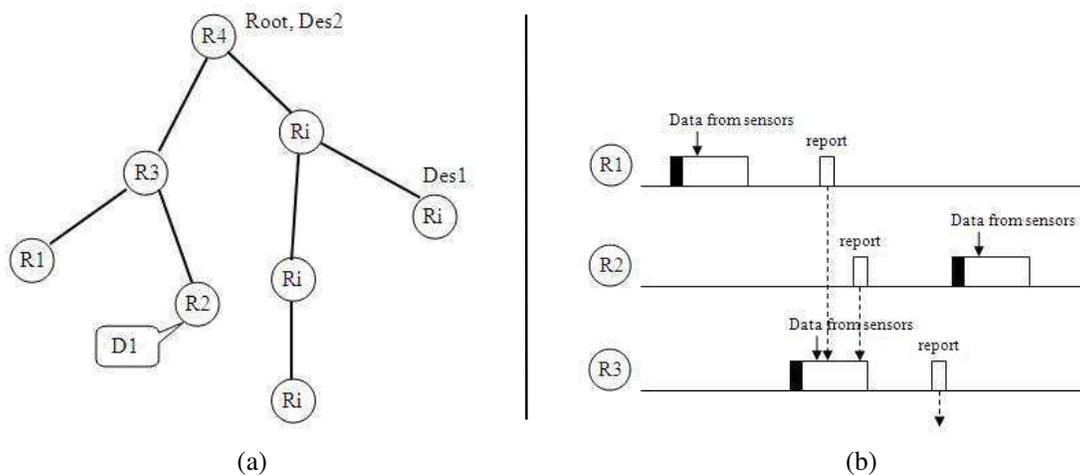


Figure 2: Problem of multi-destinations in wireless sensor network.

Point, two problems of multi-destinations in wireless sensor network shown in Figure 2. Figure 2(a) shows sensors' data wanted from destination can't send to it too. This paper called Data Incorrect Problem (DIP). If slot assignment haven't optimal procedure then Data1 (D1) which wanted from Destination1 (Des1) not send to it (Figure 2(a)) or want have special procedure for sent to it which cause low-speed in network. Figure 2(b) shows convergecast problem cause more report latency or low-speed in Zigbee wireless sensor network. From Figure 2(a), router1 (R2) want send data to Des2 but have convergecast schedule shown in Figure 2(b) show data from R2 must wait a one superframe. But, if slots assignment have optimal procedure then data from R1 not wait that data from R1 send with data from R2 at the same time. This problem called Minimum Delay Beacon Scheduling Problem (MDBP) [11]. From two problems, this paper proposes convergecast solution of multi-destinations in Zigbee tree-based wireless sensor network for solve problems and minimum report latency $L(G)$.

4. CONVERGECAST SOLUTION

Convergecast solution proposed in this paper. Given $G = (V, E)$, $G_I = (V, E_I)$ and k , we propose solution. Our Solution is composed of the following step:

Step 1: From G , we first construct a breadth first search (BFS) tree T which rooted at one in destinations.

Step 2: Check DIP, Has DIP in tree network from Step 1?

Yes, it has. We do Step 1 by change root to other destinations.

No, it hasn't. We do Step 3.

If change root to all destinations then No in all answer. We must from into group of destinations which destinations in group must can send data to each destination. Look, each group is one destination. we do Step 1.

Step 3: We traverse vertices of T in a bottom-up manner. For these vertices in depth d , we first sort them according to their degrees in G_I in a descending order. Then we sequentially traverse these vertices in that order. For each vertex v in depth d visited, we compute a temporary slot number $t(v)$ for v as follows.

If v is a leaf node, we set $t(v)$ to the minimal non-negative integer l such that for each vertex u that has been visited and $(u, v) \in E_I$, $(t(u) \bmod k) \neq l$.

If v is an in-tree node, let m be the maximum of the numbers that have been assigned to v 's children, i.e., $m = \max(t(\text{child}(v)))$, where $\text{child}(v)$ is the set of v 's children. We then set $t(v)$ to the minimal non-negative integer $l > m$ such that for each vertex u that has been visited and $(u, v) \in E_I$, $(t(u) \bmod k) \neq (l \bmod k)$.

After every vertex v is visited, we make the assignment $s(v) = t(v) \bmod k$.

Step 4: Vertices are traversed sequentially from t in a top-down manner. When each vertex v is visited, we try to greedily find a new slot l such that $(s(\text{parent}(v)) - l) \bmod k < (s(\text{parent}(v)) - s(v)) \bmod k$, such that $l \neq s(u)$ for each $(u, v) \in E_I$, if possible. Then we reassign $s(v) = l$.

If multi-destinations in Zigbee tree-base beacon-enabled wireless sensor network used convergecast solution which this paper proposed such that wireless sensor network guarantee minimum report latency and solving Data Incorrect Problem (DIP).

5. SIMULATION RESULTS

The simulations results is show compared to the random slot assignment and our solution (our algorithm). Then, results shown in Figure 3. Figure 3(a) shows relationship between report latency of multi-destinations Zigbee wireless sensor network and number of Zigbee routers for two destinations network. However, Figure 3(b) shows relationship between report latency of multi-destinations Zigbee wireless sensor network and number of Zigbee routers for three destinations network.

Explaining, Figure 3(a) and Figure 3(b) show compared to the random slot assignment and our solution in multi-destinations Zigbee wireless sensor network. They are apparent our solution for slot assignment have report latency $L(G)$ less than random slot assignment. And, increasing of number of destinations cause decreasing-speed and increasing-report latency for the equal of number of routers in multi-destinations Zigbee wireless sensor network. Finally, our solution which proposed can indeed achieve convergecast of multi-destinations.

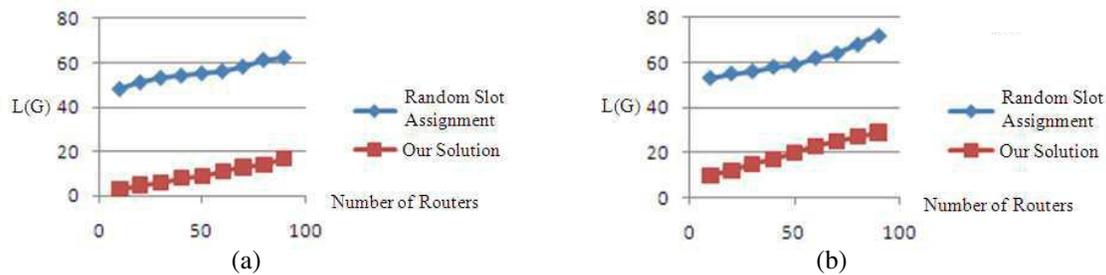


Figure 3: Graph show simulation results.

6. CONCLUSION

In this paper, we have propose convergecast solution of multi-destinations in Zigbee tree-based beacon-enabled wireless sensor network for solving DIP and decreasing report latency $L(G)$ which defined in Meng-Shiuan Pan, Yu-Chee Tseng [11]. It is a slot assignment which optimal convergecast and using accompany with network's resource for all destinations. Simulation, compared to the random slot assignment and our solution can effectively schedule the ZigBee routers' beacon times to achieve convergecast.

Future, it deserves to consider extending this work to multimedia of multi-destinations in Zigbee tree-based beacon-enabled wireless sensor network which extend from Changsu Suh, Zeeshen Hameed Mir, Young-Bea Ko [12] and multi-destinations in Zigbee mesh wireless sensor network.

REFERENCES

1. Tsou, Y.-P., C.-Y. Chen, and C.-T. Lin, "Zigbee technology development for smart home appliance application," Energy & Resource Laboratory, Taiwan, 2007.
2. Shim, S., S. Park, and S. Hong, "Parking management system using Zigbee," *International Journal of Computer Science and Network Security*, Vol. 6, No. 9B, September 2006.
3. Park, W.-C. and M.-H. Yoon, "The implementation of indoor location system to control Zigbee home network," *SICE-ICASE International Joint Conference*, Korea, 2006.
4. Yao, Q. and F.-Y. Wang, "Fellow, Location estimation in Zigbee network-based on fingerprinting," *IEEE Online*, 2007.
5. Marco, A., R. Casas, J. Falco, H. Gracia, J. I. Artigas, and A. Roy, "Location-based services for elderly and disabled people," *International Journal of Computer Communications*, Vol. 31, 2008.
6. Chae, M. J., H. S. Yoo, and J. R. Kim, "Bridge condition monitoring system using wireless network (CDMA and Zigbee)," *Proceedings of ISARC*, 332–334, 2006.
7. Kwon, J.-W., Y.-M. Park, and S.-J. Koo, "Design of air pollution monitoring system using zigbee network for Ubiquitous-City," *International Conference on Convergence Information Technology*, 2007.
8. Moraisa, R., M. A. Fernandesb, and S. G. Matosb, "A Zigbee multi-powered wireless acquisition device for remote sensing applications in precision viticulture," *International Journal of Computer and Electronic in Agriculture*, Vol. 62, 2008.
9. Huang, M.-C., J.-C. Huang, J.-C. You, and G.-J. Jong, "The wireless sensor network for home-care system using zigbee," *Proceedings of the Third International Conference on International Information Hiding and Multimedia Signal Processing*, Vol. 1, 643–646, Taiwan, 2007.
10. Ruiz-Garcia, L., P. Barreiroa, and J. I. Robla, "Performance of Zigbee-based wireless sensor nodes for real-time monitoring of fruit logistics," *International Journal of Food Engineering*, Vol. 87, 405–415, 2008.
11. Pan, M.-S. and Y.-C. Tseng, "Quick convergecast in Zigbee beacon-enabled tree-based wireless sensor networks," *International Journal of Computer Communications*, Vol. 31, 999–1011, 2008.
12. Suh, C., Z. H. Mir, and Y.-B. Ko, "Design and implementation of enhanced IEEE 802.15.4 for supporting multimedia service in wireless sensor networks," *International Journal of Computer Communications*, 2008.

Conception of Patch Antennas in the GSM and UMTS Band

M. Iftissane^{1,2}, S. Bri^{1,2}, and L. Bellarbi²

¹Microwave and Materials Group, ESTM, B. P 3103, Meknès, Morocco

²Electrical Engineering Laboratory, ENSET, Rabat, Morocco

Abstract— In this paper, we present the results of simulations of a series of patch antennas and also the experimental measures by means of a network analyzer. The conception of these patch antennas are realized by software HFSS “Ansoft-High Frequency Structure Simulator” and ADS “Advanced Design System”, based themselves essentially on the variation of the shape of the antenna and its conductive material, the nature and the thickness of the substratum to have a structure which resonates in the frequencies used for precise applications. This result is compared with those published in the bibliography.

1. INTRODUCTION

Many mobile radio services are expanding and attracting more and more users. Access to these services from a single terminal requires the use of compact antennas multifrequency and multi polarization particularly disappointing for simultaneous FM (Frequency Modulation), DVB (Digital Video Broadcasting), cellular (GSM, UMTS) and GPS (Global Positioning System). It was noted that the frequencies used by these applications are spread over several octaves and is therefore difficult to design a single structure with the characteristics required for access to these services [1–3]. To overcome this problem some authors have proposed some compact antennas GSM-GPS with interesting features. Other researchers have focused their work on the reception by the same antenna broadcasts DAB (Digital Audio Broadcasting) and DVB. The material properties (high permittivity, special forms). Have been exploited by many investigators to design variants of compact, multifrequency antennas [4–7].

In this paper, we present generalities on antennas microstrip or we define the basic parameters of an antenna (gain, bandwidth, radiation pattern), types of feed (for microstrip line, coaxial, coupled with opening ring coupled by proximity). To validate the proposed equivalent models, we present the study of patch antennas. We will compare the results of the simulation by HFSS and ADS and the results published in the literature [8–10].

2. SIMULATIONS

We have simulated the printed antennas using the software Ansoft-HFSS (High Frequency Structure Simulator) and ADS (Advanced Design System). This powerful simulation software which allows to represent the distribution of fields and calculate the parameters S_{ij} passive microwave structures. The simulation technique used to calculate the three dimensional electromagnetic field inside a structure is based on the finite element method (FEM). The principle of the method is to divide the study area into many small regions (tetrahedrons), then calculate the local electromagnetic field in each element. HFSS uses an interpolation method combined with an iterative process in which a mesh is created automatically and redefined in the critical regions. The simulator generates a solution based on the predefined initial mesh. Then, it refines the mesh in regions where there is a high density of errors, and generates a new solution.

2.1. Influence of Spatial Parameters of Printed Antennas

The following paragraph illustrates the results of simulations of several cases of printed antennas in the environment HFSS. We will submit applications, in particular calculation of S_{ij} parameters of a printed antenna fed by a microstrip line and the input impedance.

Feed can also be done by direct connection to a coaxial line. The central conductor of the coax is then connected at a point on the axis of symmetry of the radiating element, more or less close to the board to adjust the impedance. The outer conductor is connected to the plane. The major drawback of this technique is that it can be used with the polymer substrate, air or metal plates welded. We note that the resonant frequency and the level of S_{11} depends on the length of the antenna as shown in Table 2.

The resonant frequency and the reflection coefficient S_{11} is inversely proportional to the length l of the microline. For a microstrip line, the width W is given by the following formula: $W =$

Table 1: Parameters of the antenna.

Element	Airbox (mm * mm)	Thickness (mm)
Substrate	20 * 30	0.1
Ground plane	20 * 30	0
Microstrip $w_i \times l_i$	$L * W$	0
Patch, $W \times L$	10*14	0

Table 2: Simulation results of the antenna with variable length.

Length (mm)	Resonance frequency (GHz)	$ S_{11} $ (dB)
6	1.06	-21
8	0.98	-17
10	0.9	-16

Table 3: Simulation results of the antenna with variable width.

Width w (mm)	Resonance frequency (GHz)	Reflection coefficient $ S_{11} $ (dB)
2	1.16	-6.92
4	1.31	-13.65
6	1.41	-17.91

Table 4: Simulation results of the patch antenna with variable tickness.

Thickness h (mm)	Resonance frequency (GHz)	$ S_{11} $ (dB)
0.5	1.06	-20.56
1	1.09	-13.56
1.53	1.16	-6.85

$(C/2f_r) * 1/[(1 + \epsilon_r)/2]0.5$. To validate this relationship, we analyzed a line of 50Ω with a length of 6 mm. The resonant frequency and the level of S_{11} depends on the width of the antenna as shown in Table 3.

When the width of the microstrip line increases, the resonant frequency and S_{11} are also increasing, by varying the thickness of the substrate ($\epsilon_r = 4.32$ (glass epoxy)) and keeping all other parameters fixed, the curves of S_{11} as a function of the resonant frequency data in Table 4.

We note that varying thicknesses of the substrate, the resonant frequency and the level varies S_{11} . The more the thickness increases, the resonance frequency increases, the magnitude of S_{11} decreases. Note that the resonant frequency increases when ϵ_r decreases. Miniaturization of antennas can be achieved by choosing materials with high permittivity. In conclusion, we note that the resonant frequency of a microstrip line varies depending on its length, its width, the permittivity and thickness of the substrate. The antenna of rectangular microstrip resonator is similar to a microstrip line of length L and width W . The length of w is chosen close to $\lambda_g/2$ where λ_g is the wavelength in the line. It can therefore be based on the results found for a microstrip line at the antenna design rectangular or other shapes.

2.2. Antenna Line Buried

The antenna line buried structure is performed in two substrates with a microstrip line plated on the bottom substrate that ends in an open circuit in the printed patch on the upper substrate. This close coupling allows to improve bandwidth and reduce stray radiation. The dimensions $W = L = 5$ mm from the patch fix the resonance of the antenna and $f = 10$ GHz. $W_i =$ width = 0.635 mm from the microstrip feed, can have an impedance of 50Ω at the input antenna. A good adaptation is achieved for a microstrip of length $l_i = 2.5$ mm. The two alumina substrates are modeled by two identical dielectric layers of thickness 0.635 mm, the real part of permittivity is 9.6 (Figure 1).

Figure 2 shows the variation of reflection coefficient S_{11} at the entrance of the antenna versus frequency. It shows that if we consider $|S_{11}| \leq -10$ dB we will have a wide band around the resonance frequency ranging from 9.2 to 9704 GHz.

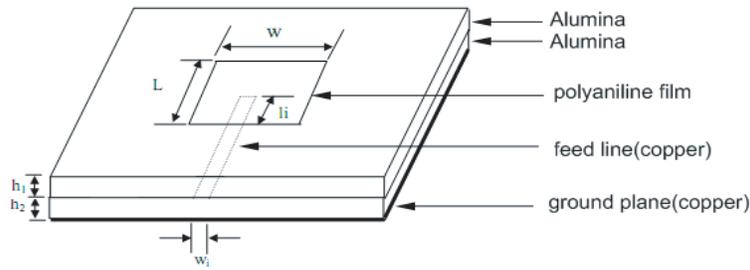


Figure 1: Schematic of the antenna line buried.

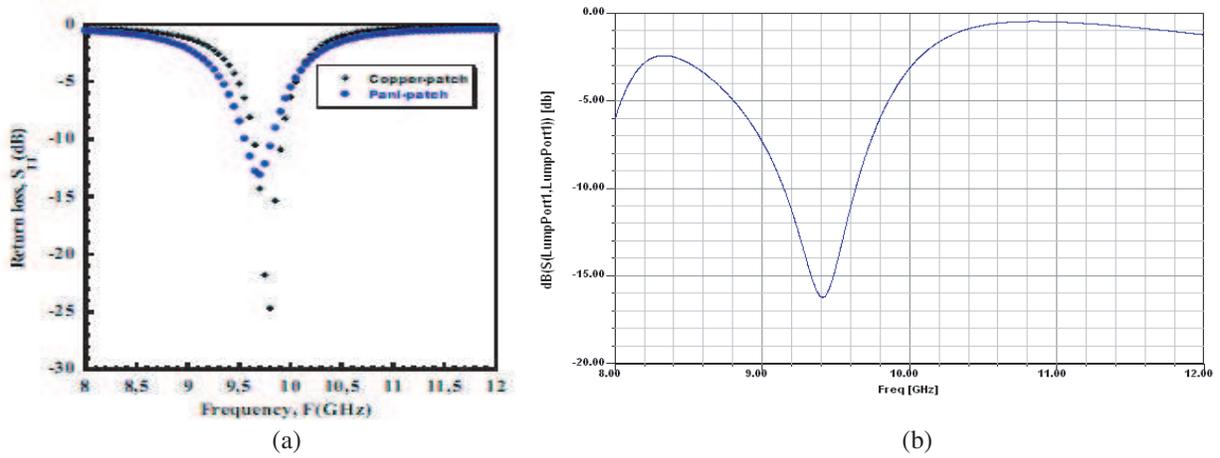


Figure 2: Reflection coefficient as a function of frequency: (a) = [7], (b) = HFSS.

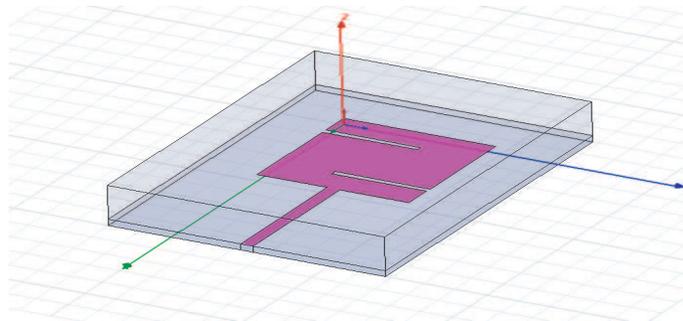


Figure 3: Design of the antenna by HFSS.

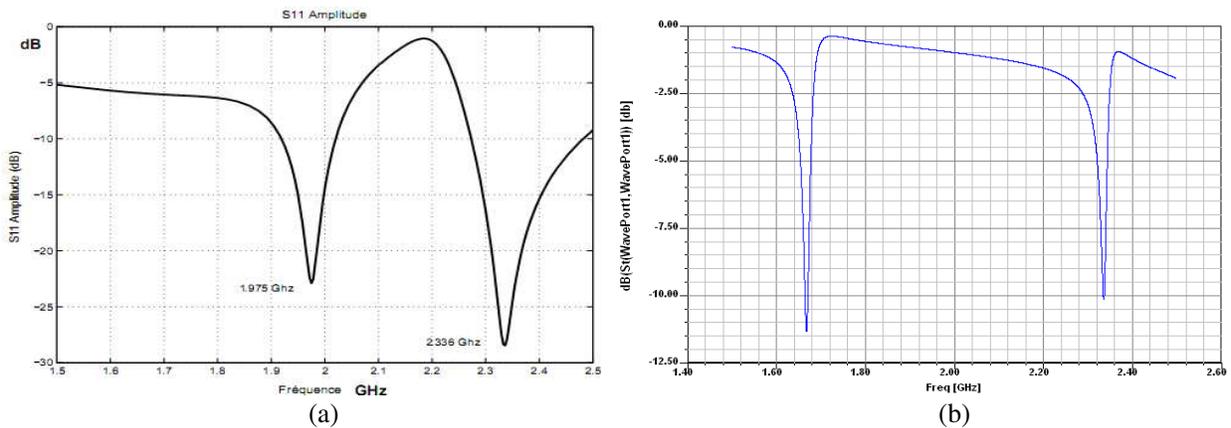


Figure 4: Variation of S_{11} as a function of frequency: (a) = [8], (b) = HFSS.

2.3. Antenna Form of ‘2’

In this case, we take $\epsilon_r = 4.32$ (glass epoxy) $h = 1.53$ mm with 2 layers of 35 μm copper ground plane with copper (80 mm/60 mm) (Figure 3).

Figure 4 presents the variation of S_{11} in two bands: $B_1 = 1918$ to 2022 GHz and $B_2 = 2273 - 2.48$ GHz $|S_{11}| < -10$ dB.

3. EXPERIMENTAL MEASUREMENTS OF PATCH ANTENNAS

Taking into account the design steps mentioned in the previous chapter, we have made various prototypes of antennas as shown in Figure 5, using as substrate the ‘glass epoxy’ type G_{11} with a relative permittivity $\epsilon_r = 4.38$ and $d = 1.6$ mm thick with 2 layers of copper 35 microns.

3.1. Measurements and Results

The antenna characteristics made, were measured with a vector network analyzer E5061A type operating in the band 300 kHz–3000 MHz. The Figure 6 below shows the variation of S_{11} as a function of the frequency band [1000–2500] MHz. This antenna structure “Antenna in (‘U’)” has two resonance frequencies: $f_{r1} = 1600$ MHz and $f_{r2} = 2000$ MHz.

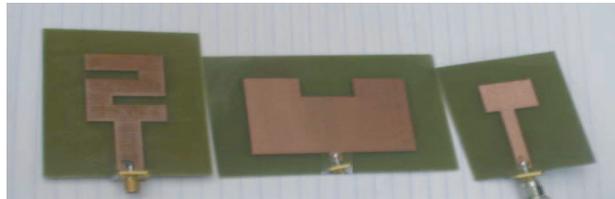


Figure 5: Prototype of the the realized antennas.

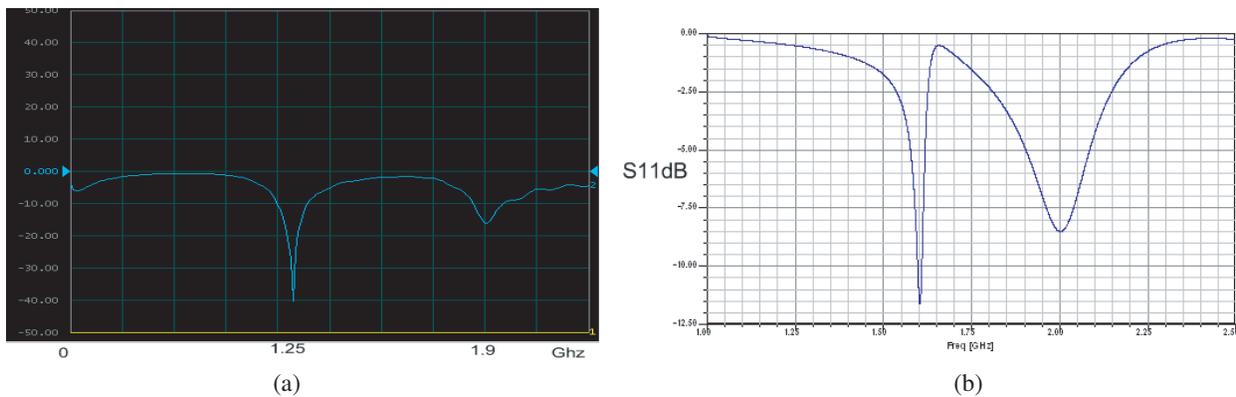


Figure 6: $S_{11} = F(f)$ in dB): (a) = measurement by network analyzer, (b) = HFSS.

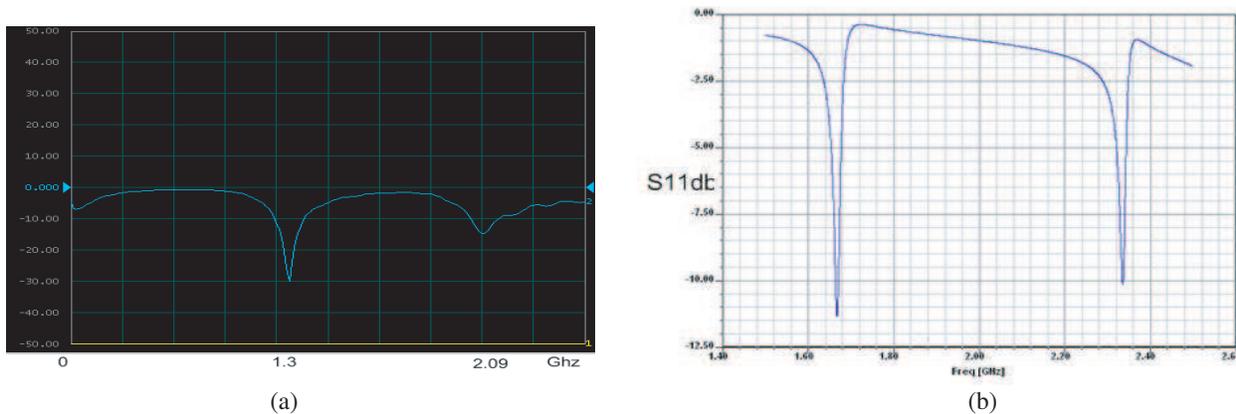


Figure 7: Variation of reflection coefficient (S_{11} in dB versus frequency): (a) = measurement by network analyzer, (b) = HFSS.

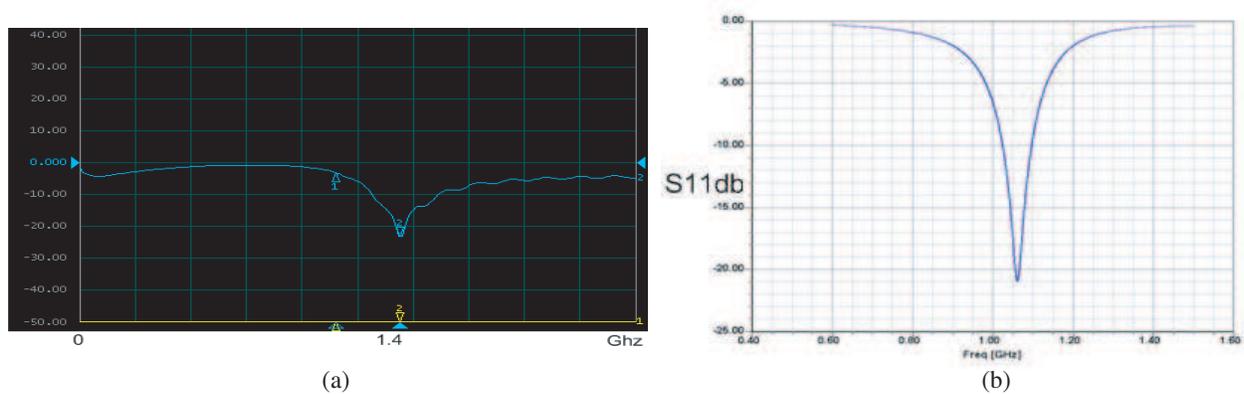


Figure 8: Variation of reflection coefficient (S_{11} in dB) versus frequency: (a) = measurement by network analyzer, (b) = HFSS.

The Figure 7 shows the variation of S_{11} as a function of the frequency band [1400, 2600] MHz. This antenna structure carried ‘2’ has two resonance frequencies: $f_{r1} = 1675$ MHz and $f_{r2} = 2232$ MHz.

The Figure 8 below shows the variation of S_{11} as a function of the frequency band [400, 1600] MHz, the antenna structure “Antenna as ‘P’” has a resonant frequency: $f_{r1} = 1060$ MHz.

Note that the values measured resonance frequencies are very close to those simulated. The difference between them is due to measurement uncertainties. It was found that the value of reflection coefficients for the resonance frequencies is very different from that which was obtained by simulation, this may be due to losses introduced by the dielectric, the mismatch between the source and antenna. Losses associated with the driver to the substrate. In fact, the bandwidth Δf is related to the quality factor of the antenna Q which is proportional to the factors Qd (the dielectric loss) and Qc is the and Qc the loss of driver.

4. CONCLUSION

The printed antenna geometries are simple and require further study by different methods developments. Ainsy, many travaux research on the design and simulation of patch antennas have been made. By following these steps, we designed several antennas while varying several parameters: the length of the antenna is connected to the resonant frequency, width, shape, thickness of the substrate, and the relative permittivity ϵ_r . For this a detailed study of the variation of these parameters and its influence on the resonant frequency is achieved in this course.

The antenna design using the software HFSS, “Ansoft High-Frequency Structure Simulator” and ADS ‘Advanced Design System’ is based mainly on the change in the shape of the antenna and its conducting material, the nature and thickness of the substrate to have a structure that resonates in the frequency desired for applications specific. The optimal variation of each of these parameters affects the resonance frequency, the reflection coefficient and the geometrical structure of the antenna patch. In conclusion, the integration of these planar structures may lead to the design of a communication system in a frequency range well defined.

REFERENCES

1. Gaha, H., F. Choubani, J. David, and A. Bouallegue, “Intrinsic correlated parameters for modified sierpinski multi-band antenna design,” *4th Conférence Internationale, JTEA’2006*, Tunisie-12-13-14-Mai, 2006.
2. Bri, S., A. Saadi, M. Habibi, A. Nakheli, L. Zenkouar, L. Bellarbi, and A. Mamouni, “Finite difference time domain (FDTD) analysis of new applicators for hyperthermia and evaluation of the SAR distribution in humain head near cellular phone,” Vol. 69, No. 2, ANSE, France.
3. Eldek, A. A., A. Z. Elsherbini, and C. E. Smith, “Square slot antenna for dual wideband wireless communication systems,” *Journal of Electromagnetic Waves and Applications*, Vol. 19, No. 12, 1571–1581, USA, 2005.
4. Almeida, J. F., C. L. D. S. S. Sobrinho, and R. O. dos Santos, “Analysis by FDTD method of a microstrip antenna with PBG considering the substrate thickness variation,” *Journal of Microwave and Optoelectronics*, Vol. 3, No. 3, 41–48, December 2003.

5. Bri, S., K. Elknani, and A. Nakheli, "Patch antenna at frequency $f = 2.35$ GHz for telecommunications applications," *Progress in Electromagnetics Research Symposium*, Prague, Czech Republic, August 27–30, 2007.
6. Çakir, G. and L. Sevgi, "Design, simulation and tests of a low-cost microstrip patch antenna arrays for the wireless communication," *Turk. J. Elec. Engin.*, Vol. 13, No. 1, 2005.
7. Rmili, H., J.-L. Miane, T. Olinga, and H. Zangar, "Design of microstrip-fed proximity-coupled conducting-polymer patch antenna," *SETIT 2005 3rd International Conference: Sciences of Electronic, Technologies of Information and Telecommunications*, Tunisia, March 27–31, 2005.
8. Ibrahim Gaha, H. B., "Analyse et conception des antennes fractales applications aux Télécommunications large bande," *Thèse de Doctorat Cotutelle No. 3068 du 27 Avril 2004 entre l'Université INPT-ENSEEIH de Toulouse & UTM-ENI de Tunis*, Tunis, Juillet 2007.

Light Propagation in Micro-optical-lattice Waveguide

Xiaofei Chen, Yali Qin, Hongliang Ren, and Fei Liu

Zhejiang Key Research Lab of Fiber-optic Communication Technology

College of Information Engineering

Zhejiang University of Technology, HangZhou 310023, China

Abstract— This paper have studied a coherent beam evolvment in an Micro-Optical-Lattice (MOL) waveguide. The waveguide here have two different structures. One is straight and the other is tilt with a small angle $\alpha = \arctan(1/600)$. The result shows that the MOL-waveguide can localize the beam's energy even under low power and when the waveguide have a small angle, it can guide the transmission direction of the beam. The light can change its direction under the guidance of the waveguide, however, it needs a certain distance before the beam can totally shift its energy to the tilt waveguide perfectly.

1. INTRODUCTION

Photorefractive crystals play a very important role in all optical communication because of its low-power, high nonlinear index and good modulability, especially bringing convenience in experience and design for the optical devices, such as optical switches, optical waveguide and so on. Many observations have been done to study the characteristic of the beam propagation in the optical lattice in the past decade [1–4]. The evolvment of the beam in nonlinear photonic crystal (PC) has been paid more attention over past decade, and the refractive index of the PC is periodically modulated by pairs of coherent beam called array lights, that is optical lattice. It has many novel characteristics [5], and many studies have been done as the light propagation along the lattice [5, 6].

However, in this paper, the light is transmitting across the lattice as showed in the Fig. 1. The lattice here is some different from the ones described in other references, such as [1]. As the spatial period is much smaller, and is comparable with the wavelength (so we call it Micro-Optical-Lattice). This structure has some new good properties. It will not make the light discrete while it has enough nonlinear to balance the diffusion and bring us the soliton. As the array light can be modulated and controlled, this structure is very important for us to control the signal light by another controllable beam.

2. THEORY AND MATHEMATICAL MODEL

As we know, the most simple form of electromagnetic waves transmitting in the periodic structure is Floquet-Bloch (FB) waves, and the most simple form of electromagnetic waves transmitting in the block structure is plane waves, the former is similar to the latter [7]. According to the FB theory, a FB wave can be described by addition of a set of plane waves which are of the same k_z , and k_z is the wave vector of the beam. These plane waves are called space harmonics. The FB theory is

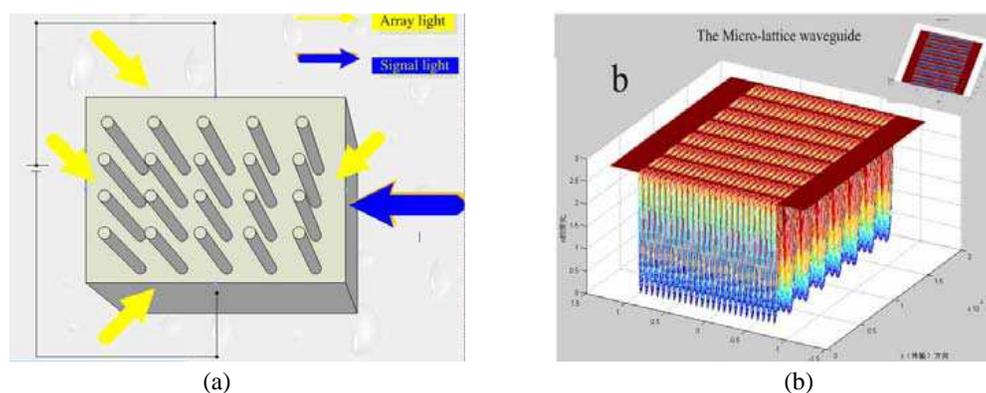


Figure 1: This is the structure of the lattices. (a) shows the experiment model, the lattice is produced by the array light described by the yellow arrow. And the signal represented by the blue arrow is incident from the side. (b) is the real refractive index distribution in the simulation calculation. The waveguide showed in picture is in straight.

widely used for analyzing and describing the kinds of phenomena when the light transmitting in the periodic lattices such as reflection, refraction, scattering, interference and diffraction, etc.

The beam propagation in optical lattices can be described by Nonlinear Schrödinger Equation (NLS) which is derived from Helmholtz Equation under the condition of paraxial approximation [8],

$$i2k \frac{\partial \phi}{\partial z} + \frac{\partial^2 \phi}{\partial y^2} + \frac{\partial^2 \phi}{\partial x^2} - k_0^2 n_e^4 \gamma_{33} E_{sc} \phi = 0 \quad (1)$$

where the ϕ is the envelope of the incident beam, and E_{sc} is the space-charge field, $E_{sc} = E_0/(1+I)$, expressed through total intensity I , normalized with respect to the dark irradiance of the crystal I_d . E_0 is extra bias field. Here, $k = k_0 n_e'$ and $n_e'^2 = n_e^2 - n_e^4 \gamma_{33} E_{sc}$, $k_0 = 2\pi/\lambda_0$, λ_0 is the wavelength of the light in vacuum, n_e is refractive index of the extraordinary light. While n_e' is the refractive index in the crystal material. γ_{33} is the electro-optical coefficient of the crystal. Set $\xi = z/kx_0^2$, $\zeta = x/x_0$, $\varsigma = y/x_0$, we can get the Normalized NLS Equation

$$i \frac{\partial \phi}{\partial \xi} + \frac{1}{2} \left(\frac{\partial^2 \phi}{\partial \varsigma^2} + \frac{\partial^2 \phi}{\partial \zeta^2} \right) - N \phi = 0 \quad (2)$$

where $N = k_0^2 n_e^4 \gamma_{33} x_0^2 E_{sc} / 2$. The Space-charge field (E_{sc}) is produced by the total intensity of the array lights and the signal light. As both Bloch waves and gap solitons are stationary solutions of Eq. (2), which are of the form $E = \phi(x, y, z) e^{ikz}$ [7, 9], the beam transmitting in the MOL waveguide can also be described by the NLS Equation. For convenience, we still use x, y, z to replace ζ, ς, ξ , so in the following result, the unit of x and y is x_0 , and the unit of z is kx_0^2 . The experiment structure model of the simulation is shown as Fig. 1. The lattice's dimension is comparable with beam's wavelength, The signal light is transmitting across the lattices.

3. SIMULATION AND RESULT

The calculation for the NLS Equation is using Alternating-Direction-Implicit-Beam-Propagation-Method (ADI-BPM) [10, 11]. To study the performance of the MOL-waveguide, the simulation is done in two different lattice structure: One is in line (showed in Fig. 1(b)) and the other is of a gentle turning with the angle $\alpha = \arctan(1/600)$ (this value is before the x, y and z normalized), Fig. 1(b) is the refractive index of the micro-optical-lattice waveguide. It can be produced by two pairs of coherent beams of which the wavelength is 488 nm. Here the distribution of integrated intensity of the array lights is described by $\cos(k_1 \sin(\theta) k_1 x_0^2 z)^4 \cdot \cos(k_1 \sin(\theta) x_0 x)^4$, where $\theta = \pi/150$, $k_1 = k_0^* n_e$. So the spatial period of the lattice d is about $d = 1 \mu\text{m}$, (so we call it MOL waveguide). The width of the waveguide is $12 \mu\text{m}$. The x and z are normalized by x_0 and $k_1 x_0^2$ respectively. The beam propagation along the z axis. The material used here is SNB: 75 with the parameters $\gamma_{33} = 1340 \text{ pv/m}$, $n_e = 2.3$, the wavelength of the incident beam is $\lambda = 488 \text{ nm}$.

3.1. Light in the Straight MOL Waveguide

The incident light is Gauss-like as described by the normalized expression $f = 2 * \exp(-x^2 - y^2)$, and external voltage $E_0 = 1000 \text{ v/cm}$. The result is displayed in the Fig. 2. The first picture in

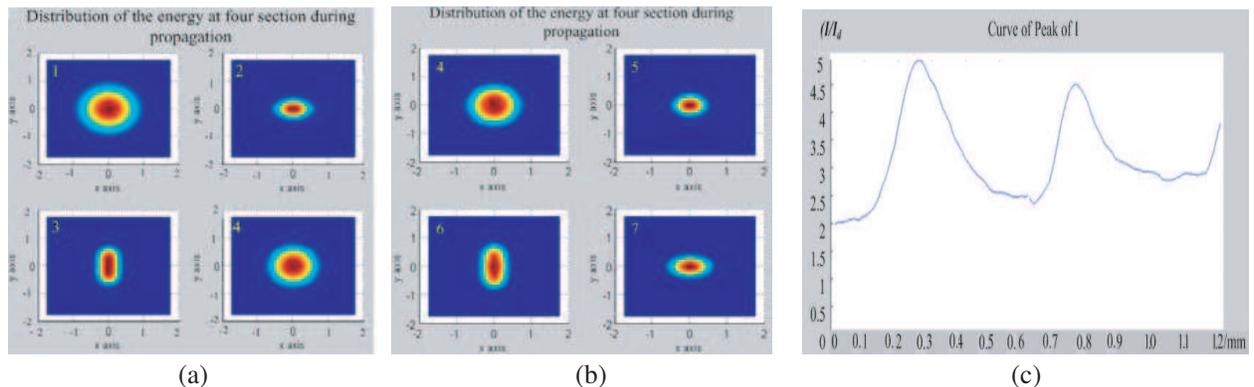


Figure 2: (a) and (b) show the distribution of the energy for every 0.2 units propagation and the first in (a) is of the incident. $f = 2 * \exp(-x^2 - y^2)$, $E_0 = 1000 \text{ v/cm}$. (c) is the curve of the energy peak value along the 1.2 units long transmission.

Fig. 2(a) is the energy distribution of the incident light, these pictures show the distribution of the energy at seven different section plane along z axis, for every 0.2 units (this value is normalized, and the following result is the same), here two pictures numbered 4th are the same as we use the final output of the first simulation as the second one's incident beam. There are two obvious phenomenon displayed in Fig. 2 that are different from the one propagating in the block material. First one, The energy first focused on y axis and then focused on x axis alternately as showed in the Figs. 2(a) and (b). and the whole changing are accompanied with the second one, the peak of the solitons energy changes periodically during 1.2-unit-long-propagation with the energy maintaining, which we can see clearly from Fig. 2(c), that is soliton.

In order to explain these phenomenon and get more characteristics, the incident energy is reduced to $f = 0.2 * \exp(-x^2 - y^2)$ with the other conditions unchanging. In fact, when the energy is reduced to this value, the beam can't be the soliton in an block material as there's not enough energy to be. However, the result showed in Fig. 3 is much different. The left picture shows the energy distribution during the dynamic propagation of 0.6 units in this structure. It seems that the energy evolvement on x is separately with it on y . In y direction, the beam breaks up into two after a short distance propagation (Fig. 3(a) 2) and then focus into one maintaining for a relative long distance (Fig. 3(a) 3–7), however, the beam finally turned to one peak after another short-distance-status of two (Fig. 3(a) 8, 9). In x direction, it seems that the beam behaves diffraction only all the time. However, when prolonging the transmission distance, the periodicity on x axis begins to show (Fig. 3(b)). Obviously, this result shows that beam's behavior in x and y direction can't be explained separately, and the behaviors affect each other mutually.

According to all the results above, when the beam transmits in the straight MOL waveguide, it shows different characteristics from the result in the block material. Although there is no periodicity on x -direction, the structure on y and z -direction is deeply affect the beam's behavior on this direction, and finally made it evolving periodically in this structure. As a result, the beam transmitting in it becomes soliton.

3.2. Light in the Tilt MOL Waveguide

In this part, the MOL waveguide has a small angle with the z axis (the direction in which the light is incident), first we set the angle $\alpha = \arctan(1/600)$, and $f = 2 * \exp(-x^2 - y^2)$, $E_0 = 1000$ v/cm. The result is displayed in the following Figure 4. The most different phenomenon compared with the one in the straight waveguide is the energy center shift forward to the plus y axis direction because of the waveguide. After 0.9-unit-propagation in the waveguide, the soliton has totally shifted its energy for 0.5 (normalized by x_0), that is 5 lattice's spatial period. The phenomenon is obviously by comparing the first picture with 9th in Fig. 4(a). With the $\alpha = \arctan(1/600)$, the distance fit the calculation result perfectly. Besides, there is a common characteristic between the two structures: The peak of the light energy changes periodically during the whole propagation

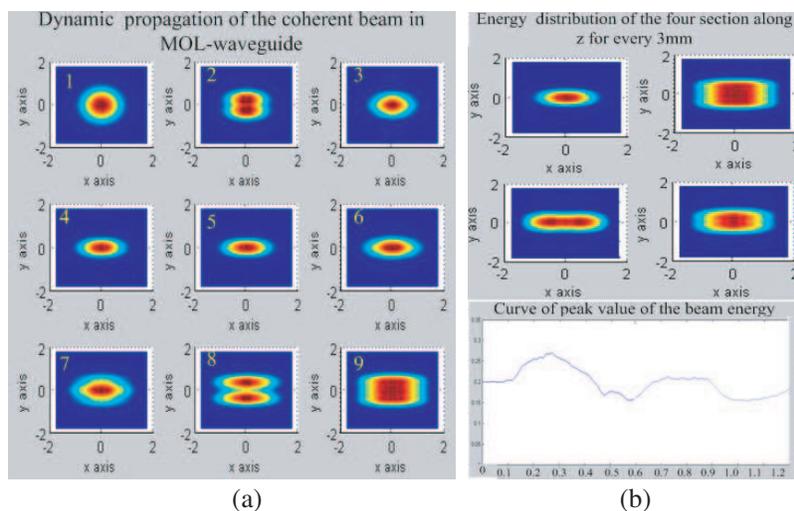


Figure 3: Shows the energy distribution of the beam for every 0.75 units. (b) is the distribution for every 0.3 units, the curve show the peak changing during the 1.2-unit-long-propagation. $f = 0.2 * \exp(-x^2 - y^2)$, $E_0 = 1000$ v/cm. the first one in (a) is the incident for both.

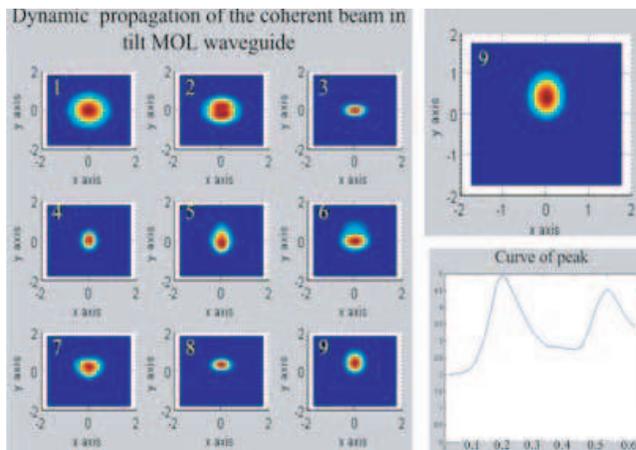


Figure 4: Shows the distribution of the energy for every 0.2 units and the first one is of the incident light with $f = 2 * \exp(-x^2 - y^2)$, $E_0 = 1000$ v/cm, $\alpha = \arctan(1/600)$. (b) is same of the 9th in a, and (c) is the curve of the energy peak value along the 0.6-unit-long transmission.

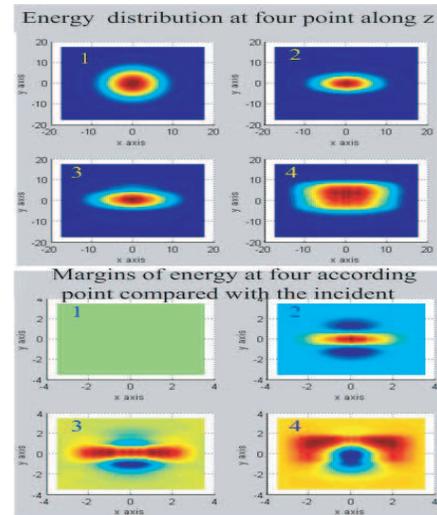


Figure 5: (a) shows the distribution of the energy at four sections for every 0.2 units and the first one is the incident. $f = 0.2 * \exp(-x^2 - y^2)$, $E_0 = 1000$ v/cm. (b) is margins of the same plane of the (a) compared with the incident light.

with the soliton compressed first on y and then on x alternatively.

Simulation is also done in this structure when $f = 0.2 * \exp(-x^2 - y^2)$ with others unchanged. The beam's evolution is displayed in Fig. 5. The upper one is the distributions of the energy at four section planes for every 0.2 units during 0.6-unit-long-propagation and the first one is the incident. In order to see more clearly about the energy shifting, the nether one is margins of the same section of the (a) compared with the incident light. It's obvious that the energy first focuses to the center of the beam and then shifts to plus y direction accompanied by diffraction on x axis during all the process within 0.6-unit-long-propagation.

4. CONCLUSION

This paper have studied a coherent beam evolution in two different Micro-optical-lattice waveguide structure. One is straight and the other is with an angle of $\alpha = \arctan(1/600)$. The result shows that the MOL-waveguide can localize the beam's energy even though the beam's energy is less then the threshold being soliton in the block with other condition being the same. And the most important, it can guide the transmission of the beam so that the light can change its direction. Besides, the beam transmitting in such structure presents a new phenomenon, it compresses energy in x and y direction alternatively with its whole energy maintaining. The light can change its direction under the guidance of MOL waveguide, however, it needs a certain distance before the beam can totally shift its energy to the tilt waveguide perfectly. More characteristics need to be studied about MOL-waveguide structure.

REFERENCES

1. Efremidis, N. K., J. Hudock, D. N. Christodoulides, J. W. Fleischer, O. Cohen, and M. Segev, "Two-dimensional optical lattice solitons," *Phys. Rev. Lett.*, Vol. 91, 213906, 2003.
2. Fleischer, J. W., M. Segev, N. K. Efremidis, and D. N. Christodoulides, "Observation of two-dimensional discrete solitons in optically induced nonlinear photonic lattices," *Nature*, Vol. 422, 147–150, 2003.
3. Martin, H., E. D. Eugenieva, Z. Chen, and D. N. Christodoulides, "Discrete solitons and soliton-induced dislocations in partially coherent photonic lattices," *Phys. Rev. Lett.*, Vol. 92, 123902, 2004.
4. Buljan, H., O. Cohen, J. W. Fleischer, T. Schwartz, M. Segev, Z. H. Musslimani, N. K. Efremidis, and D. N. Christodoulides, "Random-phase solitons in nonlinear periodic lattices," *Phys. Rev. Lett.*, Vol. 92, 223901, 2004.

5. Chen, Z., H. Martin, E. D. Eugenieva, and J. Xu, “Formation of discrete solitons in light-induced photonic lattices,” *Optics Express*, Vol. 13, No. 6, 2005.
6. Christodoulides, D. N., F. Lederer, and Y. Silberberg, “Discretizing light behavior in linear and nonlinear waveguide lattices,” *Nature*, Vol. 424, 2007.
7. Ma, C., “Wavevector diagrams of Floquet-Bloch waves in periodic dielectric structures,” *Acta Optica Sinica*, Vol. 14, No. 5, 1994.
8. Wang, H. and W. She, “Optical spatial solitons and their applications,” Master Thesis, China Zhongshan University.
9. Zhang, Y. and B. Wu, “Composition relation between gap solitons and blochwaves in nonlinear periodic systems,” *Phys. Rev. Lett.*, Vol. 102, 093905, 2009.
10. Zhang, Z., Q. M. Luo, S. Q. Zeng, X. Y. Zhang, and D. X. Huang, “Alternating direction implicit method for solving the two-dimensional time-dependent photon diffusion equation,” *Chinese Journal of Computational Physics*, Vol. 20, 359–362, 2003.
11. Tian, Z. F. and Y. B. Ge, “A fourth-order compact ADI method for solving two-dimensional unsteady convection-diffusion problems,” *Journal of Computational and Applied Mathematics*, Vol. 198, 268–286, 2007.

A Combined Cavity with Improved Performance under Simultaneous Resonance of Sub-cavities

Chih Jung Wu^{1,2}, Qiang Liu^{1,3,4}, Chung Ping Liu²,
Jong C. Wang², and Zhengbiao Ouyang^{1,3,4}

¹THz Technical Research Center, Shenzhen University, Shenzhen 518060, China

²Department of Electrical Engineering, Yuan Ze University, Taoyuan 320, Taiwan

³Shenzhen Key Lab of Micro-Nano-Photonics Information Technology
Shenzhen 518060, China

⁴College of Electronic Science and Technology, Shenzhen University
Shenzhen 518060, China

Abstract— We insert a dielectric rod in the point cavity or increase the volume of the point cavity in the combined system to shift the resonant frequency of the point cavity to be the same as that of the feedback cavity. These systems are advantageous over our previous design in that the feedback cavity is uniform and no stray scattering exists in the cavity. Simulations through finite-difference time-domain (FDTD) method demonstrate that, in double resonance, the maximum quality factor of the new structure is promoted by 34.7%, and the maximum localized field intensity in the point defect cavity is promoted by 314% over that in the earlier structure.

1. INTRODUCTION

Signal detecting unit is a basic part in electronic, electromagnetic, and optical systems for communications, radars, sensors, etc. It is very important to have high sensitivity for long distance communication, long range radars, biological signal detection, and cosmic signal hunting. Especially, at present, THz signals are usually very weak and how to promote the sensitivity of THz signal detection becomes a key problem for development and application of THz technology on which great attention has been paid in the last decade [1–10].

The idea of combined cavity with sub cavities in simultaneous resonance was presented by the authors in our group in 2007 in the hope of promoting the field intensity in a cavity, so that weak signals can be detected [8]. In the combined cavity, as shown in Fig. 1, when the point-defect cavity (PDC) and the waveguide resonator (WGR) are in simultaneous resonance, i.e., the PDC and WGR have the same resonance frequency and a matched field pattern, the field intensity in the PDC will be much higher than in a separate PDC. This is because the input field is first enhanced in the resonant WGR and further enhanced in the resonant PDC (also called as “double resonance”) as the wave in the WGR coupled to the PDC. This is useful for greatly promoting the detection sensitivity of waves, especially for THz waves, when one puts a wave sensing element in the PDC [8]. Moreover, for THz signal detections, the background radiation is out of resonance with the cavity and will have negligible influence on the detection of useful signals, and thus bulk and expensive cooling systems can be omitted.

The key point is to make signals to be in double resonance through signal feedback in a combined cavity consisting of two sub-cavities. However, a simple combined cavity would not meet the condition for double resonance. For this we tried to add a piece of dielectric material in the feedback cavity, as shown in Fig. 1, and useful results were obtained [8]. But we are not satisfied with them, and so we now do more studies on the combined cavity for higher sensitivity detection applications. We now add a dielectric rod in the point cavity or increase the volume of the point cavity in the combined system to shift the resonant frequency of the point cavity to be the same as that of the feedback cavity, as shown in Fig. 2. These systems are advantageous over our previous design in that the feedback cavity is uniform and no stray scattering exists in the cavity. Furthermore, in the system in Fig. 2 the resonance mode of the point cavity lies closer to the center of the bandgap of the photonic crystal in the system than that in Fig. 1, and thus has higher quality factors than that of Fig. 1. Simulations through FDTD method demonstrate that the systems in Fig. 2 can have higher quality factor and localized field intensity than that in the old structure shown in Fig. 1. Since Figs. 2(a) and (b) operate in a similar mechanism, we focus on the structure in Fig. 2(a) in this paper.

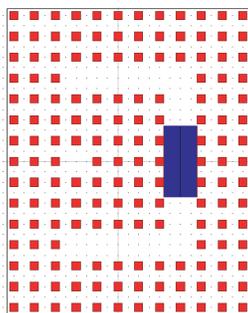


Figure 1: A combined cavity using a dielectric slab in the waveguide resonator to tune the resonance frequency.

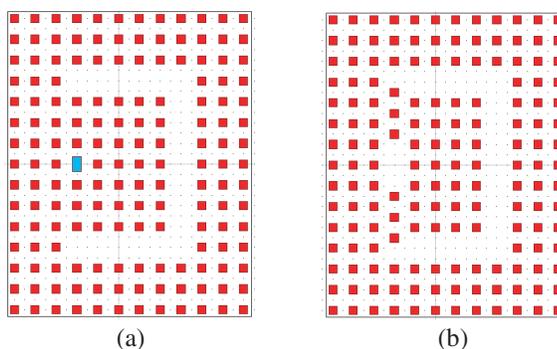


Figure 2: A combined cavity in which the resonance frequency is tuned by (a) adding a dielectric pole in the point defect cavity and (b) increasing the cavity volume.

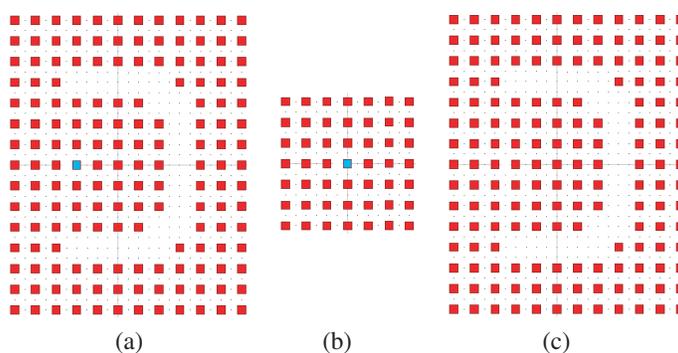


Figure 3: The structure of the combined cavity to be investigated (a) which is a combination of a PDC (b) and a WGR (c).

2. PHYSICAL MODEL AND DESCRIPTION OF SIMULATION PROCEDURE

For better performance, we modify the structure in Fig. 2(a) to be that in Fig. 3(a), where the red squares are dielectric rods, the background is air, and the blue rectangular indicates the defect rod. The combined cavity in Fig. 3(a) is a combination of the PDC in Fig. 3(b) and the WGR in Fig. 3(c).

The simulation procedure is as follows. First, we use plane wave expansion (PWE) method to determine the resonance wavelengths of the separate WGR, shown in Fig. 3(c), by treating the WGR as a super cell. A number of resonance modes can be found in the separate WGR in the bandgap region of the photonic crystal in the system. We use the PWE method to view the band structure. But we find that, for the resonant frequencies, much more accurate result with higher speed of calculation can be obtained by the FDTD method than by the PWE method. Second, we go to calculate the fundamental defect mode of the PDC, shown in Fig. 3(b), and set the resonance frequency to be the same as the WGR mode that lies near the bandgap center by proper choice of the parameters of the defect pole in the PDC. Third, with the properly chosen parameters of the defect pole in the PDC, we calculate the resonance modes of the combined cavity shown in Fig. 3(a). Fourth, we calculate the mode patterns of the resonance modes to find out the modes that are in resonance with the PDC and the WGR simultaneously. For simultaneous resonance, not only the resonance frequency, but also the field pattern of the PDC and the WGR should be matched. Fifth, we use FDTD method with perfect matched boundary (PML) conditions to calculate the quality factors of the simultaneous resonance modes in the combined cavity and the localized field intensity in the PDC at simultaneous resonance. Sixth, we compare the results with that of separate PDC and separate WGR and that of the early combined cavity shown in Fig. 1.

In the following simulations, for the convenience of comparison, the parameters of photonic crystal in the system are chosen to be the same as that in our early work [8], i.e., we consider a photonic crystal of square lattice consisting of square rods with a lattice period of a and, for the square rods, a side width of $0.4a$. Also as that in [8], the refractive index of the square rods is

$n_a = 3.4$, and the refractive index of the background air is $n_b = 1$. Furthermore, only TE-mode operation is considered. In TE-mode operation, the electric vector of the wave is perpendicular to the propagation route of waves in the waveguide and parallel to the axis of poles in the photonic crystals.

3. RESULT OF SIMULATIONS AND COMPARISON WITH THAT OF THE OLD STRUCTURE

3.1. Resonance Modes in the Separate WGR, the Separate PDC, and the Combined Cavity

Using the plane wave expansion method, we can find the band structure of the separate WGR indicated in Fig. 3(c) by treating the WGR as a super-cell. The result is shown in Fig. 4, which displays that there are 12 resonance modes indicated as horizontal lines in the bandgap region, including the bandgap edge. Accurate resonance frequencies are obtained by the FDTD method as shown in Table 1. These modes are also the defect modes in the structure since the WGR can be considered as a kind of special defect in a photonic crystal.

From Fig. 4, we see that the bandgap is from 0.2668 to 0.3882 (normalized frequency). Thus the center of the bandgap is 0.3275. As mentioned in Sec. 1, to get a separate PDC or the combined cavity with quality factors as high as possible, the parameters of the defect pole in the PDC should be so chosen that the normalized frequency of the defect mode in the PDC is near 0.3275. Through FDTD simulations and looking at the field pattern in the WGR, we find that the modes with even numbers in Table 1 are modes having inverse symmetry (which may be called as odd symmetry) about the central horizontal axis in the system, while the modes with odd numbers in Table 1 are modes having positive symmetry (which may be called as even symmetry) about that axis. Noting that the fundamental mode pattern in the PDC has positive symmetry about that axis, we set the resonant frequency of the PDC to be that of mode-7 in Table 1, i.e., 0.330543, which is realized by taking the refractive index of the defect pole to be $n_d = 3.13$ with its width being the same as the that of other poles in the photonic crystal and vertical height being $0.08a$. The center of the defect pole coincides with that the PDC.

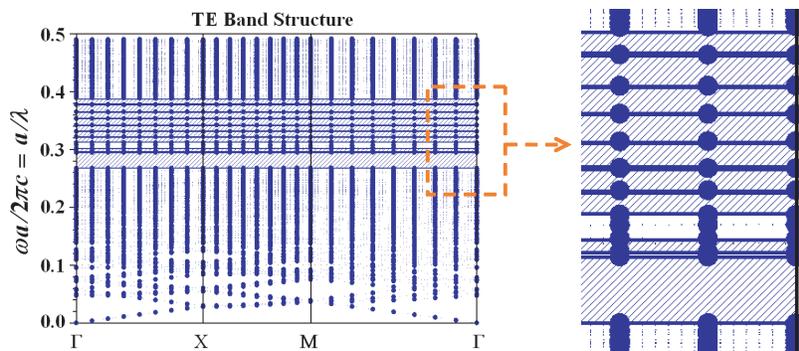


Figure 4: The band structure of the separate WGR.

Table 1: Resonance modes in the separate WGR.

Mode No.	1	2	3	4	5	6
a/λ	0.294018	0.295257	0.301647	0.306225	0.311375	0.320816
Mode No.	7	8	9	10	11	12
a/λ	0.330543	0.340939	0.352764	0.363731	0.377179	0.387764

Table 2: Resonance modes in the combined cavity.

Mode No.	1	2	3	4	5	6	7a
a/λ	0.294018	0.295257	0.301647	0.306225	0.311375	0.320816	0.329875
Mode No.	7b	8	9	10	11	12	
a/λ	0.331688	0.340939	0.352764	0.363731	0.377179	0.387764	

Table 3: Quality factors and localized field intensity in the PDC.

	Q_{SP}	Q_{SL}	Q_{SR}	E_{SP}^2	E_{SL}^2	E_{SR}^2
OLD	1156	2355	2766	687	7062	6420
NEW	1475	2934	3727	645	19997	36341

Then with the above parameters, the resonance modes in the combined cavity in Fig. 3(a) can be obtained by the FDTD method, as shown in Table 2, from which we see that there are 13 modes in the combined cavity. Comparing Tables 1 and 2 we see that mode-7 in Table 1 is split up into two modes — mode-7a and mode-7b in Table 2. The coupling is strong because the two cavities are directly connected and the strong coupling caused the splitting of the resonance mode, which is similar to that in coupled electronic resonance circuits. From Tables 1 and 2, we also see that, except for mode 7a and 7b, other modes in the combined cavity are the same as that in the separate WGR. This is understandable because the PDC has no influence on these modes in the WGR.

3.2. Quality Factors of the Combined Cavity and Localized Field Intensity in the PDC in the Combined Cavity and Comparison with That of the Separate WGR, the Separate PDC, and the Old Combined Cavity

Since the field in a cavity decays exponentially when the exciting source is suddenly removed, by applying an impulse and keeping the profile of the decaying field in the system through FDTD simulations, quality factors can be calculated from the profile of the decaying field. Localized field intensity in the separate PDC or the PDC in the combined cavity can be calculated also through FDTD simulations by putting a unit source outside the PDC. Indicate $Q_{SP}(E_{SP}^2)$, $Q_{SL}(E_{SL}^2)$, $Q_{SR}(E_{SR}^2)$ the quality factors (localized field intensity) of the separate PDC, the left ($a/\lambda = 0.329875$ for the new combined cavity, 0.364193 for the old combined cavity) and right ($a/\lambda = 0.331688$ for the new combined cavity, 0.365604 for the old combined cavity) modes of the combined cavity in double resonance, respectively. For the same parameters in the above sections, we may obtain the results as shown in Table 3. The localized field intensities are obtained for a source outside the PDC.

From Table 3, we can see that, in double resonance, the maximum quality factor of the new combined cavity is promoted by 34.7%, and the maximum localized field intensity in the PDC in the combined cavity is promoted by 314% over that in the old combined cavity. Thus for signal-detection applications, with the improved structure, the detecting sensitivity can be further promoted by three times when a THz wave sensor is inserted in the PDC in the combined cavity.

4. CONCLUSIONS

Through adding a dielectric pole in the PDC, in double resonance, the maximum quality factor of the new structure is promoted by 34.7%, and the maximum localized field intensity in the point defect cavity is promoted by 314% over that in the earlier structure. Thus the performance of the combined cavity is greatly enhanced.

ACKNOWLEDGMENT

This work was supported by the Chinese Natural Science Foundation (Grant No. 60877034), the Guangdong Natural Science Foundation (Key Project, No. 8251806001000004), and the Shenzhen Science Project (No. 200720, 200805).

REFERENCES

1. Mizuno, K., R. Kuwahara, and S. Ono, "Submillimeter detection using a Schottky diode with a longwire antenna," *Appl. Phys. Lett.*, Vol. 26, No. 11, 605–607, 1975.
2. Wu, Q., M. Litz, and X. C. Zhang, "Broadband detection capability of ZnTe electro-optic field detectors," *Appl. Phys. Lett.*, Vol. 68, No. 21, 2924–2926, 1996.
3. Klappenberger, F., A. A. Ignatov, S. Winnerl, et al., "Broadband semiconductor superlattice detector for THz radiation," *Appl. Phys. Lett.*, Vol. 78, No. 12, 1673–1675, 2001.
4. Esaev, D. G., S. G. Matsik, M. B. M. Rinzan, et al., "Resonant cavity enhancement in heterojunction GaAs/AlGaAs terahertz detectors," *J. Appl. Phys.*, Vol. 4, No. 93, 1879–1883, 2003.

5. Liu, H. C., C. Y. Song, A. J. Spring Thorpe, et al., “Terahertz quantum-well photodetector,” *Appl. Phys. Lett.*, Vol. 84, No. 20, 4068–4070, 2004.
6. Kawayama, I., T. Miyadera, H. Murakami, et al., “A tunable sub-terahertz wave generation and detection system with a photomixer and a high-Tc Josephson junction,” *Supercond. Sci. Technol.*, Vol. 5, No. 19, S403–S406, 2006.
7. Yellampalle, B., K. Y. Kim, G. Rodriguez, J. H. Glowina, and A. J. Taylor, “Details of electro-optic terahertz detection with a chirped probe pulse,” *Optics Express*, Vol. 15, No. 3, 1376–1383, 2007.
8. Ouyang, Z., X. Luo, J. C. Wang, C. P. Liu, and C. J. Wu, “A combined cavity for high sensitivity THz signal detection,” *Proc. of SPIE*, Vol. 6840, No. 684008, 2007.
9. Khan, M. J., J. C. Chen, and S. Kaushik, “Optical detection of terahertz using nonlinear parametric upconversion,” *Opt. Lett.*, Vol. 33, No. 23, 2725–2727, 2008.
10. Scheller, M. and M. Koch, “Terahertz quasi time domain spectroscopy,” *Optics Express*, Vol. 17, No. 20, 17723–17733, 2009.

Static Magnetic Field Interferes with the Physiological Removal of Circulating Apoptotic Lymphocytes

Luciana Dini

Department of Biological and Environmental Science and Technology, University of Salento
via per Monteroni snc, Lecce 73100, Italy

Abstract— In various cellular systems, including cells of the immune system, a number of biological effects induced by static magnetic fields (SMFs) have been reported and different mechanisms have been proposed to explain these effects. Although a number of theoretical models have been proposed, the variety of experimental conditions (intensity, frequency and time windows of the fields, differing characteristics of the materials used- cell type, age, treatment) makes contradictory the data present in the literature and the possibility of the replication of the experiments. However, (S)MFs have been reported to perturb distribution of membrane proteins and sugars, cytoskeleton and trans-membrane fluxes of different ions, especially calcium $[Ca^{2+}]_i$. In turn, these alterations could interfere with specific physiological activities, like phagocytosis, which are based on receptors, cytoskeleton elements and motor proteins. In a previous work we found that, sinusoidal liver cells quickly recognized and engulfed human lymphocytes exposed for up to 72 h to 6 mT SMF, by using the same receptors that mediate the clearance of apoptotic cells. Thus, aim of the present work has been to decipher the modifications exerted by the SMF on lymphocytes and/or on the process of phagocytose. We analysed the cell surfaces of normal and apoptotic human lymphocytes in the presence or absence of 6 mT SMF by immunocytochemistry and biochemistry assays. SMF increases, in a time-dependent way, the expression of GD3 ganglioside and cholesterol on the plasma membrane of normal lymphocytes and prevents GD3 removal on apoptotic-induced cells. Lipid peroxidation of plasma membrane was observed soon after lymphocytes induction of apoptosis and after 72 h of SMF exposure. The recognition and the engulfment of the control and apoptotic lymphocytes is modified by SMF exposure. In fact, normal exposed lymphocytes are recognized by the liver sinusoids at the same extent of the apoptotic non exposed cells. Conversely, the exposure to SMF promoted the binding but delayed the engulfment of apoptotic lymphocytes in *in situ* as well as in *in vitro* phagocytose assays. Further studies will clarify the eventual implication of SMF on human health.

1. INTRODUCTION

The development of modern society has been accompanied by a dramatic increase in the number of electronic devices, that, as consequence, have increased the exposure to static and/or electric fields of humans. An increasing bulk of evidence indicates that SMFs influence tissues and cells [1–3]. Indeed, in various cellular systems, including cells of the immune system, a number of biological effects have been reported and different mechanisms have been proposed to explain these effects. In our and other author' studies, it was reported that the cellular and molecular modifications induced when magnetic fields interact with biological material depend on the duration of exposure, intensity, tissue penetration and the type of cells [1]. However, difficulties remain in resolving the contradictory results that arise from the multiplicity of experimental conditions [2, 3]. The plasma membrane is considered a primary site of SMF action. Thus, it is likely that, SMFs influence the diamagnetic properties of the plasma membrane that in turn distorts embedded ion channels (i.e., calcium ions channel) to the point of altering their function [3]. MFs affect the rotation of the membrane's phospholipids and change protein distribution. Altogether, these alterations strongly interfere with physiological activities, linked to phagocytic mechanisms, like the recognition and clearance of circulating apoptotic cells. The ultimate and most favourable fate of almost all dying cells is engulfment by neighbouring or specialized cells. Efficient clearance of cells undergoing apoptotic death is crucial for normal tissue homeostasis and for the modulation of immune responses [4, 5]. Engulfment of apoptotic cells is finely regulated by a highly redundant system of receptors and bridging molecules on phagocytic cells that detect molecules specific for dying cells. The clearance of dying cells is an important fundamental process serving multiple functions in the regulation of normal tissue turnover and homeostasis. Aim of the present work has been to decipher the modifications exerted by 6 mT SMF on the process of clearance of apoptotic lymphocytes.

2. MATERIALS AND METHODS

Induction of lymphocytes apoptosis and exposure to SMF — Human lymphocytes were obtained after Ficoll gradient separation of buffy coats from blood donations of non-smoker healthy males, aged 25–50 as reported in [2]. Apoptosis was induced with 10^{-2} M cycloheximide (CHX) for 18 h. SMF was produced by Neodymium magnetic disks (10 mm in diameter and 5 mm in height) of known intensity supplied by Calamit Ltd (Milano, Italy) placed under the culture Petri dishes. The intensity of the field generated by the magnet was checked by means of a gaussmeter with a range of operating temperature of 0°C to 50°C and an accuracy (at 20°C) of $\pm 1\%$ (Hall-effect gaussmeter, GM04 Hirst Magnetic Instruments Ltd, UK). The laboratory areas between incubators, worktops and tissue culture hood measured $0.08\ \mu\text{T}$ to $0.14\ \mu\text{T}$ (50 Hz) magnetic fields. In the room the background flux density was $10\ \mu\text{T}$ (static) and the local geomagnetic field was approximately $43\ \mu\text{T}$ (for exposure details see [12]).

Lymphocyte characterization — Normal, apoptotic and SMF exposed lymphocytes (up to 72 h) were studied for their cell surface modifications by cytochemistry and immunocytochemistry of exposed saccharides (FITC conjugates lectins: $40\ \mu\text{g}/\text{ml}$ Concavalin-A, D-mannose/D-glucosamine and $2\ \mu\text{g}/\text{ml}$ Ricinus Communis, D-galactose/D-galactosamine), cholesterol ($0.05\ \text{mg}/\text{ml}$ Filipin), GD3 (Ganglyoside-Precursor-Disialohematoside) ganglyoside (monoclonal Ab anti-GD3 1:100). Samples were observed under a fluorescent microscope, Nikon 80i. The thiobarbituric acid (TBA) test was used to evaluate the presence of malonildialdehyde (MDA).

Phagocytosis test — The adhesion and internalization of normal, apoptotic and SMF exposed lymphocytes to liver sinusoidal cells was quantified after *in situ* injection of 1×10^6 Hoechst 33342 labeled normal and apoptotic lymphocytes, in the absence or in the presence of 6 mT SMF. The cells concentration represent a saturating concentration as evaluated by the presence of cells in the medium collected from the liver. Livers were perfused in a non-recirculating system at a flow rate of 1 ml/min at ice temperature for adhesion measurement and at 37°C for internalization evaluation. The adhesion specificity was tested in parallel inhibition experiments by adding appropriate inhibitors into the perfusion tube before adding apoptotic lymphocytes. Macrophages (3 days TPA-differentiated THP-1 cells) were co-incubated up to 6 h with normal and apoptotic exposed or not exposed lymphocytes, in a ratio of 10 apoptotic cells per phagocyte. Internalization of lymphocytes was allowed in the absence and in the presence of 6 mT SMF.

3. RESULTS

3.1. SMF Modifies the Exposure of Cell Surface Molecules on Control and Apoptotic Lymphocytes

Lymphocytes viability was more than 95% at the end of the isolation procedure and it was reduced to 30% when cells were induced to apoptosis by 18 h incubation with 10^{-2} M CHX. SMF exposure up to 24 h did not reduce cell viability or growth. SMF was able to reduce the yield of apoptotic cells (15–20% less) after CHX induction. Both the apoptotic treatment and the exposure to 6 mT SMF affect sugar expression as quantity and surface distribution; mannose/D-glucosamine and galactose/D-galactosamine residues were enhanced on the external surface of the plasma membrane (Tab. 1).

Table 1: Semiquantitative evaluation of FITC-conjugate lectins for the detection of carbohydrate residues on the cell surface of control (CTRL) and apoptotic (APO) lymphocytes in the absence and in the presence of 6 mT SMF.

Lectins	CTRL	CTRL+SMF	APO	APO+SMF
<i>Con-A</i>	–	+	++	+
<i>Ricinus c.</i>	–	++	+++	++

CTRL = normal lymphocytes; APO = apoptotic lymphocytes, 18 h

To verify if the SMF interferes with the plasma membrane glycoproteins distribution through the lipid alteration the measurements of thiobarbituric acid reactive substances (TBARS), the localization of GD3 ganglyoside and cholesterol have been performed. The data are reported in Figs. 1, 2. An extremely high increase of lipid peroxidation was measured (Fig. 1). GD3 ganglyoside, that is randomly distributed on the cell surface of control lymphocytes, was clustered at one pole

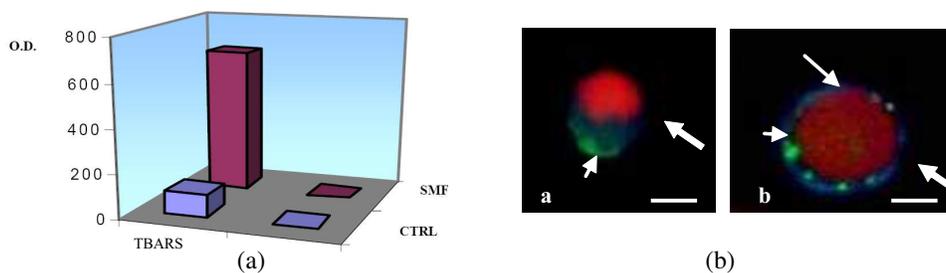


Figure 1: (a) Level of thiobarbituric acid reactive substances (TBARS) in control and 6 mT SMF exposed human lymphocytes. O.D. optical density at 233 nm. (b) Fluorescence images of triple staining of nucleus (red, Hoechst 33342), cholesterol (blue, Filipin) and GD3 ganglioside (green, primary Ab anti GD3, secondary Ab FITC conjugated) of 6 mT SMF exposed (a') and control (b') human lymphocytes. Bars = 5 μm. Arrowhead = GD3, thick arrow = cholesterol; thin arrow = nucleus.

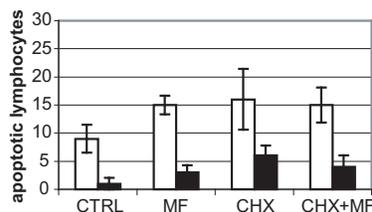


Figure 2: Number of control or 10⁻² M CHX treated lymphocytes in the absence or in the presence of 6 mT SMF for 24h, adhering (white columns) or internalized (black columns) by liver sinusoidal cells. Statistical analyses were performed using Student's *t*-test for unpaired data, and *P* values < 0.05 versus CTRL were considered significant. Data are presented as mean ± S.D..

of the cells after induction of apoptosis or after 48 h of exposure to 6 mT SMF (Fig. 2(b)). A comparable behavior was observed for cholesterol distribution (Fig. 2(b)).

3.2. The Recognition and Engulfment of Apoptotic Lymphocytes is Modified under SMF

Phagocytosis of the apoptotic cells is the last, but very important, step of the apoptotic program. *In vivo* apoptotic cells are cleared by phagocytosis with a rapid and efficient process able to remove damaged cells and components by inhibiting inflammation and by modulating the immune system [5, 6].

During phagocytosis of apoptotic cells plasma membrane has a pivotal role for the recognition of dead cells and for their engulfment through connection with cytoskeleton [7]. The phagocytosis of apoptotic cells is a multi-steps process. Apoptotic cells are first recognized by phagocytes, to which they adhere; then apoptotic cells are internalized and finally degraded. These mechanisms assure an efficient clearance of dead cells in non-pathological conditions. The data obtained in *in situ* and in *in vitro* experiments indicate that SMF interfere with both the steps of recognition and of engulfment of the apoptotic lymphocytes (Figs. 2 and 3). Apoptotic lymphocytes were retained by sinusoidal liver cells at high degree, and when internalization was allowed, they were found inside the cells. Worth noting, control non exposed lymphocytes were never phagocytosed. However, when normal lymphocytes were exposed for at least 24 h to SMF, they were retained by sinusoidal walls (three times with respect to control) (Fig. 2) and internalized when allowed.

The percentage of lymphocytes induced to apoptosis under exposure to SMF bound to sinusoidal liver cells was higher than the apoptotic ones. Correspondingly the number of internalized apoptotic lymphocytes was increased. The induction of apoptosis and SMF exposure had a synergic effect.

Time course (1, 2, 3, 4 h) of *in vitro* phagocytosis of apoptotic cells is reported in Fig. 3. Phagocytosis is progressively achieved with time of TPA-induction (i.e., from 1 to 3 days): 2% of THP-1 cells at 1 day, 25% at 2 days and more than 50% at 3 days. Only differentiated THP-1 cells could bind and internalize apoptotic cells. Non-differentiated THP-1 cells could only bind apoptotic cells at very low rate (5–7%) but not engulf them. Surprisingly, at the end of 4 h of phagocytosis, in the presence of SMF a very small percentage of non-differentiated THP-1 cells bore phagosomes containing apoptotic cells. In presence of SMF the percentage (about 25%) of non-differentiated THP-1 cells able to bind apoptotic cells increased. Exposure of 3 day-differentiated THP-1 cells to SMF during phagocytosis promoted engulfment of apoptotic cells preventing their binding (Fig. 3).

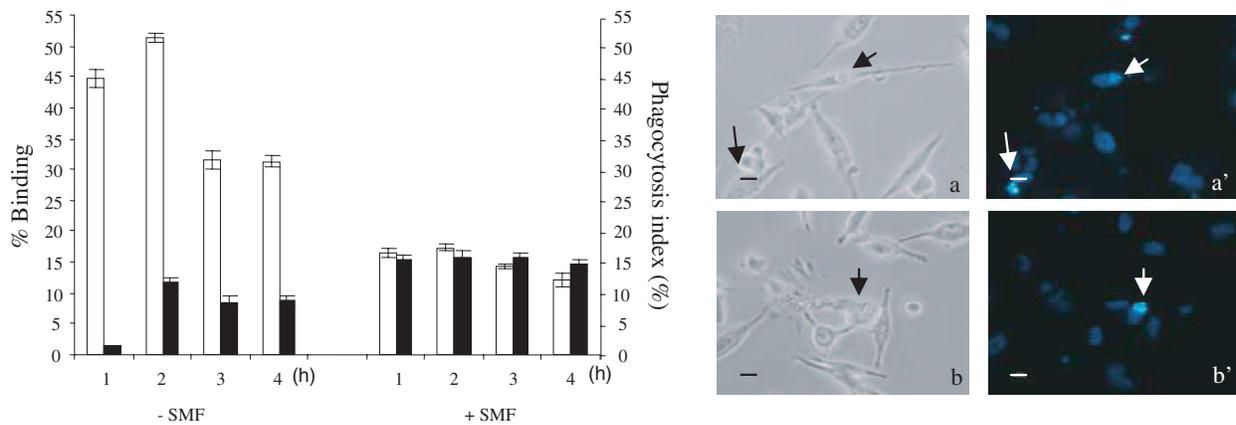


Figure 3: Percentage of the phagocytosis index (black columns) and of binding (white columns) of THP-1 cells, differentiated with TPA 50 ng/mL for 3 days in the presence (+SMF) and absence (–SMF) of 6 mT SMF. The values are the mean \pm SD of three independent experiments. The values –SMF are significantly different ($p < 0.05$) with respect to +SMF. The phagocytosis index and percentage of binding was measured scoring at least 500 cells for each experimental time. Images show contrast phase (left) and fluorescent (right) micrographs showing THP-1 cells at the third day of differentiation with TPA, internalizing (arrows) Hoechst 33342 stained apoptotic U937 cells after 4 h of phagocytosis assay in the absence (a), (a') and in the presence (b), (b') of 6 mT SMF. Bars = 10 μ m.

4. CONCLUSIONS

Our data support the bioeffects of the exposure to 6 mT SMF on the phagocytosis of human normal and/or apoptotic lymphocytes. SMF induced modifications of the plasma membrane of normal exposed lymphocytes, making these cells much more similar to the apoptotic ones. As a consequence, they are fast and silently removed by phagocytes. We showed the strong and reproducible effects that the exposure to SMF exerts on the process of liver phagocytosis of apoptotic lymphocytes. Liver clearance of apoptotic lymphocytes is a complex mechanism that can be influenced at least at three different levels: phagocyte receptors, apoptotic cell quality and SMF exposure. On the side of the phagocyte receptors, in general, the efficiency of phagocytosis is proportional to the number of expressed receptors. In the liver this aspect is particular evident, since the receptors involved in the recognition of apoptotic cells are modulated in relation to the type of the cells, to the physiological or pathological status of the organ and to the cell localization inside the lobule [7].

The extent of the SMF influence is also related to the degree of macrophage maturation. Since 6 mT SMF exposure interferes with monocyte/macrophage TPA-induced differentiation of U937 promonocytes and THP-1 monocytes (respectively: 20% increment and 15% decrement) [8], the delayed macrophage maturation could explain the defects in phagocytosis. Impairment of phagocytosis could also be due to the alteration of cytoskeleton following SMF exposure. Ruffling of plasma membrane, that subsequently close to form macropinosomes, requires reorganization of the actin filament network to the cell periphery, that has been reported to be altered by SMF exposure [9]. Microtubules polymerization is under the control of concentration of Ca^{2+} , that are modulated under SMF exposure. Ca^{2+} are also very important for signals transduction during TPA differentiation and cytoskeleton modifications during phagocytosis. Ca^{2+} intracellular concentration in U937 cells increases under exposure to SMF as well as during TPA treatment also in the presence of SMF [10, 11]. The modulation of the index and rate of phagocytosis of apoptotic cells could further be influenced by modifications of cell surface molecules exerted by SMF exposure. These modifications, involving a large set of molecules, can, in turn, affect related functions like cell-cell recognition, interaction and attachment; all activities that are very important during phagocytosis of apoptotic cells.

Altogether, our data gave evidence of a strong influence of SMF on phagocytosis. The implications for human health needs to be deeply investigated, because the pathological and/or physiological implications are not yet understood and need to be further investigated. In particular, further studies will clarify if SMF exposure could be used as a novel therapeutic tool to produce “apoptotic dressed cancer cells” and thus promote their elimination.

ACKNOWLEDGMENT

This work was financed by MIUR grant (ex 60%).

REFERENCES

1. Tenuzzo, B., A. Chionna, E. Panzarini, R. Lanubila, P. Tarantino, B. Di jeso, M. Dwikat, and L. Dini, "Biological effects of 6 mT static magnetic fields: A comparative study in different cell types," *Bioelectromagnetics*, Vol. 27, No. 7, 560–77, 2006.
2. Dini, L. and L. Abbro, "Bioeffects of moderate-intensity static magnetic fields," *Micron*, Vol. 36, 195–217, 2005.
3. Rosen, A. D., "Mechanism of action of moderate-intensity static magnetic fields on biological systems," *Cell Biochem. Biophys.*, Vol. 39, 163–173, 2003.
4. Krysko, D. V., K. D'Herde, and P. Vandenabeele, "Clearance of apoptotic and necrotic cells and its immunological consequences," *Apoptosis*, Vol. 11, 1709–1726, 2006.
5. Fadok, V. A. and G. Chimini, "The phagocytosis of apoptotic cells," *Semin. Immunol.*, Vol. 13, No. 6, 365–72, 2001.
6. Michlewska, S., A. McColl, A. G. Rossi, I. L. Megson, and J. Dransfield, "Clearance of dying cells and autoimmunity," *Autoimmunity*, Vol. 40, No. 4, 267–73, 2007.
7. Dini, L., P. Pagliara, and E. C. Carlà, "Phagocytosis of apoptotic cells by liver: A morphological study," *Microscopy Research Techniques*, Vol. 57, 530–540, 2002.
8. Tenuzzo, B., M. Dwikat, and L. Dini, "Static magnetic field selects undifferentiated myelomonocytes from low-glutamine concentration stimulated U937 cells," *Tissue & Cells*, Vol. 40, 177–184, 2008.
9. Pagliara, P., R. Lanubile, M. Dwikat, L. Abbro, and L. Dini, "Differentiation of monocytic U937 cells under static magnetic field exposure," *Eur. J. Histochem.*, Vol. 49, No. 1, 75–86, 2005.
10. Fanelli, C., S. Coppola, R. Barone, C. Colussi, G. Gualandi, P. Volpe, and L. Ghibelli, "Magnetic fields increase cell survival by inhibiting apoptosis via modulation of Ca⁺⁺ influx," *Faseb J.*, Vol. 13, 95–102, 1999.
11. Chionna, A., M. Dwikat, E. Panzarini, B. Tenuzzo, E. C. Carlà, T. Verri, P. Pagliata, L. Abbro, and L. Dini, "Cell shape and plasma membrane alterations after static magnetic fields exposure," *Eur. J. Histochem.*, Vol. 47, 299–308, 2003.
12. Dini, L., et al., *Bioelectromagnetics*, Vol. 30, 352–364, 2009.

Studies on the Effect of Static Magnetic Fields on Biological Systems

Arthur D. Rosen

Department of Biological Sciences, Purdue University, West Lafayette, IN 47907, USA

Abstract— Considerable evidence now exists relating exposure to static magnetic fields (SMFs) to changes in a number of biological systems, particularly those whose function is linked to the properties of membrane ion channels. Most of the reported effects of moderate SMF may be explained on the basis of alterations in membrane calcium ion flux as well as intracellular calcium availability. The mechanism suggested to explain these effects is based on the diamagnetic anisotropic properties of membrane phospholipids. It is proposed that reorientation of these molecules during SMF exposure will result in the deformation of imbedded ion channels, thereby altering their activation kinetics. Channel inactivation kinetics would not be expected to be influenced by these fields because the mechanism for inactivation is not located within the intramembraneous portion of the channel. Patch-clamp studies of calcium channels have provided support for this hypothesis, as well as demonstrating a temperature dependency that is understandable on the basis of the membrane thermotropic phase transition.

1. INTRODUCTION

The influence of static magnetic fields (SMFs) on biological systems has been an area of considerable interest for many years. These fields, unlike time-varying (electromagnetic) fields, are not associated with induced electric currents except during activation and deactivation or when there is movement within the field. Our interest has been primarily on moderate intensity fields, those with strengths of 1 mT to 1 T. Many of the initial reports on the bioeffects of these fields were phenomenological descriptions without attempts to identify a mechanism of action [1–5]. The present report describes the development of a hypothesis sufficient to explain the action of moderate SMFs on biological systems. This proposed mechanism involves the effect of SMFs on the molecular structure of excitable membranes, an effect sufficient to modify the function of imbedded ion-specific channels. This hypothesis would explain virtually all of the bioeffects attributed to these fields and is testable using several different neurophysiological techniques.

2. EXPERIMENTAL STUDIES

An important clue to understanding the mechanism of SMF influence on excitable membranes was the time course of this action. In examining the effect of a 120 mT SMF on the evoked potential in cats [6], a decrease in amplitude was first seen 50–95 seconds after the field was turned on and persisted for several minutes after it was turned off. If the field exerted its influence on the central nervous system simply by altering ionic current flow in excitable membranes, the effect should have been immediate. The amplitude of the evoked potential has been shown [7] to be largely a function of the activity of those cells in the lateral geniculate body that project to visual cortex. The spontaneous discharge frequency of these cells was found [8] to significantly decrease during and immediately following exposure to the 120 mT SMF. The time course was exactly as that previously observed for the evoked visual potential. In addition, the latency of the response of these cells to optic tract stimulation was not altered by exposure to a SMF. That latency, reflecting axonal conduction time, would be prolonged if the magnetic field were associated with sufficient Lorentz force to influence local ionic currents. The failure to delay central conduction is consistent with theoretical studies [9] that predicted that fields of at least 24 T would be required to slow axonal conduction. The slow change in cellular response to SMFs suggests a chemically mediated effect at the synapse, one that might influence neurotransmitter release.

A useful system for the study of synaptic neurotransmitter release is the murine neuromuscular junction. This preparation allows one to precisely control the temperature and ionic environment at the recording site, something that is virtually impossible when dealing with intact animals. Even in the absence of a presynaptic action potential, small quantities of acetylcholine are spontaneously released from presynaptic nerve terminals and trigger brief but partial depolarization of the postsynaptic membrane. These depolarizations, called miniature endplate potentials (mepps), have been shown to be dependent on the movement of Ca^{2+} through the presynaptic membrane [10]. Using an excised murine phrenic nerve-diaphragm preparation perfused with Tyrode's solution, exposure to a 120 mT SMF resulted in a temperature dependent change in mepp frequency [11]. At and

below 34°C, minimal changes in mepp frequency were noted. Above 34.5°C, a prominent decrease in mepp frequency was observed. The onset of this change was relatively slow, reaching maximum 50–100 seconds after the field was turned on. The response was reversible, with return to baseline within 50 seconds after the field was turned off. When Ca²⁺ was eliminated from the perfusate, while retaining its osmolality, the change in mepp frequency during exposure to the SMF was abolished, regardless of temperature. When exposure duration was examined [12] it was found that with a 120 mT field, a minimum exposure of 50 seconds was necessary for mepp inhibition. The efficacy of the field in inducing further inhibition was a function of its duration, but only for periods up to 150 seconds. At and above this limit, recovery time remained constant at 135 seconds. The threshold for mepp inhibition was found to be 38 mT [13].

To better understand those events occurring at the membrane in response to SMFs, direct measurement of membrane calcium currents were made using the whole-cell patch clamp technique in cultured GH3 cells [14]. Current-voltage relationships, activation kinetics, and inactivation kinetics of voltage activated calcium channels were examined before, during, and following exposure to a 120 mT SMF. The most significant change during SMF exposure was an increase in the activation time constant, evident one minute after exposure onset. There was no change in the inactivation time constant. Increase in the activation time constant was a temperature dependent phenomenon, present only above 27°C. All of the observed changes were reversible, with return to preexposure values within 3 minutes after the field was turned off.

In order to evaluate the possible effect of long term exposure to SMFs on basic cellular functions, we examined the effect of exposure of GH3 cells to a 0.5 T field, for periods of up to 5 weeks [15]. Following a 1 week exposure, cell growth declined by 22% and returned to control levels in 1 week. Although this was not statistically significant, a 4 week exposure was associated with a statistically significant decline in growth of 51% with return to control levels only after 4 weeks. Cell diameter, on the other hand, significantly increased following 3 weeks of exposure and did not return to control levels for 3 weeks after termination of exposure.

3. DISCUSSION

The one property of a biomembrane's physical structure that has the potential to be influenced by moderate intensity SMFs is its diamagnetic anisotropy. Diamagnetic anisotropic molecules will rotate in a homogeneous magnetic field and ultimately achieve an equilibrium orientation, representing the minimum free-energy state. Weakly diamagnetic molecules will exhibit a preferred orientation in the presence of a sufficiently intense SMF. In a moderate intensity magnetic field the actual degree of orientation will be quite small for single molecules, even if they are strongly anisotropic. However, for molecules aligned parallel to one another and functionally linked, individual anisotropies summate [16]. Molecular orientation in a magnetic field is opposed by the randomizing effect of thermal energy. For a domain containing N cylindrical diamagnetic anisotropic molecules with axial symmetry and a volume V , in a homogeneous magnetic field H , the degree of orientation is given by β , the ratio of magnetic to thermal energy,

$$\beta = \frac{-NH^2V(\chi_r + \Delta_\chi \cos^2 \phi)}{2k_B T} \quad (1)$$

where χ_r is the radial magnetic susceptibility vector, Δ_χ is the diamagnetic anisotropy, θ is the angle between the symmetry axis and the field direction, k_B is Boltzmann's constant and T is absolute temperature.

The energetics for orientation in a magnetic field is favorable for structures made up of a large number of parallel molecules, characteristic of biomembranes. Moderate intensity SMFs have been shown to interfere with the mechanism for calcium channel activation. This activation is a function of the intramembranous portion of the channel, specifically it is the α_1 subunit that is the primary voltage sensitive molecule in the calcium channel complex. This is the largest of the five proteins that make up the channel and is located almost entirely within the membrane. As such, it would be expected to be especially vulnerable to any membrane deformation induced by a SMF. Calcium channel inactivation was not influenced by moderate intensity SMFs since voltage dependent inactivation is linked to movement of a peptide moiety in the cell's cytoplasmic domain and membrane deformation would have little effect on such a mechanism.

Most of the diamagnetic anisotropy of lipids is contributed by their acyl chains and biological membranes, with their highly ordered phospholipid bilayer structure, would be expected to exhibit

substantial diamagnetic properties. The actual molecular reorientation within a phospholipid bilayer is the balance between the theoretical value of β as given in Eq. (1) and those intermolecular forces which limit movement. At low temperatures, the membrane exists in a gel phase similar to that of a crystalline dehydrate and denoted as the L_C phase. In this phase the hydrocarbon chains are packed tightly and rotation about their long axis is restricted. With increasing temperature, a thermally induced rotational excitation of the hydrocarbon chain occurs. This is at the subtransition temperature where the L_C phase is converted to the lamellar gel, L_β , phase. In this phase, the hydrocarbon chains exhibit limited rotational motion about their long axis. With further increase in temperature there is an abrupt rotameric disordering of the lipid acyl chain, marking the transition to the liquid-crystal, L_α phase. Although diamagnetic anisotropic molecules may assume a preferred orientation in the presence of an adequate SMF, this effect is enhanced when the membrane transitions to a less rigid phase.

Reorientation of diamagnetic structures in a magnetic field is an inherently slow process and a function of both their geometry and anisotropy, as well as of the field strength and the nature of the suspension medium. Rotatory motion of large ensembles of oriented diamagnetic molecules in a homogeneous SMF has been described by a nonlinear differential equation [17]. From that equation, the time course for rotation becomes

$$\ln \theta_1 - \ln \theta_2 = \frac{-N\Delta_\chi H^2}{\zeta} t \quad (2)$$

where θ is the angle between the molecular axial diamagnetic vector and the field direction, and t is the time required for rotation from θ_1 and θ_2 . ζ is the rotatory frictional coefficient, the value of which is determined by the size and shape of the ensemble and the viscosity of the suspension. For large diamagnetic ensembles, the randomizing effect of thermal energy is negligible. Eq. (2) provides a reasonable description of the relationship between exposure parameters and membrane molecular reorientation. At any given temperature, the size and diamagnetic anisotropy of a molecular ensemble will be constant as will the rotatory frictional coefficient. Therefore, the time course for rotation becomes a function of exposure time and the square of the flux density (H^2t). This correlates with the observed effects of SMFs at the neuromuscular junction [13]. In that study, Ca^{2+} flux through the presynaptic membrane was a linear function of H^2t , with deviation from linearity only at the highest product values. That nonlinearity suggests a mechanical limit imposed on the free rotation of membrane phospholipids. Any structure capable of restricting such movement should be distributed uniformly in close proximity to the membrane and, in some manner, bound to it. Additionally, it would have to be less influenced by a SMF than the membrane itself. These conditions are met by the cellular cytoskeleton [18].

The effect on cellular growth and size during long term exposure to SMFs is believed to be the result of changes in cell division following mitosis (cytokinesis) and is the cell's actin cytoskeleton that controls cytokinesis. Considering the low diamagnetic properties of proteins, it is unlikely that actin would be directly influenced by moderate intensity SMFs but an indirect of SMFs on the actin cytoskeleton is a possibility. This structure is not a static component of the cell but is involved in a process of continuous dynamic reorganization, modulated by intracellular Ca^{2+} . Although this is a relatively slow process, its disruption would be reflected not only in plasma membrane structure but in cytokinesis. Disruption of cytokinesis would explain the decreased rate of growth coupled with an increase in cell size seen with long term exposure to a 0.5 T SMF [15].

4. CONCLUSION

Moderate intensity SMFs influence a number of biological systems, notably those whose function is linked to transmembrane ion flux. These fields can result in a rotational displacement of the membrane's phospholipid molecule by virtue of their collective diamagnetic properties. Evidence is presented indicating that molecular rotation within the membrane matrix will influence imbedded ion channels, most likely by producing some degree of deformity of their intramembraneous segment, that part of the structure which is responsible for activation. Channel inactivation, a function of that portion of the channel not within the membrane per se, would not be expected to be altered by changes in the membrane's matrix. It is the structural properties of biological membranes that allow for summation of individual molecular anisotropies and, therefore, rotation of those ensembles at moderate field intensities.

REFERENCES

1. Barnothy, J. M., M. F. Barnothy, and I. Boszormenyi-Nagy, "Influence of a magnetic field upon the leukocytes of the mouse," *Nature (London)*, Vol. 181, 1785–1786, 1956.
2. Young, W. and J. W. Gofman, "Magnetic fields, vagal inhibition and acetylcholinesterase activity," *UCRL Rep. 12389*, Lawrence Livermore Laboratory, Livermore, CA, 1965.
3. Brewer, H. B., "Some preliminary studies on the effects of a static magnetic field on the life cycle of *lebetes reticulatus* (guppy)," *Biophys. J.*, Vol. 28, 305–314, 1979.
4. Hong, C., P. Huestis, R. Thompson, and J. Yu, "Learning ability of young rats is unaffected by repeated exposure to a static electromagnetic field in early life," *Bioelectromagnetics*, Vol. 9, 269–273, 1988.
5. Azanza, M. J. and A. Del Moral, "Cell membrane biochemistry and neurobiological approach to biomagnetism," *Prog. Neurobiol.*, Vol. 44, 517–601, 1984.
6. Rosen, A. D. and J. Lubowsky, "Magnetic field influence on central nervous system function," *Exp. Neurol.*, Vol. 95, 679–687, 1987.
7. Rosen, A. D. and E. F. Vastola, "Unit signs of visual cortex modulation by the lateral geniculate body," *EEG Clin. Neurophys.*, Vol. 20, 38–43, 1966.
8. Rosen, A. D. and J. Lubowsky, "Modification of spontaneous unit discharge in the lateral geniculate body by a magnetic field," *Exp. Neurol.*, Vol. 108, 261–265, 1990.
9. Wikswo, J. P. and J. P. Barach, "An estimate of the steady magnetic field strength required to influence nerve conduction," *IEEE Trans. Biomed. Eng.*, Vol. 27, 722–724, 1980.
10. Katz, B. and R. Miledi, "The effect of calcium on acetylcholine release from motor nerve terminals," *Proc. R. Soc. Lond. B*, Vol. 161, 496–503, 1965.
11. Rosen, A. D., "Magnetic field influence on acetylcholine release at the neuromuscular junction," *Am. J. Physiol. (Cell Physiol. 31)*, Vol. 262, C1418–C1422, 1992.
12. Rosen, A. D., "Membrane response to static magnetic fields: effect of exposure duration," *Biochim. Biophys. Acta*, Vol. 1148, 317–320, 1992.
13. Rosen, A. D., "Threshold and limits of magnetic field action at the presynaptic membrane," *Biochim. Biophys. Acta*, Vol. 1193, 62–66, 1994.
14. Rosen, A. D., "Inhibition of calcium channel activation in GH3 cells by static magnetic fields," *Biochim. Biophys. Acta*, Vol. 1282, 149–155, 1996.
15. Rosen, A. D. and E. Chastney, "Effect of long term exposure to 0.5 T static magnetic fields on growth and size of GH3 cells," *Bioelectromagnetics*, Vol. 30, 114–119, 2009.
16. Maret, G. and K. Dransfeld, "Macromolecules and membranes in high magnetic fields," *Physica*, Vol. 86B, 1077–1083, 1977.
17. Hong, F. T., D. Mauzerall, and A. Mauro, "Magnetic anisotropy and the orientation of retinal rods in a homogeneous magnetic field," *Proc. Natl. Acad. Sci.*, Vol. 68, 1283–1285, USA, 1971.
18. Carraway, K. L. and C. A. C. Carraway, "Membrane-cytoskeleton interactions in animal cells," *Biochim. Biophys. Acta*, Vol. 988, 147–171, 1989.

Cellular Perception and Static Magnetic Fields Active Penetration Depth for Pain Magnetotherapy

Pierre Le Chapellier and Badri Matta
Soissons General Hospital, Soissons, France

Abstract— Cellular perception concerns the process by which stimulation induces events through mechanical signaling pathways, according to a sense order whose nature is discussed by epistemology: It is supposed that the signification base at the origin of the cell behavior answers to a “being-envelope” whose physical nature can be approached in topological thermodynamics terms. In joined cells with gap-junctions, sense order can be covered by the signaling effect of calcium waves. But as it is related to the cell functional status, it can explain the biphasic response of the cell behavior after exposure to Static Magnetic Fields, SMF.

Four magnetobiological mechanisms solve the “kT paradox” about thermal agitation, (whose energy is $4 \cdot 10^{-21}$ J, at 310°K): 1/ Magnetosomes; (endogenous ferromagnetic nanoparticles can have magnetic moments of $2 \cdot 10^{-15}$ JT⁻¹). For ionic channels activation, these magnetosomes must operate two other kT independent mechanisms: 2/ Radicals’ pair, and 3/ Interference in angular modes of proteins. A fourth mechanism is: 4/ Protons subsystems, (inducing topological evolution in the medium).

In these mechanisms, the SMF determining parameter is magnetic flux density, or induction B : When comparing clinical trials results about pain relief by SMF, an active induction threshold of 0.5 mT is suggested. With the field decay from the magnet surface, when B is around 0.5 mT, the lateral field gradient, dB/dx , is far under its active threshold, thus secondary. Therapy must determine the SMF Active Penetration Depth from the skin surface, as compared to the tissue receptors depth. It depends on the magnet characteristics and setting.

1. INTRODUCTION TO THE POSTULATE OF CELLULAR PERCEPTION

Cellular perception concerns the process by which the cell receptors stimulation induces biological events through mechanical signaling pathways, according to a *sense order* whose nature is discussed by *epistemology*¹. *Cellular perception* supplements the molecular aspects of cell biology, because it includes two orders, *the causal order* which is about mechanical signaling and *the sense order* which is dependent on a more global mechanism. By example, the causal bioeffects of weak Electromagnetic Fields, EMF, may be apparently subtle [1]. But EMF gradients could offer cell guidance criteria capable of influencing cell migration, orientation and function [2]. Now articular cartilage health state depends on the functional state and differentiated functions of the cartilage matrix cells which are the chondrocytes [3]. These chondrocyte state and functions may be related to changes in the actine fibers of the cytoskeleton [4]. These spatial changes may result from transduction of signals coming from the cartilage matrix, which are associated with the control of gene expression [5]. Through changes in surface topologies, spatial changes can help to overcome the effect of inflammatory factors on chondrocyte response [6]. It is conjectured that changes could depend on a topological *sense order* in the cellular perception:

According to *epistemology* (as it results from *the philosophy of nature* [7, 8]) any independent cell without gap junction² is a *simple natural being* which has an intrinsic unit. In this case, its perception *sense order* is stated to be neither local nor subjective. It is asserted that the *signification base* which is at the origin of *the cell behavior* answers a “*being-envelope*”³ whose physical nature is *global* and *topological*. Insofar as the *sense order* of the individual cell answers its *topological being-envelope*, the cell behavior can be modulated by its *surrounding topological structure*. Cell-cell contact may inhibit this individual effect: The possible Static Magnetic Fields, SMF, action on chondrocyte metabolism is thus inhibited if the concentration of cultured cells is too high [10]. In those *in vitro* joined cells, the lack of individual modulation by medium topology is covered by the gap-junction functioning, particularly through a calcium waves’ effect [11].

¹With simplest, *epistemology* is the philosophy of sciences. But it also aims at locating science in an experiment of knowledge which overflows it and which is reduced finally to *the biological problem of perception*

²A *gap junction* is a local membrane vacuum acting as a ionic channel across membranes of two joint cells.

³In Whitehead and Merleau-Ponty doctrines of *organic mechanisms*, as Toscano [9] indicated, *life* is considered as a structure or a *being-envelope* for micro-phenomena; it is a relational *architectonic* that sets constraints on phenomena at lower scales. *Architectonic* means here: globally structuring, as by a qualitative topology envelopment effect.

Because of the cellular sense order, a same Ca^{2+} influx in cells exposed to SMF can be translated in a biphasic biological response. Xu [12], then Okano [13] described SMF enhance vasodilatation if vessels are relatively vasoconstricted and enhance vasoconstriction if vessels are relatively vasodilated. Mayrowitz [14] suggested the local thermodynamics modulates this bioeffect. This idea agrees with the theory of the topological thermodynamics environment, presented by Kiehn [15].

Sense order in the cellular perception is epistemologically asserted to be the soil containing the root of *cell behavior*. While acting on the medium topological thermodynamics structures, then acting upon the cellular *sense order*, stimulation by EMF may indirectly modulate the *general behavior* of the cell. For cultured chondrocytes, the starting *guidance* is that of the only electrostatic structure of the medium, such as a galvanotaxis [2, 16]. Under Pulsed EMF, PEMF, stimulation, (not SMF) a *topological structuring* can translate into morphological changes [17]. More generally the topological aspect of the full organism sense order could allow perturbations in ambient EMF activity to impact on human behavior in a clinically meaningful manner [18]. And workers exposed to perturbed EMF may have an increased brain tumor incidence attested by epidemiologic data, even if EMF mechanical effects, through an increased cytosolic calcium concentration, $[\text{Ca}^{2+}]_c$, are neither mutagenic nor tumorigenic [19].

2. BIOLOGICAL INTERACTION WITH STATIC MAGNETIC FIELDS

Sense order and possible bioeffects of PEMF and SMF gradients: From 0.6 to 6 mT, SMF may inhibit apoptosis in cells hit by apoptogenic agents, by interfering with the apoptotic process through an alteration of Ca^{2+} fluxes [20]. But as calcium flux plays a role in cell shape regulation [21], a possible occurrence of morphological changes in cell surface, depending on cells type, has also to be considered [22]. Now without any modification of cell shape, a majority of hit cell types rescue from apoptosis after a 6 mT exposure, when a promotion of apoptosis may be observed in some other types of hit cells [23]. Could it be because of a SMF topological effect acting upon the cellular *sense order*, but not on the cell morphology?

Eichwald and Walleczek [24] proposed a model of EMF regulation of cell calcium dynamics through a *feed-back* leading to a modulation of calcium entry. Such a *feed-back* is linked to the cell biological-functional status, itself linked to the *sense order* in perception. This status, then *sense order*, would induce the biphasic response behavior from the same EMF action on the cell calcium signaling. But there is a big empirical difference between the PEMF and SMF possible effects on morphological changes:

The biggest difference between PEMF and SMF structuring effects concerns the time derivative of magnetic induction, the *temporal gradient*, $\delta B/\delta t$, which induces electric fields. Many PEMF clinical tests highlighted the therapeutic interest to apply pulsations with a *temporal gradient*, $\delta B/\delta t$, set between 1 and 10 mT/ms [25]. Now, applied SMF offer no *temporal gradient* in a resting body. SMF can only offer a *spatial gradient*, particularly a *lateral gradient*, $\delta B/\delta x$. With SMF produced by homopolar disk magnets, a peak gradient region exists in a limited ring, just peripheral to magnets. Such a peak gradient region is reinforced in multipolar magnets. This way, Cavopoli [26] observed bioeffects in this peak region and suggested that the *lateral gradient threshold* for effective pain relief by neurone blockade, is around 0.5 mT/mm. Okano [27] showed a behavioral effect, (an endothelial tubular formation,) is promoted by SMF *lateral gradient*, 28 mT/mm, (when uniform SMF, 20 or 120 mT, did not). No effect had been shown for low intensity SMF gradients, even if they affect the medium topological structure.

Now, as showed by Kiehn [15], for non-equilibrium systems, topological dimensions of $\delta B/\delta t$ and $\delta B/\delta x$ are different. It implies the integral over time cannot be the same as the integral over space. This way, $\delta B/\delta t$ cannot be biologically equivalent to $\delta B/\delta x$, and therefor PEMF and SMF effects on topological structuring are not equivalent. However SMF may induce bioeffects where PEMF do not: Chiu [28] observed that SMF could promote osteoblast-like cell differentiation, via increasing *membrane rigidity*. With contrast, Jahns [17] supposed PEMF effects involved the *reorientation of membrane molecules*.

The most general target: The target for SMF interactions is generally assumed to be the *cellular membrane*. Ayrapetyan [29] suggested that the general target for SMF & PEMF interactions could be the *aqueous medium* bathing the membrane, then acting on the cellular membrane components. Further more, the most general target could be the *aqueous physical environment* which consist of the fields that bathe H_2O molecules and other chemical components in the environmental space.

A model of topological thermodynamics for non-equilibrium systems has been described by Kiehn [15]. A key result is that EMF interaction is a cubic curvature not involving the forces

generated by the E and B fields, but the A and Φ potentials and the charge current densities. Similarly, in the quantum approach of Binhi, 2007 [30], the state of a molecular target interacting with low frequency MF can be described not through forces, E and B , but through potentials A and A_O , as in the Bohm-Aharonov effect.

Andocs [31] used this idea to demonstrate that when two charged particles create a chemical bond, their interaction could be changed by an external *magnetic vector potential* A . The effect would apply to the bifurcative phenomena of hydrogen bonds in water, bringing some topological thermodynamics changes in protons subsystems.

The four possible primal mechanisms for SMF interactions: Simultaneously, the applied magnetic induction B may interact on molecular targets only through a few physical processes. Binhi [30] demonstrated that four possible mechanisms of magnetobiological effects directly solve the kT paradox about thermal agitation, (whose energy is $4 \cdot 10^{-21}$ J at 310°K [32]):

- 1/ Magnetosomes. (Endogenous ferromagnetic nanoparticles can have magnetic moments of $2 \cdot 10^{-18}$ J/mT, whose reorientation, under more than 2 mT, exceed thermal agitation., for a ionic channel activation, it must be associated with a stochastic resonance and operate the two following mechanisms [33]:
- 2/ Interferences in angular modes of proteins. (Interference of quantum states of bound heavy ions involve natural active biophysical structures [34]. And, ion-channel gating interactions may stem from locally SMF enhanced drift velocities [35].)
- 3/ Spin orientations in free radicals. (There are two aspects: Radicals' pair mechanism [36] and Mechanism of enzyme reactions involving metal radical ions with large hyperfine constants, as in anti-oxidative system, easily altered by SMF [37].)
- 4/ Aqueous medium structuration, through proton subsystems. (Hydration structure and translocation of protons affect aqueous domains [38], and modulate some membrane-bound hydrated protein activity [39, 40].)

When these primal physical mechanisms act on cellular receptors molecules, what can be the magnetic induction threshold which induces a pain relief effect?

3. THE ACTIVE PENETRATION DEPTH OF STATIC MAGNETIC FIELDS

Free radicals mechanisms act on biological reactions through non-linear complex processes and could be preponderant when the applied SMF exceed 1 mT [33, 41, 42]. However a causal effect of SMF depends also on the other coupled mechanisms. So that Fanelli [20] showed a 0.6 mT minimal threshold in cells cultured *in vitro*. Can the calcium ion-channel gating which depend on cell receptors be acted *in-vivo* for a 0.5 mT SMF threshold?

3.1. Objective: To Check the Weintraub's SMF Active Threshold

The empirical results of Weintraub [43] suggested that the minimum threshold for MF detection by the receptors of *in vivo* biological targets is around 0.5 mT. This Weintraub minimal limit comes out from his comparison of trials of weak SMF acting upon neuropathic pain. In one of these pathologies, the receptors of the sensitized afferent pain fibers are only at few mm below the skin surface. Despite the very strong field decay, the used multipolar magnets placed at some mm over the skin could then offer more than 0.5 mT at the receptors distance. But in another trial the concerned neuroreceptors were too deeply seated in the body, at a distance exceeding the SMF 0.5 mT Active Penetration Depth, APD. For this reason, the first study results showed a significant pain relief and the second not.

3.2. Method: As Weibtraub, to Carry out Another SMF APD Comparison in Two Similar Trials

Field decay and the Weintraub limit can explain clearly the opposite results between the Alfano and Colbert trials of SMF for fibromyalgia treatment. Alfano [44] used powerful magnets, offering more than 400 mT on surface, but he sets them under mattress, well below the patient. Colbert [45] used less powerful magnets, 110 mT on surface, but she sets them closer to the patient, in upper mattress. We have calculated the field decay at a distance d from the surface of similar magnets, by checking the approximation of the square law, k/d^2 , through systematic measurements at variable distances with the help of a Gaussmeter, (model HT23). The result yields two response curves, A , Colbert setting, and B , Alfano setting, drawn on Figure 1.

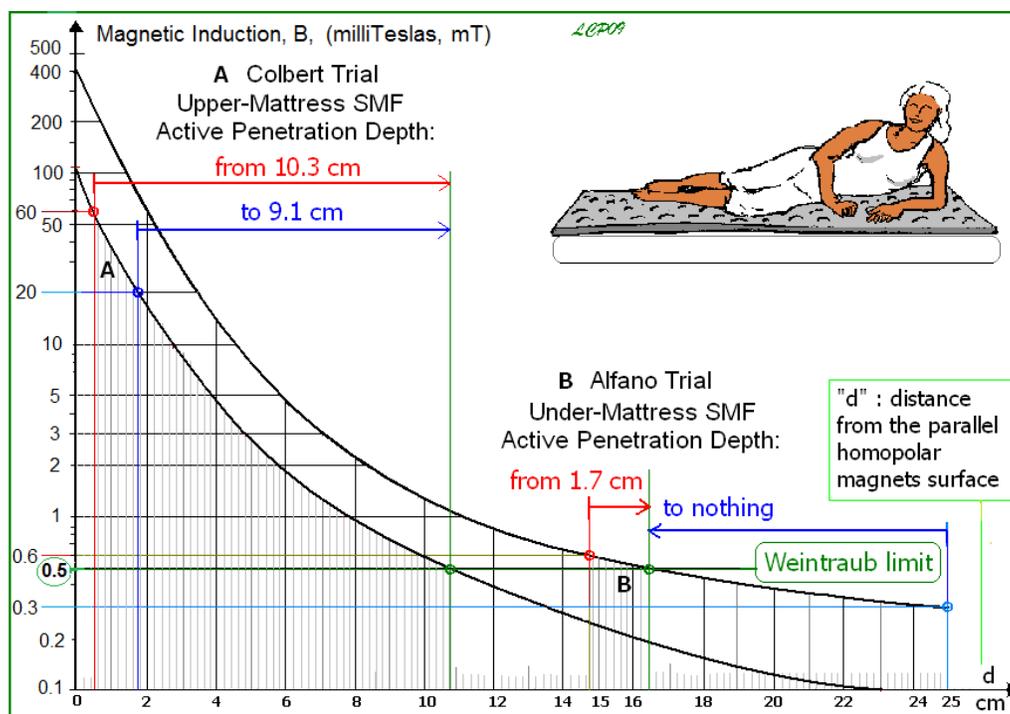


Figure 1: Static magnetic fields active penetration depth, in compared clinical trials for fibromyalgia treatment (Magnetic induction is scaled in milliTeslas, mT, with 1 mT = 10 Gauss).

3.3. Results

In the Colbert trial, the upper mattress magnetic pad offers between 60 et 20 mT at the skin surface level. For the 5 Gauss, or 0.5 mT, threshold, a graphic calculation using Figure 1 shows an active penetration depth, APD, of 10.3 to 9.1 cm, inside the patient body. This way Colbert SMF bioeffects were able to stimulate the deep peri-rachidians receptors which are relative to the pain ways.

In the Alfano study, the under mattress magnetic pad allows an active penetration offers between 0.6 and 0.3 mT at skin surface level. Graphic calculation from the Figure 1 shows an active penetration depth of 1.7 to zero cm. Alfano SMF bioeffects could thus concern only some surface receptors of the patient body.

Finally Colbert could observe a significant pain relief and a functional status improvement, when Alfano could only observe non-significant bioeffects, which is in good agreement with a 0.5 mT threshold. In the region around this threshold, the measured maximum of lateral gradient, $\delta B/\delta x$, is less than $15 \mu\text{T}/\text{mm}$, far under the active limit of $0.5 \text{ mT}/\text{mm}$, thus as a secondary causal parameter for pain relief.

4. CONCLUSION

In order to obtain a pain relief in physiopathological situation, it is assumed that the causal order in the cellular perception needs SMF induction more than 0.5 mT at the receptors level. Based on this threshold, the SMF active penetration depth, SMF-APD, appears to be an important SMF therapy parameter. This parameter can be determined with the help of a Gaussmeter. Moreover, the principle of sense order in the cellular perception could explain the bioeffects of qualitative tiny magnetic fields, through an interaction within the topological thermodynamic structure of the aqueous medium.

REFERENCES

1. Liboff, A. R., "Towards an electromagnetic paradigm for biology and medicine," *The Journal of Alternative and Complementary Medicine*, Vol. 10, No. 1, 41–47, 2004.
2. Funk, R. H. W. and T. K. Monsees, "Effects of electromagnetics fields on cells: Physiological and therapeutical approaches and molecular mechanisms of interaction," *Cells Tissues Organs*, Vol. 182, No. 2, 59–78, 2006.

3. Glowacki, J., “Chondrocyte gene expression in osteoarthritis,” *ACR Basis Research Conference, Bone and Cartilage Biology*, San-Francisco, CA, Nov. 11–12, 2001.
4. Adolphe, M., S. Demignot, et al., “La versatilité des fonctions différenciées du chondrocyte articulaire en culture,” (Modulation of differentiated functions of cultured chondrocytes, Advantages for therapy), *Bulletin De L’académie Nationale De Médecine*, Vol. 184, No. 3, 593–604, 2000.
5. Juliano, R. L. and S. Haskill, “Signal transduction from the extracellular matrix,” *The Journal of Cell Biology*, Vol. 120, No. 3, 577–585, 1993.
6. Iscrua, D. F., M. Anghelinab, S. Agarwalb, and G. Agarwala, “Changes in surface topologies of chondrocytes subjected to mechanical forces: An AFM analysis,” *Journal of Structural Biology*, Vol. 162, No. 3, 397–403, 2008.
7. Merleau-Ponty, M., *La Nature. Notes — Cours Au Collège De France*, Le Seuil, Paris, 1995.
8. Barbaras, R., “Merleau-ponty et la nature, atelier Merleau-ponty,” Université De Toulouse, <http://pedagogie.ac-toulouse.fr/philosophie/forma/barbaras.htm> (extracts translated in English, by Toscano, 2003, [9].), Sept. 9, 2000.
9. Toscano, A., “Merleau-ponty, whitehead and the politics of nature,” *Merleau-ponty and the Philosophy of Nature, Conference*, Univ. of Warwick, May 2003.
10. Stolfi, S., M. Skorvanek, P. Stolfi, J. Rosocha, G. Vasko, and J. Sabo, “Effects of static magnetic field on viability of human chondrocytes in vitro,” *Physiological Research*, Vol. 56, S45–S49, 2007.
11. Tonon, R. and P. D’Andrea, “Interleukin-1 β increases the functional expression of connexion 43 in articular chondrocytes: Evidence for a Ca²⁺-dependent mechanism,” *Journal of Bone and Mineral Research*, Vol. 15, No. 9, 1669–1677, 2000.
12. Xu, S., N. Tomita, K. Ikeuchi, and Y. Ikada, “Recovery of small-sized blood vessels in ischemic bone under static magnetic field,” *ECAM*, Vol. 4, No. 1, 59–63, 2007.
13. Okano, H., K. Gmitrov, and C. Ohkubo, “Biphasic effects of static magnetic fields on cutaneous microcirculation in rabbits,” *Bioelectromagnetics*, Vol. 20, No. 3, 161–171, 1999.
14. Mayrowitz, H. N. and E. E. Groseclose, “Effects of a static magnetic field of either polarity on skin microcirculation,” *Microvascular Research*, Vol. 69, No. 1–2, 24–27, 2005.
15. Kiehn, R. M., “Non-equilibrium systems and irreversible processes,” *Topological Torsion and Macroscopic Spinors*, Vol. 5, Lulu, <http://www.lulu.com/kiehn>, 2008.
16. Mycielska, M. E. and M. B. A. Djamgoz, “Cellular mechanisms of direct-current electric fields effects: galvanotaxis and metastatic disease,” *Journal of Cell Science*, Vol. 117, 1631–1639, 2004.
17. Jahns, M. E., E. Lou, N. G. Durdie, et al., “The effect of pulsed electromagnetic fields on chondrocyte morphology,” *Med. Bio. Eng. Comput.*, Vol. 45, 917–925, 2007.
18. Berk, M., S. Dodd, and M. Henry, “Do ambient electromagnetic fields affect behaviour? A demonstration of the relationship between geomagnetic storm activity and suicide,” *Bioelectromagnetics*, Vol. 27, No. 2, 151–155, 2005.
19. Teodoni, L., W. Göhde, M. G. Valente, F. Tagliaferri, et al., “Static magnetic fields affect calcium fluxes and inhibit stress-induced apoptosis in human glioblastoma cells,” *Cytometry Part A*, Vol. 49, No. 4, 143–149, 2002.
20. Fanelli, C., S. Coppola, R. Barone, C. Colussi, L. Ghibelli, et al., “Magnetic fields increase cell survival by inhibiting apoptosis via modulation of Ca²⁺ influx,” *The FASEB Journal*, Vol. 13, 95–102, 1999.
21. Onuma, E. and S. Hui, “Electric-field directed cell shape changes, displacement and cytoskeletal reorganization are calcium dependent,” *Journal of Cell Science*, Vol. 106, 2067–2075, 1988.
22. Belton, M., K. Commerford, J. Hall, F. S. Prato, and J. J. L. Carson, “Real-time measurement of cytosolic free calcium concentration in HL-60 cells during static magnetic field exposure and activation by ATP,” *Bioelectromagnetics*, Vol. 29, No. 6, 449–446, 2008.
23. Tenuzzo, B., A. Chionna, E. Panzarini, L. Dini, et al., “Biological effects of 6 mT static magnetic fields: A comparative study in different cell types,” *Bioelectromagnetics*, Vol. 27, No. 7, 560–577, 2006.
24. Eichwald, C. and J. Walleczek, “Activation-dependent and biphasic electromagnetic field effects: Model based on cooperative enzyme kinetics in cellular signaling,” *Bioelectromagnetics*, Vol. 17, No. 6, 427–435, 1996.

25. Shupak, N. A., “Therapeutic uses of pulsed magnetic fields exposure: A review,” *The Radioscience Bulletin*, Vol. 307, 9–32, Dec. 2003.
26. Cavopoli, A. V., A. W. Wamil, R. R. Holcomb, and M. J. McLean, “Measurement and analysis of static magnetic fields that block action potential in cultured neurons,” *Bioelectromagnetics*, Vol. 16, 197–205, 1995.
27. Okano, H., N. Tomita, and Y. Ikada, “Spatial gradient effects of 120 mT static magnetic field on endothelial tubular formation in vitro,” *BEMS2007*, 134–137, Kanazawa, Japan, Aug. 5, 2007.
28. Chiu, K. H., K. L. Ou, S. Y. Lee, C. T. Lin, W. J. Chang, C. C. Chen, and H. M. Huang, “Static magnetic fields promote osteoblast-like cells differentiation via increasing the membrane rigidity,” *Ann. Biomed. Eng.*, Vol. 35, No. 11, 1932–1939, 2007.
29. Ayrapetyan, S. N., “Cell aqua medium as a primary target for the effect of electromagnetic fields,” *Bioelectromagnetics Current Concepts, NATO Security through Science Services*, 31–63, Springer, 2006.
30. Binhi, V. N. and A. B. Rubin, “Magnetobiology: The kT paradox and possible solutions,” *Electromagnetic Biology and Medicine*, Vol. 26, 45–62, 2007.
31. Andocs, G., G. Y. Vincze, O. Szasz, P. Szendro, and A. Szasz, “Effects of curl-free potentials on water,” *Electromagnetic Biology and Medicine*, Vol. 28, No. 2, 166–181, 2009.
32. Swerdlow, A. J., et al., *Static Magnetic Fields*, Report of the independent Advisory Group on Non-ionising Radiation, Documents of the Health Protection Agency, Chilton, Didcot, UK, 2008.
33. Scaiano, C., S. Monahan, and J. Renaud, “Dramatic effect of magnetite particles on the dynamics of photogenerated free radicals,” *Photochemistry and Photobiology*, Vol. 65, No. 4, 759–762, 1997.
34. Binhi, V. N. and A. Savin, “Effects of weak magnetic fields on biological systems: physical aspects,” *Physics-Uspekhi*, Vol. 46, No. 3, 259–291, 2003.
35. Liboff, A. R., G. Vincze, and A. Szasz, “Ion velocity resonances at cell membranes,” *BEMS2008*, Vol. 12-4, 199, San Diego, Jun. 8–12, 2008.
36. Okano, H., “Effects of static magnetic fields in biology: Role of free radicals,” *Front Biosci.*, Vol. 1, No. 13, 6106–6125, 2008.
37. Pedersen, J. B., “Effects of magnetic fields on biochemical reactions,” *BEMS2008*, Vol. 12-3, 198–199, San Diego, Jun. 8–12, 2008.
38. Bankura, A. and A. Chandra, “Hydration and translocation of an excess proton in water clusters: An ab initio molecular dynamics study,” *Pranam Journal of Physics*, Vol. 65, No. 4, 763–768, 2005.
39. Wikström, M., M. I. Verkhosky, and G. Hummer, “Water-gated mechanism of proton translocation by cytochrome c oxidase,” *Biochimica & Biophysica Acta, (B & A)/Bioenergetics*, Vol. 1604, No. 2, 61–65, June 5, 2003.
40. Mentré, P., “Interfacial water: a modulator of biological activity,” *Journal of Biological Physics and Chemistry*, Vol. 4, 115–123, 2004.
41. Grissom, C. B., “Biochemical magnetic fields effects through changes...MS16-1,” F. Bersani, *El. & Mag. in Biol. & Med.; Proceedings of the 2nd World Congress*, Bologna, Jun. 8–13, 1997.
42. Le Chapellier, P. and B. Matta, “Pain relief by magnetic fields and theories of radicals pair and methionine oxidation,” *3rd Int. Forum on Pain Medicine, IFPM*, Vol. 28, 128, Montreal, Jun. 28–Jul. 01, 2007.
43. Weintraub, M. I., “Weak static magnetic fields reduce neuropathic pain dependent upon field strength and penetration,” *Bioelectromagnetics 2005*, Vol. 20-9, 232, Dublin, 2005.
44. Alfano, A. P., “Static magnetic fields for treatment of fibromyalgia: a randomized controlled trial,” *J. Alternat Complement Med.*, Vol. 7, 53–64, 2001.
45. Colbert, A. P., M. S. Markov, M. Banerji, and A. A. Pilla, “Magnetic mattress pad use in patients with fibromyalgia: a randomized double blind pilot study,” *J. Back Musculoskeletal Rehabil.*, Vol. 13, 19–31, 1999.

Analysis of Inhomogeneous Static Magnetic Field-Induced Antinociceptive Activity in Mice

János F. László¹ and Klára Gyires²

¹Institute for Research Organization, Hungarian Academy of Sciences
Nádor u. 18, 1051 Budapest, Hungary

²Department of Pharmacology and Pharmacotherapy, Semmelweis University, Budapest, Hungary

Abstract— The effect of inhomogeneous static magnetic field (SMF) on visceral pain elicited by the intraperitoneal injection of 0.6% acetic acid (writhing test) was studied in mice in an environment, where animals could freely move. 30 min, whole-body exposure of mice to SMF (permanent NdFeB N50 grade 10×10 mm cylindrical magnets with alternating poles) following the nociceptive challenge resulted in a 74% inhibition of the pain reaction ($p < 0.001$). With the help of several inhomogeneous SMF configurations, where magnets were grouped in partitions and a 2D model of ambulation, motion-induced electric current density, MR-equivalent switching, and slew rate were estimated. Their potential contribution to peripheral nerve stimulation is discussed in correlation to pain inhibition.

1. INTRODUCTION

An increasing number of evidence suggest that static magnetic field (SMF) can induce analgesic action in humans — as tested mostly in chronic pain. The beneficial effect of SMF was observed in patients under different pain conditions [1–5]. SMF exposure also shortened the inflammatory period [6].

Beside human data preclinical studies demonstrated the antinociceptive activity (AA) of SMF under experimental conditions [7–15].

Motion in an inhomogeneous SMF can induce electric currents in the moving body. The most sensitive parts of a mammal to electrical stimulus are peripheral nerves. If an electric current is high enough to stimulate visceral nerves, it certainly stimulates peripheral nerves. Consequently, pain inhibition in the viscera will definitely be effective, if induced electric currents remain below peripheral nerve stimulation (PNS) threshold.

The aim of the present study was to answer the following questions concerning motion-induced electric current density, MR-equivalent switching, and slew rate in the mouse ambulating freely while exposed to an inhomogeneous SMF: (i) Do these quantities correlate with AA? (ii) Can they be made responsible for PNS?

2. MATERIALS AND METHODS

2.1. Magnetic Field Exposure

SMF was generated by an exposure system described by László et al. [10]. This system contained two magnet holding matrices, one below and another on top of the animal cage. The exposure volume was $140 \times 140 \times 50$ mm. The individual magnets used in the matrices were NdFeB N50 type ($B_r = 1.47$ T), axially magnetized, cylindrical magnets ($d = 2r = 10$ mm, $h = 10$ mm) provided by ChenYang Technologies GmbH & Co. KG, Finsing, Germany. The poles below and above the cage facing each other were of opposite polarity allowing field lines cross the cage, i.e., the main SMF was vertically directed. SMF between the matrices was adjusted by magnetically connecting the opposite poles of the lower and upper magnets in a horseshoe-like coupling. When magnets sat one next to another with alternating polarity, the resultant SMF was inhomogeneous also laterally with a matrix constant ($\lambda =$)10 mm. It had a peak-to-peak amplitude of 754 ± 7 mT at 3 mm from the surface of the magnets, 109 ± 1 mT at 10 mm and 3 mT at 15 mm measured along the axis of a magnet and averaged over all pairs of neighbouring magnets. Magnetic induction (flux density) as well as its gradient have horizontal components, but they were a minimum of one order of magnitude below the vertical components at any position in-axis and therefore, were neglected as fringe field/gradient components. For the same reason, the vertical components off-axis were ignored.

This arrangement allowed us to insert a $140 \times 140 \times 46$ mm Plexiglas animal cage with air holes into the exposure chamber. An air permeable opaque material covered the cage on four sides to make illumination conditions similar in the exposure chamber and in the sham experiment.

A special arrangement called M -SMF denoted that the above magnet arrangement was modified in the sense that some magnets sat in the matrix with identical polarity in quadratic groups, while the neighbouring groups were of opposite polarity. We used the square root of the number of magnets (M partitioning number) in each group for distinguishing between partitionings. The notation of these arrangements follows the rule: 2×2 for the quadratic groups of 4 magnets ($M = 2$), 3×3 for groups of 9 magnets ($M = 3$), etc. The original SMF corresponds to 1×1 ($M = 1$) in this notation.

The switching rate of an MR machine is often connected to the time-varying gradient field component. We defined a similar quantity called MR-equivalent switching rate ($MEWR$). Its definition is the maximum peak-to-peak amplitude of the magnetic induction at the boundary of two partitions, divided by the average duration it takes the mouse to move through the boundary. This boundary means a distance of 10 mm, if $M = 1$, 20 mm, if $M = 2$, etc. for a mouse crossing it perpendicularly. We also defined a quantity similar to MR's slew rate, the MR-equivalent slew rate ($MESR$). This is similar to $MEWR$, but instead of the peak-to-peak amplitude of the magnetic induction it has magnetic induction gradient between two partitions in the numerator.

The magnetic induction in all experimental setups was measured with a calibrated 5 V Hall probe with 12.3 mV/T sensitivity (model UGN3503, Allegro Micro Systems, Inc., MA, USA). These measurements were executed separately from the animal experiments.

2.2. Animals, Pain Essay, Ethics

Male CFLP mice (24–26 g) were used in the experiments. Animals were housed in groups of five, the room was held under a 12 : 12 light/dark cycle at $20 \pm 2^\circ\text{C}$.

The writhing test as described by Wende and Margoli [16] and modified by Witkin et al. [17] was applied. The visceral pain was elicited by the intraperitoneal (i.p.) injection of 0.6% acetic acid in a volume of 0.2 ml/mouse. As a result of chemical irritation a characteristic stretching and writhing movement could be observed. AA of a treatment in a given time interval was defined as $AA = 100(1 - \bar{x}/\bar{y})\%$, where \bar{x} is the average of the number of writhings for the treated and \bar{y} is that for the control animals. Immediately after the administration of acetic acid some mice were exposed to SMF for 30 min. Two mice were placed in the plastic cage simultaneously and the number of writhings was monitored in 0–5, 6–20 and 21–30 min following the acetic acid challenge. In all experiments a daily control was used, when the mice were given acetic acid, placed to the sham cage, but were not exposed to SMF.

All experimental procedures were carried out according to the 1998/XXVIII Act of the Hungarian Parliament on Animal Protection and Consideration Decree of Scientific Procedures of Animal Experiments (243/1988) and complied with the recommendations of the International Association for the Study of Pain [18] and the Helsinki Declaration. The studies were approved by the Animal Care Committee of Semmelweis University, Budapest (permission number: 1810/003/2004).

2.3. Ambulation Model — Approximation of $B(x; y; z)$, Mouse Motion and Induced Electric Current Density

When approximating the m measured values of the SMF distribution along the x axis (in plane with the magnets' surface), we fitted normal (Gaussian) probability density functions to the individual peaks, where the contribution of the k -th peak to the distribution was $\pm \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{[x-(k-1)\lambda]^2}{2\sigma^2}}$, with $k = 1, 2, 3, \dots, m$, σ 's (standard deviations) were considered equal for all peaks, and λ 's (matrix constants) equal in the lateral directions. For symmetry reasons this simple 2D approach holds for the y axis as well. This model is widely used for finding the spatial resolution in laterally or vertically periodic systems [19]. $B(x; y)$ can then be approximated by the linear combination of the periodic functions for x and y , c.f., the Error Function Superposition Approximation (EFSA) introduced by Marton et al. [20]:

$$B(x; y) = B_x \frac{1}{\sigma\sqrt{2\pi}} \sum_{k=1}^m e^{-\frac{[x-(k-1)\lambda]^2}{2\sigma^2}} + B_y \frac{1}{\sigma\sqrt{2\pi}} \sum_{k=1}^m e^{-\frac{[y-(k-1)\lambda]^2}{2\sigma^2}} \quad (1)$$

with B_x and B_y magnetic induction in the x and y directions, respectively. In the present model $B_x = B_y \equiv B_0$ and we further approximate $B(x; y)$ with:

$$B(x; y) \approx \frac{B_0}{\sigma\sqrt{2\pi}} \sum_{k=1}^m e^{-\frac{(\frac{M\lambda}{2} + 1 - k)^2 \lambda^2}{2\sigma^2}} \left\{ \sin \left[\pi \left(\frac{x}{M\lambda} - 1 \right) \right] + \sin \left[\pi \left(\frac{y}{M\lambda} - 1 \right) \right] \right\} \quad (2)$$

where M is the partitioning number. The limit of this approximation is the finiteness of the resolution, namely the following condition for the resultant derivative of $B(x)$ in the range $0 \leq x \leq M\lambda$ must hold:

$$\frac{\partial B(x)}{\partial x} = B_0 \frac{(k-1)\lambda - x}{\sigma^3 \sqrt{2\pi}} \sum_{k=1}^m e^{-\frac{[x - (k-1)\lambda]^2}{2\sigma^2}} \begin{cases} = 0, & \text{if } x = \frac{M\lambda}{2} \\ \neq 0 & \text{elsewhere} \end{cases} \quad (3)$$

and similarly for $\partial B(y)/\partial y$. This gives a non-analytic solution for the $\sigma(\lambda)$ function (and thus for the resolution, $\Delta z = 2\sigma$) with the geometrical constraint $\lambda \geq d$.

In the z direction of the arrangement (perpendicular to the cylindrical magnets' surface), the magnetic induction along the axis of an individual magnet in the isocenter of the arrangement is primarily z -dependent and can be expressed in an analytic form from the integration of the Poisson equation:

$$B(z) = \frac{B_r}{2} \left\{ \frac{h+z}{[R^2 + (h+z)^2]^{1/2}} - \frac{z}{(R^2 + z^2)^{1/2}} + \frac{h+D-z}{[R^2 + (h+D-z)^2]^{1/2}} - \frac{D-z}{[R^2 + (D-z)^2]^{1/2}} \right\} \quad (4)$$

Its gradient is:

$$\frac{\partial B(z)}{\partial z} = \frac{B_r R^2}{2} \left\{ \frac{1}{[R^2 + (h+z)^2]^{3/2}} - \frac{1}{(R^2 + z^2)^{3/2}} - \frac{1}{[R^2 + (h+D-z)^2]^{3/2}} + \frac{1}{[R^2 + (D-z)^2]^{3/2}} \right\} \quad (5)$$

taking into account two (lower and upper) magnets along their common axis (in-axis) in a distance of $D = 50$ mm from each other. We fitted B_0 so that at specifically $z = 20$ mm, where AA is supposed to take place in the mouse, $B(z)$ should well estimate the measured vertical values in a first-neighbour approximation.

We also introduced N , the number of gradients per distance as the number of inflection points of the vertical magnetic induction component per average linear ambulation distance.

If SMF is also inhomogeneous laterally, mice sense a time-dependent magnetic flux (ϕ) by crossing the field lines during their ambulation in the cage. Our assumption is that the mouse moves in a 2D plane always along a straight line with constant speed (v) as long as it reaches a boundary/wall (at $0 \leq x \leq a$ or $0 \leq y \leq b$), then changes direction semi-randomly (in a degree of α — pointing into the cage) and continues its motion. This simple model can be justified by the fact that the mouse's body is 3–4 times shorter than the side length of the cage. The iterative equations of motion of the i -th linear section are as follows:

$$x_i = \Delta t |v| \cos \alpha_i + x_{i-1}; \quad y_i = \Delta t |v| \sin \alpha_i + y_{i-1} \quad (6)$$

where Δt is an equidistant time step. Since $x = x(t)$ and $y = y(t)$, then $B(x; y) = B(t)$ as if SMF would in effect be time-dependent.

As described by Maxwell's 4th (Faraday's) law, the induced electromotive force ($\varepsilon m f$) in the mouse is:

$$\varepsilon m f = \oint_{\ell_0} \vec{E} d\vec{\ell} = -\frac{d\phi}{dt} = -\frac{d}{dt} \iint_{A_0} \vec{B} d\vec{A} \quad (7)$$

where \vec{E} is the induced electrostatic field in the mouse, $d\vec{\ell}$ is an infinitesimal section of the circumference (ℓ_0) of the mouse's area exposed to SMF. We assume that the surface of a mouse perpendicular to the magnetic induction is A_0 (with the assumption $\vec{B} \parallel d\vec{A}$), \vec{B} can be regarded as constant on A_0 , and the homogeneous specific electric resistivity of the mouse is constant (ρ). Then the motion-induced current density is:

$$\vec{j}(t) = -\frac{1}{\rho \ell_0} \frac{d\vec{B}}{dt} \quad (8)$$

Using Eqs. (2) and (6), the absolute value of the induced current density of the i -th section can be expressed as:

$$j_i = -\frac{2\pi |v| B_0}{\rho \ell \lambda} \left\{ \cos \alpha_i \cos \left[\frac{2\pi}{\lambda} (\Delta t |v| \cos \alpha_i + x_{i-1}) \right] + \sin \alpha_i \cos \left[\frac{2\pi}{\lambda} (\Delta t |v| \sin \alpha_i + y_{i-1}) \right] \right\} \quad (9)$$

Data used for the estimate were: $|v| \approx 1$ m/s [21], $\rho \approx 1.82 \Omega\text{m}$ (in point 7.4.1 in [22]), $\ell_0 \approx 100$ mm, $B_0 = 0.5$ mT (after fitting to actual dosimetric values), $\lambda = 10$ mm, $k = 20\pi \text{ mm}^{-1}$, $-1 \leq \cos \alpha_i \leq 1$, $x_i \leq \min(a; b) \approx 140$ mm, $y_i \leq \min(a; b) \approx 140$ mm.

3. RESULTS

3.1. The Effect of Static Magnetic Field on Acetic Acid-induced Abdominal Pain in Mice

The effect of SMF ($M = 1$) on the writhing response measured in 0–5th, 6–20th and 21–30th min following i.p. injection of acetic acid resulted in a decrease from 9 ± 0.7 , 44 ± 0.9 and 32 ± 0.7 (sham exposed group) to 1 ± 0.3 , 12 ± 1.0 and 8 ± 0.9 , respectively, when mice were exposed to SMF (Fig. 1). The data shown are pooled averages for multiple experiments. During the total 30 min observation period the number of writhings 85 ± 1.4 decreased to 22 ± 1.6 ($p < 0.001$). The average number of writhings correspond to 85, 72, 75% AA in the three time intervals, respectively, and AA=74% for the complete 30 min time period.

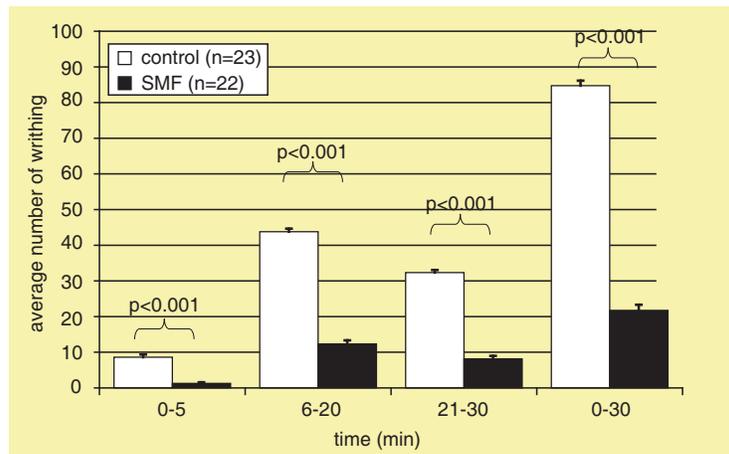


Figure 1: The effect of static magnetic field (SMF) on the writhing response induced by the intraperitoneal injection of 0.6% acetic acid measured in the 0–5, 6–20, 21–30 and 0–30 min following the administration of the acetic acid in mice. Open columns: pooled average number of writhings and S.E.M. in control animals that were not exposed to SMF. Dark columns: pooled average number of writhings and S.E.M. in mice exposed to SMF for 30 min after the administration of acetic acid. Animal numbers are shown as n . $p < 0.05$ already means significant difference to control by ANOVA-test.

3.2. The Effect of Partitions of Static Magnetic Field on Acetic Acid-induced Abdominal Pain in Mice

Figure 2 shows the average number of writhings of mice exposed to SMF ($M = 1, 2, 3, 4, 5$, and 6) for a total of 30 min exposure time. The corresponding AA values are listed in Table 1. p values were always below 0.05.

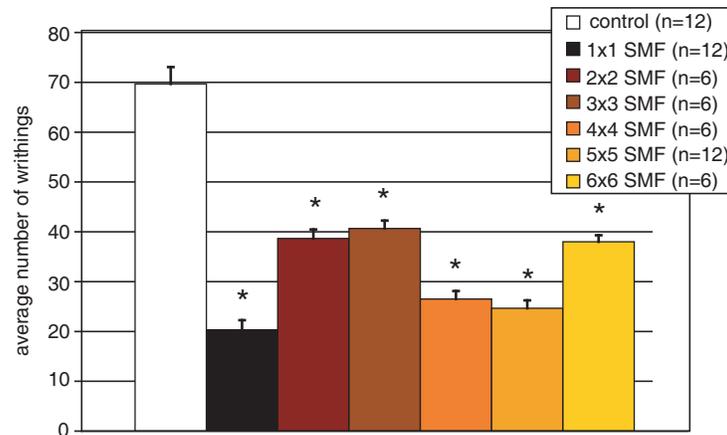


Figure 2: The average number of writhings and S.E.M. in the writhing test in mice for 30 min exposure to SMFs with different partitionings ($M = 1, 2, 3, 4, 5, 6$). For the meaning of partitionings see the text. * denotes $p < 0.05$ significant difference to control by ANOVA-test.

We also simulated $MEWR$ and $MESR$ in the M -SMF system. Table 1 presents how dosimetric data and data from the ambulation model correlate with the pain assay results.

Table 1: Correlation of the antinociceptive activities (AA) to the measured values of the peak-to-peak amplitudes of the magnetic induction, to their gradients (∇B), to the numbers of gradient per distance (N), to the MR-equivalent switching and slew rates (both $MEWR$ and $MESR$ as estimated via the ambulation model) in the M -SMFs. Vertical dosimetric data measured at $z = 20$ mm are shown with the assumption that this is the average target distance from the magnets' surface, where AA takes place in the mouse. The absolute values of induction gradient are averaged values in first-neighbour approximation. N reflects numbers of inflection points of the vertical magnetic induction component per average linear ambulation distance (140 mm). For $v = 1$ m/s average ambulation speed $MEWR$ s are equal in their numerical value to the magnetic induction gradients.

partitioning number, M	M -SMF dosimetry			ambulation model		SMF-induced antinociceptive activity, AA , % for 30 min exposure time
	peak-to-peak amplitude of the magnetic induction at $z = 20$ mm, mT	absolute value of induction gradient at $z = 20$ mm, ∇B , mT/m	numbers of gradient per distance, N	MR-equivalent switching rate ($MEWR$) at $z = 20$ mm, mT/s	MR-equivalent slew rate ($MESR$) at $z = 20$ mm, T/m/s	
1	1.6	162	14	162	16.2	74
2	17.7	46	7	46	2.3	44
3	64.9	90	3	90	3.0	42
4	133.8	140	3	140	3.5	62
5	191.2	164	2	164	3.3	65
6	212.5	154	2	154	2.6	45
correlation to AA	-3%	70%	45%	70%	68%	
correlation to AA excluding case $M = 5$ and 6	-1%	91%	62%	91%	80%	

4. CONCLUSION

Present data confirm earlier observations [10–12] that a 30 min, whole-body, inhomogeneous SMF exposure exert AA up to 74% against visceral pain elicited by acetic-acid in the mouse.

When the mouse crosses a boundary between SMF partitions, it undergoes a change of the magnetic induction in a short time. From the physical point of view this effect is identical to

either, when (i) a living object is moved into the bore of a high field MR apparatus from practically stray field, or (ii) the MR proceeds gradient switching on a static living object inside its bore. Applying the ambulation model for the motion-induced current densities (Eq. (9)), two conclusions could be drawn: (1) its long-time average in the mouse moving in a periodic SMF is zero, since the trigonometric terms result in zero in the integration. This fact can be generalized for every rectilinear motion in a laterally periodic SMFs, supposing that the typical “wavelength” of the periodicity is not bigger than the typical linear ambulation distance of the animal, the resolution obeys the rule defined in Eq. (3), and the decay time of PNS is much shorter than the time it takes for an animal to cross a boundary with strong SMF induction gradient. The partitioned SMFs are periodic with wavelengths $\lambda = 10, 20, 30,$ and 40 mm. The periodicity in case of $M = 5$ and 6 is questionable. The boundary widths are shorter than the typical linear ambulation distances in the cage (140 mm) with the exception of $M = 5$ and 6 . Thus, the discrepancy of the *AA* results for these cases may be due to the failure of the 2D model assumptions. According to models of electric signal propagation along the axon (Chapter 3 in [22]), we estimate that an impulse strong enough to stimulate a peripheral nerve would decay at a certain location in less than 2 ms, meanwhile a mouse needs 10–60 ms to intersect a boundary. (2) For a single linear path the time dependence of the motion-induced current density can be non-zero. Can this non-zero induced current density contribute to *AA*?

The threshold of the motion-induced current density for PNS was first estimated to be at 0.01 mA/m^2 [23], later 0.48 A/m^2 was proposed for human nerves [24]. Now the threshold values are accepted to be at $1\text{--}10 \text{ mA/m}^2$ [25]. The size-corrected value for mice is $0.16\text{--}1.6 \text{ mA/m}^2$. In our ambulation model the absolute value of the motion-induced current density remains below 1.8 mA/m^2 — not excluding an effect. We must assume that the *AA* effect of SMF exposure based on the endogenous μ -opioid receptors is either more effective and/or has a preventive manner as compared to the nerve stimulatory effect leading to pain sensation.

The clinically determined PNS threshold for the MRs' switching rate is between 50 [26] and 60 T/s [27]. The mean pain nerve stimulation threshold is at 90 T/s, the mean cardiac stimulation threshold is by 2 orders of magnitude higher [27]. In our experiments *MEWRs* lay between 46 and 164 mT/s, 2 orders of magnitude below threshold. Switching rates were shown not to contribute to pain inhibition in clinical studies, rather the contrary was found [28]. On the other hand Table 1 suggests a strong correlation between *AA* and *MEWR*. We may resolve this contradiction remembering that the inhomogeneity of SMF was demonstrated earlier not to be necessary to achieve *AA*; the homogeneous SMF of a 3 T clinical MR also exerts 69% *AA* in the writhing test in mice [12]. Although the peak-to-peak amplitude of the magnetic induction does not have a strong correlation directly to *AA* (Table 1), but it may have through its gradient. We may interpret these facts that *AA* may not exclusively be related with the inhomogeneity of SMF.

The 1D single axis slew rates of clinical MRs are typically 20 T/m/s and above. *MESR* calculated in our model for case (i) above remains well below threshold, at cca. 1 T/m/s for a clinical MR. The estimated *MESR* values in Table 1 are all below typical slew rates. From the correlations of the *M*-SMF-induced *AA* to the *MESRs* (Table 1), we may conclude for case (ii) that the higher the *MESR*, the more expressed the *AA*.

REFERENCES

1. Holcomb, R. R., W. B. Worthington, B. A. McCullough, and M. J. McLean, “Static magnetic field therapy for pain in the abdomen and genitals,” *Pediatr. Neurol.*, Vol. 23, No. 2, 261–264, 2000.
2. Alfano, A. P., A. G. Taylor, P. A. Foresman, P. R. Dunkl, G. G. McConnell, M. R. Conaway, and G. T. Gillies, “Static magnetic fields for treatment of fibromyalgia: A randomized controlled trial,” *J. Altern. Complement Med.*, Vol. 7, No. 1, 53–64, 2001.
3. Brown, C. S., F. W. Ling, J. Y. Wan, and A. A. Pilla, “Efficacy of static magnetic field therapy in chronic pelvic pain: A double-blind pilot study,” *Am J. Obstet. Gynecol.*, Vol. 187, No. 6, 1581–1587, 2002.
4. Hinman, M. R., J. Ford, and H. Heyl, “Effects of static magnets on chronic knee pain and physical function: A double-blind study,” *Altern. Ther. Health. Med.*, Vol. 8, No. 3, 50–55, 2002.
5. Smania, N., E. Corato, A. Fiaschi, P. Pietropoli, S. M. Aglioti, and M. Tinazzi, “Therapeutic effects of peripheral repetitive magnetic stimulation on myofascial pain syndrome,” *Clin. Neurophysiol.*, Vol. 114, No. 1, 350–358, 2003.

6. Weinberger, A., A. Nyska, and S. Giler, "Treatment of experimental inflammatory synovitis with continuous magnetic field," *Isr. J. Med. Sci.*, Vol. 32, No. 12, 1197–1201, 1996.
7. Kavaliers, M. and K. P. Ossenkopp, "Magnetic fields differentially inhibit mu, delta, kappa and sigma opiate-induced analgesia in mice," *Peptides*, Vol. 7, No. 2, 449–453, 1986.
8. Del Seppia, C., L. Mezzasalma, E. Choleris, P. Luschi, and S. Ghione, "Effects of magnetic field exposure on open field behaviour and nociceptive responses in mice," *Behav. Brain Res.*, Vol. 144, No. 1–2, 1–9, 2003.
9. Prato, F. S., J. A. Robertson, D. Desjardins, J. Hensel, and A. W. Thomas, "Daily repeated magnetic field shielding induces analgesia in CD-1 mice," *Bioelectromagnetics*, Vol. 26, No. 1, 109–117, 2005.
10. László, J., J. Reiczigel, L. Székely, A. Gasparics, I. Bogár, L. Bors, B. Rácz, and K. Gyires, "Optimization of static magnetic field parameters improves analgesic effect in mice," *Bioelectromagnetics*, Vol. 28, No. 8, 615–627, 2007.
11. Gyires, K., B. Rácz, Z. Zádori, and J. László, "Pharmacological analysis of static magnetic field-induced antinociceptive action in the mouse," *Bioelectromagnetics*, Vol. 29, No. 6, 456–462, 2008.
12. László, J. and K. Gyires, "3 T MR significantly inhibits pain in the writhing test in mice," *Life Sciences*, Vol. 84, No. 1–2, 12–17, 2008.
13. László, J., J. Tímár, Zs. Gyarmati, Zs. Fürst, and K. Gyires, "Pain-inhibiting inhomogeneous static magnetic field fails to influence locomotor activity and anxiety behaviour in mice: No interference between magnetic field- and morphine-treatment," *Brain Research Bulletin.*, Vol. 79, 316–321, 2009.
14. Antal, M. and J. László, "Exposure to inhomogeneous static magnetic field ceases mechanical allodynia in neuropathic pain," *Bioelectromagnetics*, Vol. 30, No. 6, 438–445, 2009.
15. Sándor, K., Zs. Helyes, K. Gyires, J. Szolcsányi, and J. László, "Static magnetic field-induced antinociceptive effect and the involvement of capsaicin-sensitive sensory nerves in this mechanism," *Life Sciences*, Vol. 81, 97–102, 2007.
16. Wende, V. C. and S. Margoli, "Analgesic test based upon experimentally induced acute abdominal pain in rats," *Fed. Proc.*, Vol. 15, 494, 1956.
17. Witkin, L. B., C. F. Heuter, F. Galdi, E. O'Keefe, P. Spitaletta, and A. J. Plummer, "Pharmacology of 2-amino-indane hydrochloride (Su-8629), a potent non-narcotic analgesic," *J. Pharmacol. Exp. Ther.*, Vol. 133, 400–408, 1961.
18. Zimmermann, M., "Ethical guidelines for investigations of experimental pain in conscious animals," *Pain*, Vol. 16, 109–110, 1983.
19. Hofmann, S., "Quantitative depth profiling in surface analysis," A Review, *Surf. Interf. Anal.*, Vol. 2, No. 3, 148–160, 1980.
20. Marton, D., J. László, J. Giber, and F. G. Rüdener, "EFSA — A new evaluation method of the depth resolution in depth profiling of multilayer structures," *Vacuum*, Vol. 35, No. 12, 523–526, 1985.
21. Chappell, M. A., T. Garland, E. L. Rezende, Jr., and F. R. Gomes, "Voluntary running in deer mice: Speed, distance, energy costs and temperature effects," *The Journal of Experimental Biology*, Vol. 207, 3839–3854, 2004.
22. Malmivuo, J. and R. Plonsey, *Bioelectromagnetism*, Oxford University Press, New York, Oxford, 1995.
23. International Programme on Chemical Safety, Environmental Health Criteria 69, Magnetic Fields, Section 1.4.2, <http://www.inchem.org/documents/ehc/ehc/ehc69.htm>, 1987.
24. Guidelines on Exposure to Electromagnetic Fields from Magnetic Resonance Clinical Systems, Environmental Health Directorate Health Protection Branch, Published by Authority of the Minister of National Health and Welfare, 87-EHD-127, 1987.
25. Kangarlu, A. and P. M. L. Robitaille, "Biological effects and health implications in magnetic resonance imaging," *Concepts. Magn. Reson.*, Vol. 12, 321–359, 2000.
26. Formica, D. and S. Silvestri, "Biological effects of exposure to magnetic resonance imaging: An overview," *Bio. Med. Eng. OnLine*, Vol. 3, 1–12, 2004.
27. Safety guidelines for conducting magnetic resonance imaging (MRI) experiments involving human subjects, Center for Functional Magnetic Resonance Imaging, University of California, San Diego, http://fmriserver.ucsd.edu/pdf/center_safety_policies.pdf, 2007.
28. Shellock, F. G. and J. V. Crues, "MR procedures: Biologic effects, safety, and patient care," *Radiology*, Vol. 232, 635–652, 2004.

Radiation Induced Forward Emitter Current Gain Degradation of Lateral and Vertical PNP Power Transistors in Voltage Regulators

Vladimir Vukić¹ and Predrag Osmokrović²

¹Institute of Electrical Engineering “Nikola Tesla”, Belgrade, Serbia

²Faculty of Electrical Engineering, University of Belgrade, Belgrade, Serbia

Abstract— Voltage regulators “National Semiconductor” LM2940CT5 and “STMicroelectronics” L4940V5 were examined in gamma radiation field. It was perceived that voltage regulators LM2940CT5, manufactured by the use of conventional monolithic bipolar process, characterized by lateral pnp transistors with round emitters, became unfunctional after absorption of low doses of radiation. On the other hand, voltage regulators L4940V5, BiCMOS integrated circuits created with implementation of local oxide side isolation process, showed much higher radiation hardness.

Detailed examinations of voltage regulator L4940V5 pointed on significant reduction of serial transistor’s forward emitter current gain, causing the great increase of quiescent current. The main reason of high radiation susceptibility of examined vertical serial pnp transistor is implementation of interdigitated emitter, with high perimeter-to-area ratio. Level of vertical pnp power transistor’s forward emitter current gain degradation was similar to the current gain degradation of transistors with round emitters in circuits LM2940CT5, with small perimeter-to-area ratio. Experiment showed similar influences of emitter geometry and technological process of pnp transistor creation on degradation of power transistor’s characteristics, yet the functioning of two types of integrated circuits was completely different.

1. INTRODUCTION

In the previous papers were presented results of the examination of low-dropout voltage regulators “National Semiconductor” LM2940CT5 and “STMicroelectronics” L4940V5 in ionizing radiation fields [1–3]. It was perceived that voltage regulators LM2940CT5, manufactured by the use of conventional monolithic bipolar process with lateral pnp transistors with round emitters, became unfunctional after absorption of low doses of radiation (less than 300 Gy (SiO₂)). On the other hand, voltage regulators L4940V5, BiCMOS integrated circuits created with implementation of local oxide side isolation process, showed much higher radiation hardness (more than 10 kGy (SiO₂)). Before experiment it was assumed that local oxide may cause creation of parasitic MOSFET, i.e., radiation-induced leakage currents between collector and emitter of serial transistor in voltage regulators L4940V5. However, implementation of isolated collector in integrated circuit made by “HDS²/P² Multipower 20 V” process prevented negative influence of local oxide isolation. Hypothesis was that the main cause of L4940V5’s high radiation hardness was small degradation of serial vertical pnp transistor’s forward emitter current gain, mainly due to the shift of current flow from the surface towards the substrate, while the assumption of L2940CT5’s low radiation hardness was that it is caused by the rapid loss of lateral pnp transistor’s forward emitter current gain in ionizing radiation field.

In the new series of experiments were measured voltage regulator’s output current, output voltage and quiescent current, obtaining the possibility to calculate the serial pnp transistor’s forward emitter current gain.

2. EXPERIMENT

Integrated 5-volt positive commercial-off-the-shelf voltage regulators “ST Microelectronics” L4940V5 and “National Semiconductor” LM2940CT5 were tested in Institute of Nuclear Sciences “Vinča”, Belgrade, in Metrology-dosimetric laboratory.

⁶⁰Co was used as a source of γ radiation and it was situated in the device for the realization of γ -field, IRPIK-B. Accepted mean energy of γ -photons is $E_\gamma = 1.25$ MeV. Samples were irradiated in the mouth of collimator.

Exposition doses measurement was exerted with the cavity ionizing chamber “Dosimentor” PTW M23361, volume $3 \cdot 10^{-5}$ m³, with uncertainty of measurement $\pm 2\%$. With cavity ionizing chamber, reader DI4 was used [4].

Samples of voltage regulators LM2940CT5 and L4940V5 were irradiated in groups of four circuits. Ten meters long cables supplied devices. Beside the supply cables were laid sense cables of the same length. Current and voltage measurements were carried out with laboratory instruments “Fluke” 8050A and “Hewlett-Packard” 3466A. All measurements and the irradiation of components were performed on room temperature of 20°C.

The main values used for detection of voltage regulator’s degradation due to exposure to ionizing radiation were forward emitter current gain and maximum load current. Measured electrical values were voltage regulator’s output current, output voltage and quiescent current. Examination of maximum collector current change was performed in the following way: for the constant input voltage equal to 8 V, load current was increased until output voltage dropped to 4.7 V. Lower output voltages are unacceptable for voltage regulator, since the device is beginning to shutdown [5]. The next step was measurement of output voltage and quiescent current for unloaded voltage regulator, with input voltage of 8 V. In voltage regulators with serial pnp power transistor, quiescent current represents the sum of control circuit’s internal consumption current and serial transistor’s base current. Measurement of quiescent current for unloaded voltage regulator obtains value of internal consumption, with minor influence of serial transistor’s base current. Subtraction of unloaded circuit’s quiescent current from quiescent current of maximally loaded device, for the same input voltages, gives the value of serial transistor’s base current. Additional measurement of output current, i.e. serial power transistor’s collector current, obtains possibility to calculate the forward emitter current gain of serial transistor, both during the irradiation and after the absorption of specified amount of total dose.

Devices had been irradiated until predetermined total doses were reached. To avoid the effects of recombination in semiconductor after irradiation, all measurements were performed up to half an hour after the exposure. Devices in γ radiation field were exposed to total dose of 500 Gy (SiO_2), with dose rate of 4 cGy (SiO_2)/s.

More details about experiment, sources of ionizing radiation, test procedure and technological processes implemented in manufacture of examined circuits are provided in references [1–4].

3. RESULTS

Data presented in Figures 1–3 were obtained by the tests of circuits LM2940CT5 from batch PM44AE, made by “National Semiconductor’s” subcontractor in China. Circuits were packaged in Malacca, Malaysia.

It was perceived earlier [1, 3] that voltage regulators LM2940CT5 became unfunctional after absorption of low doses of radiation. The main reason for circuit failures for low total doses was not loss of forward emitter current gain, but the degradation of error amplifier circuit [3].

The very beginning of irradiation of voltage regulators LM2940CT5 brought rapid decrease of maximum output current, but also increase of serial transistor’s forward emitter current gain for unbiased devices and circuits operating with low currents.

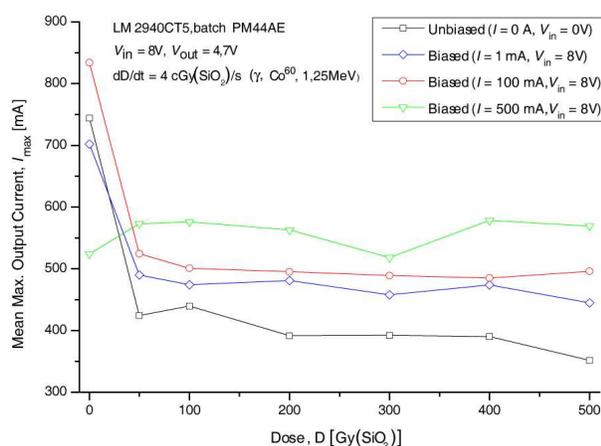


Figure 1: Change of mean maximum output current in voltage regulator LM2940CT5 under influence of γ radiation.

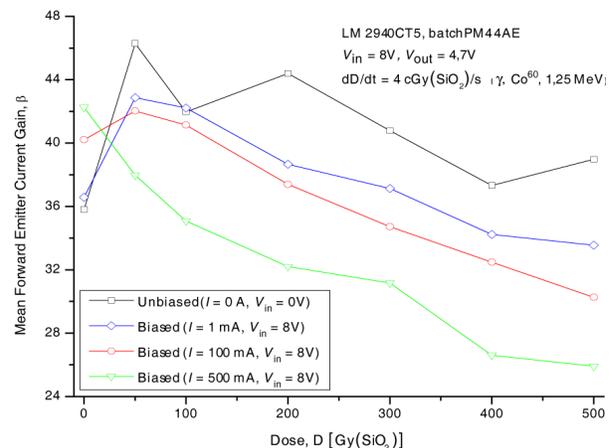


Figure 2: Change of serial transistor’s forward emitter current gain in voltage regulator LM2940CT5 under influence of γ radiation.

In the beginning of irradiation of voltage regulator LM2940CT5, due to the slow formation of interface traps, oxide traps had dominant influence, causing the rapid increase of electric field generated by the positive trapped charge in the oxide. Electric field caused by the positive trapped charge rises quickly in the highly contaminated oxides [6]. Under the influence of external electric field starts transport of holes and hydrogen ions towards the interface oxide-semiconductor and generation of interface traps. Owing to the influence of “electrostatic barrier”, generated by space charges in oxide, hole and hydrogen ion transport towards the interface was very slow, causing the dominant influence of oxide trapped charge on lateral pnp transistor. This effect induced abrupt reduction of emitter efficiency and decrease of the base current due to reduction of space-charge region in the base area [4].

Specified interpretation relates to operation with small emitter currents. When high current flows through the emitter, due to the degradation of lightly doped emitter, caused by the positive trapped charge in the oxide, space-charge region spreads deep into the p-type emitter. Examined samples with load current of 500 mA operate with significantly lower forward emitter current gain, with high carrier injection into emitter. In the case of lateral pnp transistors, high holes current flows through the n-type base, suppressing the influence of positive trapped charge in the oxide on surface accumulation in the base area. Increase of interface traps concentration proportional to total dose causes additional rise of base current and degradation of forward emitter current gain. High current flow through the lateral pnp transistors does not cause significant recombination of positive trapped charge in the oxide because the minority carriers in the pnp transistor’s base area are holes.

Forward emitter current gain degradation of serial pnp transistor is less than expected for lateral pnp transistors with lightly doped emitters (decrease up to 40% in regard with current gain before irradiation). Anyhow, it appeared that older technological process, based on round, not interdigitated emitters, is the primary reason for moderate degradation of forward emitter current gain of power transistor. On the other hand, small perimeter-to-area ratio had significant effect on degradation of emitter injection efficiency, especially during the operation with high currents, reducing the maximum current supplied to the load. Therefore, information of forward emitter current gain in operating point was not sufficient data for evaluation of serial transistor’s radiation hardness, demanding also information of maximum current.

In Figures 4–6 are presented data of examinations obtained for voltage regulators “STMicroelectronics” L4940V5 from batch WKOOGO 408, made in China.

Despite the verified great radiation hardness of voltage regulators L4940V5, from Figures 5 and 6 can be noticed significant decrease of forward emitter current gain, reaching more than 50% in regard with initial values. During the first series of experiments it was assumed that the main reason for high radiation tolerance of voltage regulators L4940V5 was small influence of ionizing radiation on vertical pnp transistor’s forward emitter current gain.

The main reason why radiation susceptibility of examined vertical serial pnp transistor was

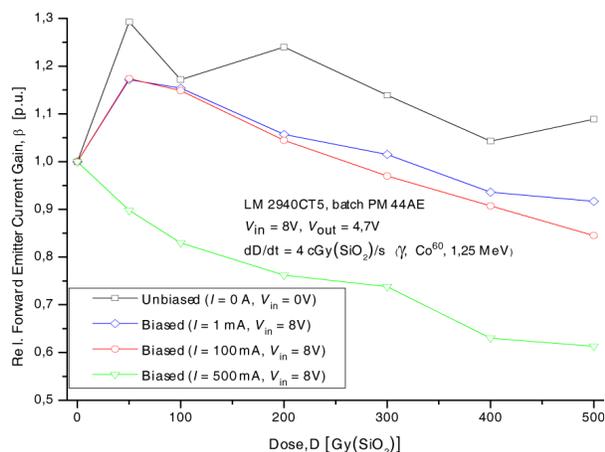


Figure 3: Relative change of serial transistor’s forward emitter current gain in voltage regulator LM2940CT5 under influence of γ radiation.

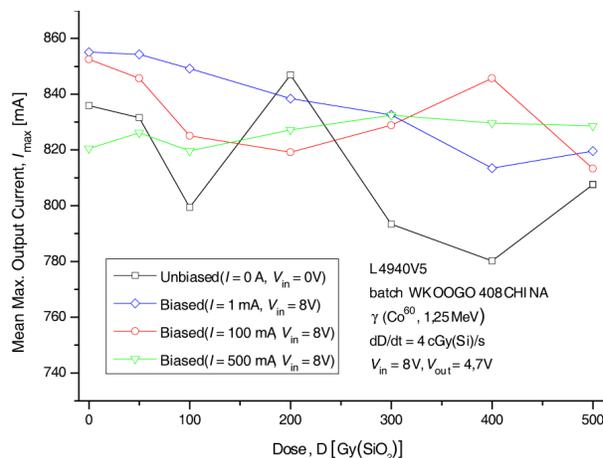


Figure 4: Change of mean maximum output current in voltage regulator L4940V5 under influence of γ radiation.

higher than expected is implementation of interdigitated emitter, with high perimeter-to-area ratio, applied in order to increase the emitter efficiency during the operation with high currents. In total, 36 groups of elementary pnp transistors occupy high chip area, with very high perimeter-to-area ratio [2]. Consequently, positive trapped charge in oxide has great impact on emitter injection efficiency and spread of space-charge region deep into the emitter area. After the initial rapid decrease, forward emitter current gain fell on one half of initial value (for samples with low load currents) after absorption of total dose 200 Gy (SiO₂), remaining in saturation for higher doses. Yet, for the exact evaluation of forward emitter current gain degradation are necessary data of serial transistor's base-emitter voltage (V_{BE}). Although the base and emitter currents are known values, voltage V_{BE} during the measurement of serial transistor's collector and base current remained unknown, since it was not possible to measure this value during the experiment. Only if values of current gain was known in various conditions for constant values of V_{BE} , reconstruction of characteristic $\beta(I_E)$ could be possible. Therefore, conclusion is that owing to the negative feedback reaction the serial transistor's operating point moved away from maximum of characteristic $\beta(I_E)$, not that for all values of emitter current appeared twofold decrease of forward emitter current gain.

Examinations of maximum output current (Figure 4) pointed on small change of this parameter, although the forward emitter current gain was halved in some cases. Examinations of maximum output current demand serial transistor operation in high level injection mode (ideality factor $n = 2$). In this mode serial transistor operates in far right part of characteristic $\beta(V_{BE})$, where change of V_{BE} has the least influence on forward emitter current gain. Some authors consider reduction of voltage regulator's maximum output current proportional to serial transistor's forward emitter current gain [7], so the change of forward emitter current gain in high-level injection mode of operation may consider basic for mutual comparison of sample's radiation tolerance. In the worst case, decline of maximum output current of voltage regulator L4940V5 didn't exceed 7%. Increase of base current, also as stable regulation of output voltage pointed on primary significance of driver transistor and control circuits (especially voltage reference and error amplifier) on voltage regulator's radiation hardness, while the current gain loss of serial pnp transistor had less influence.

Mentioned effects didn't have influence on proper functioning of circuit L4940V5, even after absorption of very high total doses.

The only similarity in radiation responses between two types of low-dropout voltage regulators was response of samples that operated with high load current during the irradiation. In serial transistors operating in conditions of high-level carriers injection in emitter (total circuit's current of 500 mA), noticed decline of lateral serial transistor's forward emitter current gain (voltage regulator LM2940CT5) was some 40%, while the current gain decline of vertical transistor, situated in circuit L4940V5, was about 30%.

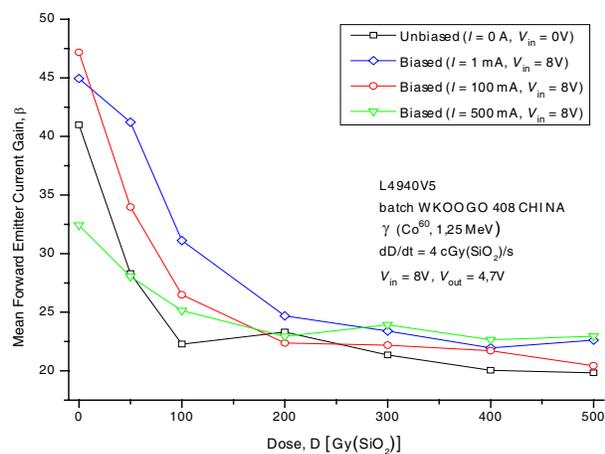


Figure 5: Change of serial transistor's forward emitter current gain in voltage regulator L4940V5 under influence of γ radiation.

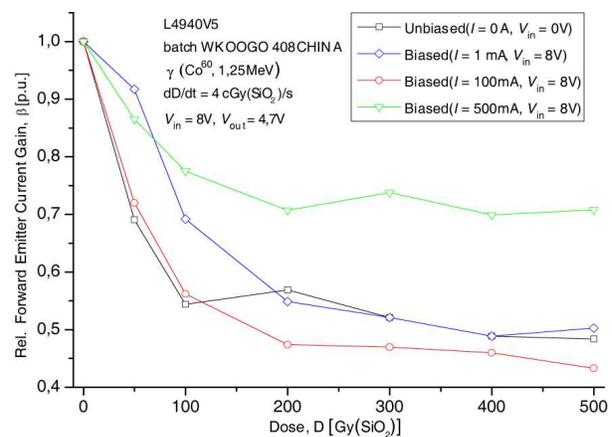


Figure 6: Change of relative value of serial transistor's forward emitter current gain in voltage regulator L4940V5 under influence of γ radiation.

4. CONCLUSION

Voltage regulators “National Semiconductor” LM2940CT5, made by the use of conventional monolithic bipolar process with lateral pnp transistors and round emitters, even due to the use of older technological process and creation with highly contaminated isolation oxide, showed notable degradation of lateral pnp transistor’s forward emitter current gain, but much less than expected. However, all samples had functional failure after exposure to low total doses of medium-dose-rate ionizing radiation. On the other hand, voltage regulators “STMicroelectronics” L4940V5, BiCMOS integrated circuits created by the use of vertical process with side local oxides, demonstrated exceptionally high radiation hardness.

Increase of absorbed total dose caused higher electron injection from the base into the emitter area of lateral pnp transistors, having impact on increase of lateral pnp transistor’s base current during the irradiation of voltage regulators LM2940CT5.

Detailed examinations of voltage regulator L4940V5 pointed on significant degradation of serial transistor’s forward emitter current gain, that exerted by the great increase of voltage regulator’s quiescent current. Nevertheless, this effect didn’t affect device’s proper functioning. The main reason for noticed sensitivity of examined vertical serial pnp transistor was application of interdigitated emitter, with great perimeter-to-area ratio, used to increase the emitter efficiency in conditions of operation with high current density.

Beside the assumed high radiation hardness of vertical pnp transistors, primarily due to the relocation of current flow from the surface towards the substrate, experiment pointed on the great influence of emitter geometry on characteristics of power transistor, comparable with the influence of implemented technological process. Also, it was noticed that changes of serial transistor’s parameters don’t need to be the decisive factor in the correct operation of voltage regulators in the radiation environment.

ACKNOWLEDGMENT

This work was supported by the Ministry of Science and Technological Development of Republic Serbia under project 141046G, “Physics of electromagnetic and radiation compatibility of electrical materials and components”.

REFERENCES

1. Vukić, V., P. Osmokrović, and S. Stanković, “Influence of medium dose rate X and gamma radiation and bias conditions on characteristics of low-dropout voltage regulators with lateral and vertical serial PNP transistors,” *8th European Conference on Radiation and Its Effects on Components and Systems RADECS 2005*, Cap d’Agde, France, September 19–23, 2005.
2. Vukić, V. and P. Osmokrović, “Total ionizing dose response of commercial process for synthesis of linear bipolar integrated circuits,” *Journal of Optoelectronics and Advanced Materials*, Vol. 8, No. 4, 1538–1544, 2006.
3. Vukić, V. and P. Osmokrović, “Total ionizing dose degradation of power bipolar integrated circuit,” *Journal of Optoelectronics and Advanced Materials*, Vol. 10, No. 1, 219–228, 2008.
4. Vukić, V., “Radiation tolerance of analog bipolar integrated circuits,” Ph.D. Thesis, University of Belgrade, 2008 (in Serbian).
5. McClure, S. S., J. L. Gorelick, R. L. Pease, and A. H. Johnston, “Dose rate and bias dependency of total dose sensitivity of low dropout voltage regulators,” *IEEE Radiation Effects Data Workshop*, 100–105, 2000.
6. Shaneyfelt, M. R., R. L. Pease, J. R. Schwank, M. C. Maher, G. L. Hash, D. M. Fleetwood, P. E. Dodd, C. A. Reber, S. C. Witzzak, L. C. Riewe, H. P. Hjalmanson, J. C. Banks, B. L. Doyle, and J. A. Knapp, “Impact of passivation layers on enhanced low-dose-rate sensitivity and thermal stress effects in linear bipolar IC’s,” *IEEE Transactions on Nuclear Science*, Vol. 49, 3171–3179, 2002.
7. Pease, R. L., S. McClure, J. Gorelick, and S. C. Witzzak, “Enhanced low-dose-rate sensitivity of a low-dropout voltage regulator,” *IEEE Transactions on Nuclear Science*, Vol. 45, 2571–2576, 1998.

Influence of Gamma Radiation on Some Commercial EPROM and EEPROM Components

Boris Loncar¹, Srboljub Stankovic², Koviljka Stankovic¹, and Bojan Jovanovic¹

¹University of Belgrade, Belgrade, Serbia

²VINCA Institute of Nuclear Sciences, Belgrade, Serbia

Abstract— This paper compares the reliability of standard commercial Erasable Programmable Read Only Memory (EPROM) and Electrically Erasable Programmable Read Only Memory (E²PROM) components exposed to gamma rays. Results obtained for CMOS-based EPROM (NM27C010) and E²PROM (NM93CS46) components provide evidence that EPROMs have a greater radiation hardness than E²PROMs. Moreover, the changes in EPROMs are reversible, and after erasure and reprogramming all EPROM components restore their functionality. On the other hand, changes in E²PROMs are irreversible. The obtained results are analyzed and interpreted on the basis of gamma ray interaction with the CMOS structure.

1. INTRODUCTION

Major advantages of Electrically Erasable Programmable Read Only Memory (EEPROM or E²PROM) over Erasable Programmable Read Only Memory (EPROM) components are the elimination of UV erase equipment and the much faster in-the-system erasing process (in milliseconds compared with minutes for high-density EPROM). On the other hand, a major drawback of E²PROMs is the large size of their two transistor memory cells compared to single transistor cells of EPROMs [1]. Following the shift from NMOS to CMOS transistor technology, presentday programmable non-volatile memories are mostly CMOS-based, as is the case with both memory models investigated in this paper. The influence of neutron displacement damage, primarily reflected in the change of minority carrier lifetime, is negligible in all MOS (Metal-Oxide Semiconductor) structures, since their operation is not significantly affected by minority carrier lifetime. Other types of neutron damage, including secondary ionization and carrier removal, are minimal and indirect [2, 3]. CMOS structure is naturally immune to alpha radiation, due to the shallow well. The formation of electron-hole pairs by an alpha particle will primarily take place in the substrate below the well. The well forms an electrical barrier to the carriers, preventing them from reaching the gate and influencing transistor operation. Any carriers generated in the well itself recombine quickly or get lost in the flow of majority carriers [1, 4]. Gamma radiation may cause significant damage to programmable memories, deteriorating properties of the oxide layer, and was therefore considered in this paper.

2. EXPERIMENTAL PROCEDURE

Examination of EPROM and E²PROM radiation hardness was carried out in cobalt-60 (⁶⁰Co) gamma radiation field at the “Vinca” Institute of Nuclear Sciences, Belgrade, Serbia. Absorbed dose dependence of the changes in memory samples caused by irradiation was monitored. The ⁶⁰Co source was manufactured at Harwell Laboratory. Air kerma rate was measured at various distances from the source with a Baldwin-Farmer ionization chamber. Absorbed dose was specified by changing the duration of irradiation and the distance between the source and the examined memory samples. Absorbed dose in Si was calculated from the absorbed dose in air, by using the appropriate mass energy-absorption coefficients for an average energy of ⁶⁰Co gamma quanta equal to 1.25 MeV. Mass energy-absorption coefficients for silicon and air ($\mu_{\text{enSi}}(1.25 \text{ MeV}) = 0.02652 \text{ cm}^2/\text{g}$, $\mu_{\text{enAIR}}(1.25 \text{ MeV}) = 0.02666 \text{ cm}^2/\text{g}$) were obtained from the NIST tables [5]. Testing was performed on samples of COTS (Commercial Of The Shelf) EPROMs and E²PROMs. EPROMs used for the investigations were NM27C010 components, with 1,048,576-bit storage capacity, organized as 128 K-words of 8 bits each, in a 32-Lead PLCC package. E²PROM samples used were NM93CS46 components, with 1024-bit storage capacity, organized as 64 × 16-bit array, packaged in an 8-pin DIP chip carrier. Fourty samples were used for both EPROM and E²PROM testing, based on which the average results presented in the paper were obtained. All tests were performed at room temperature (25°C). Irradiation of a 40-sample batch was conducted in consecutive steps, corresponding to the increase of total absorbed dose. Dose increment was 30 Gy per irradiation step for EPROMs and 50 Gy for E²PROMs. All memory locations (cells) were initially written into a logic

'1' state, corresponding to an excess amount of electron charge stored on the floating gate. This state has been shown to be more radiation sensitive than the '0' state, responding with a greater threshold voltage shift for the same absorbed dose [6]. Effects of gamma radiation were examined in terms of the number of "faults" in memory samples following irradiation. A fault is defined by the change of a memory cell logic state as a consequence of irradiation. The content of all memory locations was examined after each irradiation step, where by the number of read logic '0' states equaled the number of faults. Although ionizing radiation effects in MOS structures are generally dose-rate dependent, effects in EPROM and E²PROM cells don't depend on dose rate. Radiation induced charge changes on the floating gate occur extremely fast, and so are in phase with any incident radiation pulse [7].

3. RESULTS AND DISCUSSION

Plots presented in the paper are based on mean values taken over 40 samples. Both the differential and aggregate relative change of the number of faults with the absorbed dose in EPROM samples is shown in Figures 1(a) and (b). First faults, of the order of 2%, appeared at 1120 Gy. The number of faults increased with the rise of the absorbed dose. At dose values above 1240 Gy, significant changes in memory content were observed.

Changes in EPROMs proved to be reversible, i.e., after UV erasure and reprogramming all EPROM components became functional again — consecutive erasing, writing and reading of the previously irradiated samples was efficiently performed several times. A repeated irradiation procedure of EPROM samples, following erasure and reprogramming to '1' state, produced faults already at 220 Gy, with significant failures in memory content occurring above 310 Gy, as shown in Figures 2(a) and (b). The lower threshold of fault occurrence upon repeated irradiation testifies of the cumulative nature of radiation effects. The differential and aggregate relative change of the

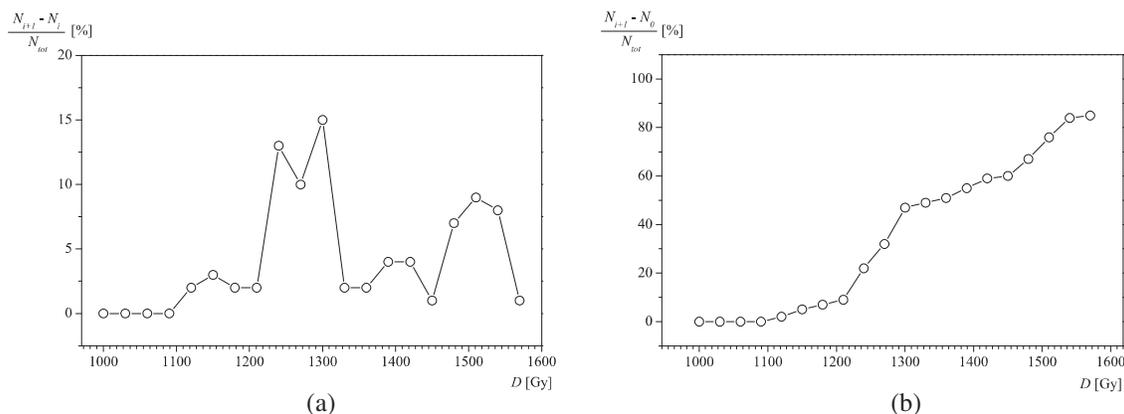


Figure 1: Average relative change of the number of faults versus the absorbed dose in irradiated EPROM samples: (a) differential, (b) aggregate ($N_{tot} = 1,048,576$ bits, $N_0 = 0$).

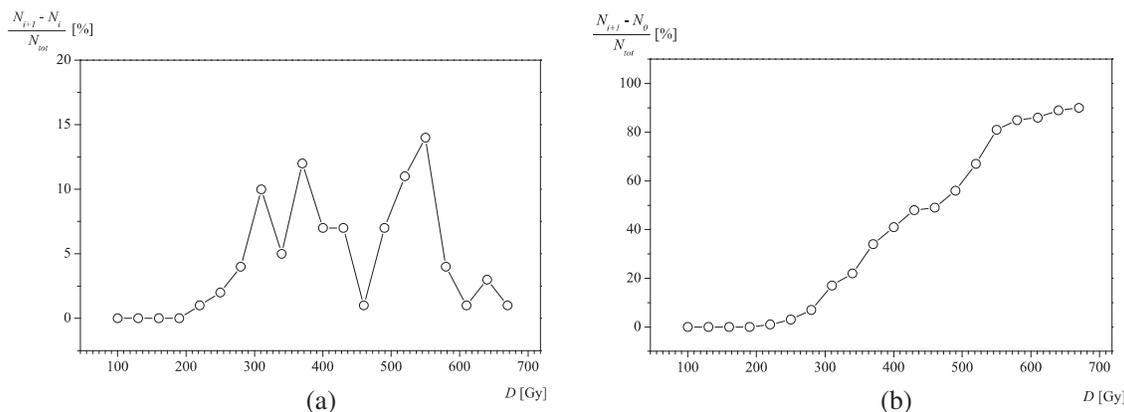


Figure 2: Average relative change of the number of faults versus the absorbed dose in reprogrammed and repeatedly irradiated EPROM samples: (a) differential, (b) aggregate ($N_{tot} = 1,048,576$ bits, $N_0 = 0$).

number of faults with the absorbed dose in irradiated E²PROM samples is shown in Figures 3(a) and (b). First faults appeared at 900 Gy, proving E²PROMs to be more sensitive to gamma radiation than EPROM components. With further dose increase, the number of faults also increased. Moreover, the changes in E²PROMs appeared to be irreversible. Irreversibility of radiation damage in E²PROMs was established based on the fact that the standard erasure procedure was unable to erase the contents of any of the irradiated memory samples. In CMOS EPROMs and E²PROMs, utilizing either *N*-well or *P*-well technology, the dual polysilicon gate, consisting of the control and the floating gate, resides over an *N*-channel transistor. Polysilicon layer floating gate, insulated from the control gate above it and the silicon channel below it by the gate oxide, is used to store charge and thus maintain a logical state. Charge is stored on the floating gate through hot electron injection from the channel in EPROMs, and through cold electron tunneling from the drain in E²PROMs. The stored charge determines the value of transistor threshold voltage, making the memory cell either ‘on’ or ‘off’ at readout [1].

Passing through the gate oxide (SiO₂), gamma radiation breaks Si-O and Si-Si covalent bonds, creating electron/hole pairs. The number of generated electron/hole pairs depends on gate oxide thickness. Recombination rate of these secondary electrons and holes depends on the intensity of electric field in the irradiated oxide, created by charge stored at the floating gate, and modulated by the change in charge carrier concentration and their separation within the oxide. The greater the electric field, the larger the number of carriers evading recombination. Incident gamma photons generate relatively isolated charge pairs, and recombination is a much weaker process than in the case of highly ionizing particles.

Secondary electrons which escape recombination are highly mobile at room temperature. In the ‘1’ state of the memory cell, excess amount of electrons stored on the floating gate maintains an electric field in both oxide layers, that swiftly drives the secondary electrons away from the oxide to the silicon substrate and the control gate. The direction of electron motion is generally dependent

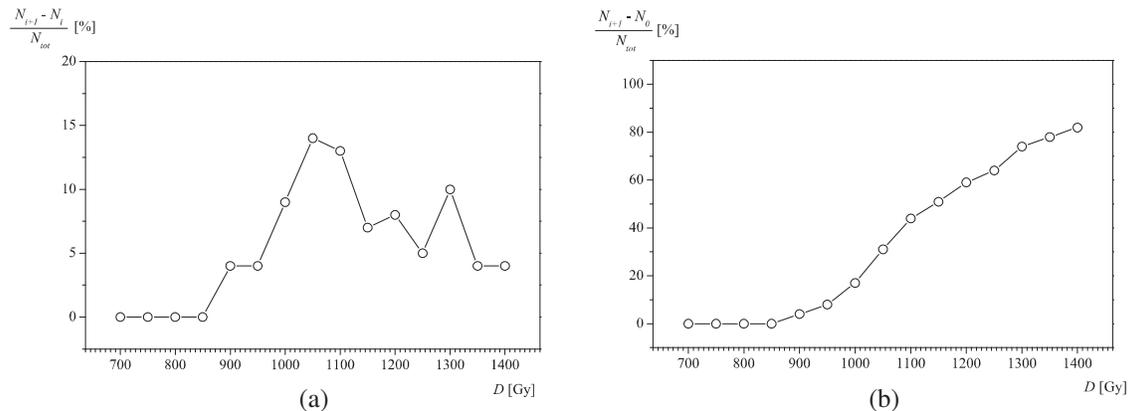


Figure 3: Average relative change of the number of faults versus the absorbed dose in irradiated E²PROM samples: (a) differential, (b) aggregate ($N_{tot} = 1024$ bits, $N_0 = 0$).

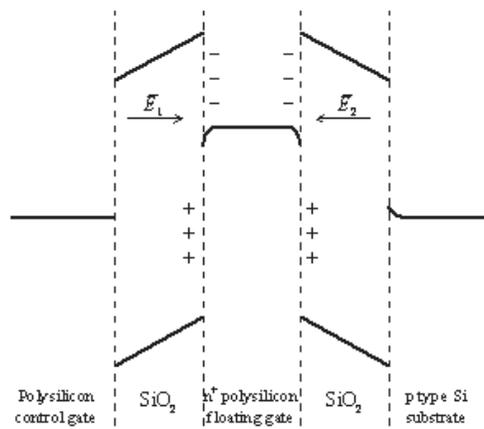


Figure 4: Energy band diagram of the the dual polysilicon gate when programmed into the logic ‘1’ state.

on the gate voltage polarity at the time of irradiation. Electron drift occurs even with no external voltage applied to the gate, due to work function potentials. The logic '1' state (excess amount of negative charge stored on the floating gate) has been chosen as the starting state in this paper, since it is more liable to fault occurrence. Energy band diagram of the the dual polysilicon gate when programmed into the excess electron (logic '1') state is shown in Figure 4. The diagram also illustrates electric field direction in the oxide (SiO_2) and the charge stored on the floating gate.

In addition to electron/hole pair creation, secondary electrons may produce defects in the oxide by way of impact ionization. Colliding with a bonded electron in either an unstrained silicon-oxygen bond ($\equiv\text{Si-O-Si}\equiv$), a strained silicon-oxygen bond, or a strained oxygen vacancy bond ($\equiv\text{Si-Si}\equiv$), a secondary electron may give rise to one of the hole trapping complexes. Interaction with an unstrained silicon-oxygen bond gives rise to one of the energetically shallow complexes ($\equiv\text{Si-O}\bullet^+\text{Si}\equiv$ or $\equiv\text{Si-O}^+\bullet\text{Si}\equiv$, where \bullet denotes the remaining electron from the bond). Strained silicon-oxygen bonds, distributed mainly near the oxide/substrate and oxide/floating gate interfaces, are easily broken by the passing electrons, giving rise to the amphoteric non-bridging oxygen (NBO) center ($\equiv\text{Si-O}\bullet$) and the positively charged $\equiv\text{Si}^+$ center (known as the E'_s center). Collision of the secondary electron with one of the strained oxygen vacancy bonds, also concentrated near the interfaces, leads to the creation of the $\equiv\text{Si}^+\bullet\text{Si}\equiv$ center (known as the E'_γ center). Hole traps generated in the bulk of the oxide are shallow, while the centers distributed in the vicinity of the interfaces (NBO, E'_s , E'_γ and their variants) act as deep hole traps at which long-term trapping of holes occurs [9, 10]. The latter are referred to as interface traps, surface states, or border traps [11].

While traversing the oxide, radiation-generated secondary electrons themselves create additional electron/hole pairs. Some of the secondary electrons may be trapped within the oxide, but this is a low-probability event, due to their high mobility and the low concentration of electron trapping sites in thermally grown SiO_2 [8].

Holes generated in the oxide by incident gamma radiation and through secondary ionization are far less mobile than the electrons. They are either trapped in the oxide, or move toward the floating gate under the influence of the electric field in the logic '1' state. Hole transport through the oxide occurs by means of two mechanisms: hopping transport via direct hole tunneling between localized trap sites, and trap-mediated valence band hole conduction.

The probability of holes moving through the oxide breaking unstrained silicon-oxygen bonds is low. However, since hydrogen atoms and hydroxyl groups are always present in thermally grown oxides as impurities, migrating holes may create defects by interacting with either $\equiv\text{Si-H}$ or $\equiv\text{Si-OH}$ bonds, whereby hydrogen atoms and ions (H° and H^+) are released. Once reaching the oxide/floating gate interface, holes can break both strained silicon-oxygen bonds and strained oxygen vacancy bonds, producing NBO, E'_s and E'_γ centers. Holes trapped at the oxide/substrate interface which recombine with electrons injected from the substrate may produce another kind of amphoteric defect ($\text{Si}_3 \equiv\text{Si}^\bullet$, a silicon atom at the interface back bonded to three silicon atoms from the floating gate), called the P_{b0} center [12–14].

Radiation produced bulk defects may themselves migrate in the strained region near either the oxide/floating gate or the oxide/substrate interface and result in the formation of interface traps [15, 16].

Another mechanism of interface trap buildup includes hydrogen atoms and ions released by the holes in the oxide. Hydrogen atoms and ions diffuse and drift toward the oxide/floating gate interface. When a H^+ ion arrives at the interface, it picks up an electron from the floating gate, becoming a highly reactive hydrogen atom H° , which is able to produce interface traps [17, 18]. Interface traps may also be generated through direct interaction of incident gamma photons [19]. The holes not trapped in the oxide are injected into the floating gate, reduce the net amount of electron charge stored on it, and thereby decrease the threshold voltage of the cell's NMOS transistor. The trapping of holes occurs mostly at the oxide/floating gate interface, where the concentration of deep hole traps is high. The positive charge of these trapped holes, will tend to mask the negative electron charge on the floating gate, again reducing the transistor threshold voltage. Thus, the trapped and the injected holes both produce a negative threshold voltage shift. Small oxide thickness gives rise to considerable fluctuation of absorbed energy, directly influencing the number of faults in the examined samples. Moreover, the amount of radiation-induced defects acting as electron and hole traps is a complex function of the gate oxide material, as well as of the doping and processing methods used in securing the oxide onto the silicon surface. These are the reasons for the observed variation in the number of faults among the examined memory samples. Another effect caused by gamma radiation is electron emission from the floating gate. This kind

of emission is the basis for standard EPROM erasure by UV radiation. During irradiation, gamma photons cause electrons to be emitted over the floating gate/oxide potential barrier. Once in the oxide, electrons are swept to the substrate or control gate by the electric field. The loss of electrons from the floating gate causes additional decrease of the threshold voltage. The net effect of charge trapping in oxide and at oxide/floating gate interfaces, as well as of floating gate hole injection and electron emission, is the change of the NMOS transistor threshold voltage. Radiation induced change of the threshold voltage may affect memory cell logic state at readout. Threshold voltage V_T is, hence, the key parameter of memory cell state. Modeling charge stored at the NMOS floating gate as charge on a parallel plate capacitor, threshold voltage can be expressed as:

$$V_T = V_{T0} + \frac{q_s d}{\varepsilon} \quad (1)$$

where V_{T0} is the initial transistor threshold voltage due to processing, q_s is the floating gate surface charge density, d is the oxide thickness between the control and floating gate, ε is the oxide dielectric constant. This model disregards the dependence of threshold voltage on the actual position of the trapped charge sheet within the oxide. The influence of gamma irradiation on programmable memories is manifested through the change of the net gate surface charge density. Threshold voltage as a function of the absorbed dose can be represented by the empirical relation:

$$V_T(D) = V_T^{eq} + (V_{T0} - V_T^{eq}) e^{-\alpha D} \quad (2)$$

where α is a constant dependent on the type and energy level density of the traps in the oxide, V_T^{eq} is the threshold voltage at extremely high doses (also called the radiation saturation voltage), when an equilibrium of the dominant processes causing the change of gate surface charge density — hole trapping, hole injection, electron emission and electron-hole recombination — is achieved. UV photons with an energy lower than the bandgap of silicon dioxide (≈ 9 eV) are incapable of creating electron-hole pairs in the oxide, but are capable of exciting electrons from the silicon substrate into the oxide, where they recombine with the trapped holes. Irradiation of EPROMs by UV light during erasure partially reduces the radiation-induced trapped charge from previous exposure to gamma photons. This light-induced annealing of trapped holes can account for the observed reversibility of changes in EPROMs. The cumulative nature of gamma radiation effects observed in EPROM components can be attributed to the fact that not all holes trapped at radiation induced interface states are annealed during UV erasure at ambient temperature. Since E²PROM erasing process involves no UV irradiation, there can be no light-induced annealing of trapped holes in these components. Thermal annealing of holes trapped at deep interface traps is not evident at ambient temperatures. Current-induced annealing of trapped holes, due to recombination of holes with electrons being driven from the floating gate to the drain, could be expected to occur during E²PROM electrical erasure. However, this kind of annealing is known to require much longer times compared to the duration of a standard E²PROM erasing procedure. On the whole, no significant annealing of trapped holes occurs in E²PROMs, and hence radiation-induced changes in these components appeared irreversible on the time scale of experiments performed in this paper (~ 10 hours). Higher sensitivity of the tested E²PROM components to gamma radiation is a consequence of a more pronounced radiation induced electron emission from the floating gate over the thin oxide region (≈ 10 nm) between the floating gate and the drain, due to a lower potential barrier [2].

4. CONCLUSION

This paper presents the results of the examination of programmable ROMs' radiation hardness. Influence of cobalt-60 gamma radiation was tested on EPROM and E²PROM components. EPROM components exhibited higher radiation reliability than E²PROMs. Significant faults in EPROM and E²PROM components appeared at 1240 Gy and 900 Gy, respectively. Changes in EPROMs are reversible, and after erasing and reprogramming, all EPROM components restored their functionality. Reversibility of changes in EPROMs is attributed to partial light-induced annealing of trapped holes during UV erasure. Due to the cumulative radiation effects, first failures of the previously irradiated EPROMs appear at significantly lower doses. On the other hand, E²PROM changes are irreversible. All observed phenomena have a plausible theoretical explanation, based on the interaction of gamma rays with the oxide layer of memory cell MOS transistors. The influence of gamma radiation is basically manifested through the change of the net floating gate surface charge density, and consequently of transistor threshold voltage.

ACKNOWLEDGMENT

The Ministry of Science and Environmental Protection of the Republic of Serbia supported this work under contract 141046.

REFERENCES

1. Prince, B., *Semiconductor Memories, A Handbook of Design, Manufacture and Applications*, 2nd Edition, John Wiley & Sons, New York, 1991.
2. Messenger, G. C. and M. S. Ash, *The Effects of Radiation on Electronic Systems*, 2nd Edition, Van Nostrand Reinhold, New York, 1992.
3. Ma, T. P. and P. V. Dressendorfer, *Ionizing Radiation Effects in MOS Devices and Circuits*, John Wiley & Sons, New York, 1989.
4. Srour, J. R., *Basic Mechanisms of Radiation Effects on Electronic Materials, Devices and Integrated Circuits*, DNA-TR-82-20, New York, 1982.
5. Hubbell, J. H. and S. M. Seltzer, *Tables of X-ray Mass Attenuation Coefficients and Mass Energy-absorption Coefficients*, Version 1.4, National Institute of Standards and Technology, Gaithersburg, MD, 2004. Available Online: <http://physics.nist.gov/xaamdi>.
6. Snyder, E. S., P. J. McWhorter, T. A. Dellin, and J. D. Sweetman, "Radiation response of floating gate EEPROM memory cells," *IEEE Transactions on Nuclear Science*, Vol. 36, 2131, 1989.
7. Osmokrovic, P., M. Stojanovic, B. Loncar, N. Kartalovic, and I. Krivokapić, *Nucl. Instrum. & Methods in Phys. Res. B*, Vol. 140, 143, 1998.
8. Ristic, G. S., M. M. Pejovic, and A. B. Jaksic, "Modelling of kinetics of creation and passivation of interface traps in metal-oxide-semiconductor transistors during postirradiation annealing," *J. Appl. Phys.*, Vol. 83, 2994, 1998.
9. Ristic, G. S., M. M. Pejovic, and A. B. Jaksic, "Analysis of postirradiation annealing of N-channel power vertical double-diffused metal-oxide-semiconductor transistors," *J. Appl. Phys.*, Vol. 87, No. 7, 3468, 2000.
10. Ristic, G. S., M. M. Pejovic, and A. B. Jaksic, *J. Non-Cryst. Sol.*, Vol. 353, 170, 2007.
11. Fleetwood, D. M., "'Border traps' in MOS devices," *IEEE Transactions on Nuclear Science*, Vol. 39, No. 2, 269, 1992.
12. McLean, F. B., "A framework for understanding radiation-induced interface states in SiO₂ MOS structures," *IEEE Transactions on Nuclear Science*, Vol. 27, No. 6, 1651, 1980.
13. Lai, S. K., "Interface trap generation in silicon dioxide when electrons are captured by trapped holes," *J. Appl. Phys.*, Vol. 54, No. 5, 2540–2546, 1983.
14. Shaneyfelt, M. R., J. R. Schwank, D. M. Fleetwood, P. S. Winokur, K. L. Hughes, and F. W. Sexton, "Field dependence of interface-trap buildup in polysilicon and metalgate MOS devices," *IEEE Transactions on Nuclear Science*, Vol. 37, No. 6, 1632–1640, 1990.
15. Loncar, B., S. Stankovic, D. Novakovic, M. Kovacevic, and P. Osmokrovic, *Proc. 2nd Yugoslav Nuclear Society Conference*, 575, Beograd, 1998.
16. McGarity, J. M., "Considerations for hardening MOS devices and circuits for low radiation doses," *IEEE Transactions on Nuclear Science*, Vol. 27, No. 6, 1739–1744, 1980.
17. Pejovic, M., A. Jaksic, and G. Ristic, "The behaviour of radiation-induced gate-oxide defects in MOSFETs during annealing at 140°C," *J. Non-Cryst. Sol.*, Vol. 240, 182–192, 1998.
18. Ristic, G. S., M. M. Pejovic, and A. B. Jaksic, *Microel. Eng.*, Vol. 40, 51, 1998.
19. Pejovic, M. and G. Ristic, "Creation and passivation of interface traps in irradiated MOS transistors during annealing at different temperatures," *Solid-St. Electron.*, Vol. 41, No. 5, 715–720, 1997.

Ambiguous Influence of Radiation Effects in Solar Cells

Aleksandra Vasic, Milos Vujisic, Koviljka Stankovic, and Bojan Jovanovic

University of Belgrade, Serbia

Abstract— Regardless of the very high standards in the production of electronic devices, all of them are more or less, prone to the effects of aging even if they are not exposed to extreme (hostile) working conditions. One of the most limiting factors for all kinds of detectors is their noise, such as frequency dependent generation-recombination noise, burst noise and $1/f$ noise. Common source of noise that is connected to the hostile working conditions is radiation. Photovoltaic (PV) conversion of solar energy is one of the most up-to-date semiconductor technologies that enables application of PV systems for various purposes. Solar cells, the basic elements for photovoltaic conversion of solar energy, are especially susceptible to radiation damage, primarily due to their large surface. The lifetime of the solar cell is restricted by the degree of radiation damage that the cell receives. This is an important factor that affects the performance of the solar cell in practical applications. Introduction of radiation-induced recombination centers reduce the minority carrier lifetime in the base layer of the p-n junction increasing series resistance, and lead to an enormous increase of noise in solar cells. After very high doses of radiation series resistance of the base layer could be so high that most of the power generated by the device is dissipated by its own internal resistance. The aim of this paper is to investigate the influence of radiation on output characteristics of solar cells in connection to their noise level that could lead to better understanding of transport properties, electron-hole pair creation etc.

1. INTRODUCTION

Photovoltaic (PV) conversion of solar energy is one of the most up-to-date semiconductor technologies that enables application of PV systems for various purposes. The wider substitution of conventional energies by solar energy lies in the rate of developing solar cell technology. Silicon is still the mostly used element for solar cell production, so efforts are directed to the improvement of physical properties of silicon structures. Silicon solar cells belong to a wide group of semiconductor detector devices, though somewhat specific in its design (larger than most of the detectors). One of the most important characteristic of detectors is their energy resolution that primarily depends on noise. That is why lowering noise is important for obtaining good quality detectors. It is known that low frequency noise ($1/f$ and burst noise) is manifested as random fluctuation of the output current or voltage, leading to lowering of the efficiency of the device. Various experiments suggests [1, 2] that the origin of this noise is fluctuation of the number free charge carriers connected to existence of the traps located in the vicinity or directly in the junction area, or fluctuation of the mobility of charge carriers. In both cases these fluctuations arise from the interactions of carriers with defects, surface states and impurities, that are either introduced during manufacturing of the device, or as a consequence of the hostile working conditions (radiation, high temperature, humidity). Because of the large surface to volume ration, surface effects are expected to be a major cause of $1/f$ noise, so good quality contacts are of great importance. Silicides belong to a very promising group of materials with low resistivity and good temperature stability that are used for fabrication of reliable and reproducible contacts. Even so, surface effects such as surface recombination fluctuations in carrier mobility, concentration of surface states, etc., have great influence on frequency dependent noise in silicides also. It has been found [3, 4] that ion implantation of As^+ ions in the formation of the silicides could improve electrical characteristics of silicides regarding their noise level.

Other most common source of noise that is connected to the hostile working conditions is radiation. For example, spent nuclear fuels emit simultaneously, in addition to γ -rays, several neutrons also, so semiconducting device (e.g., solar cell) placed in the vicinity of these fuels sustains different kind of radiation damage both from γ -rays and from neutrons. The lifetime of the solar cell is restricted by the degree of radiation damage that the cell receives. This is an important factor that affects the performance of the solar cell in practical applications.

The main effect of the radiation is an increase of the saturation current generated within or at the surface of the depletion region. The permanent damage in the solar cells materials is caused by collisions of the incident radiation particles with the atoms in the crystalline lattice, which are displaced from their positions. These defects degrade the transport properties of the material and particularly the minority carrier lifetime [5–9]. This lifetime decrease produces degradation of the

parameters of the cell ultimately leading to an increase of the noise level. The interaction between vacancies, self-interstitials, impurities, and dopants in Si leads to the formation of undesirable point defects such as recombination and compensator centers which affects performance of the solar cells, especially in space. Introduction of radiation-induced recombination centers reduce the minority carrier lifetime in the base layer of the p-n junction increasing series resistance. After very high doses of radiation series resistance of the base layer could be so high that most of the power generated by the device is dissipated by its own internal resistance [10, 11]. However, small doses of radiation carefully introduced and monitored, could have some beneficial effects on device performance due to possible relaxation of crystal lattice, leading to lowering of series resistance.

In this paper, two aspects of improvement of solar cell characteristics via lowering noise level were presented. First one is related to methods of lowering $1/f$ noise in silicides that could be used as reliable contacts, and the other to the application of small dose radiation.

2. EXPERIMENTAL PROCEDURE

Experimental measurements in this paper concerning $1/f$ noise in silicides included studies of arsenic ion implantation effects on the frequency noise level characteristics of TiN/Ti contacts on p-type Si. Ion implantation with As^+ ions, annealing and electrical characterization were performed on the samples. Thin films were deposited by ion sputtering in a Balzers Sputtron II system, at the background pressure in the chamber of 1×10^{-6} mbar, and the argon pressure during sputtering of 1×10^{-3} mbar. Implantation of arsenic was performed at 350 keV with the dose range between 1×10^{15} ions/cm² to 1×10^{16} ions/cm². Thermal treatment was performed at different temperatures for 20 min. Structural analysis of TiN/Ti/Si samples was performed by Rutherford backscattering spectrometry (RBS), with a 1.5 MeV He^+ ion beam at normal incidence and the detector position at 160° . Noise level measurements were performed with the measurement equipment consisting of the multichannel analyzer ND-100, low noise pre-amplifier, and amplifier (standard ORTEC equipment). MAESTRO II code was used for automatic energy calibration. Frequency noise measurements were performed at room temperature.

Experimental measurements concerning solar cells were carried out on the commercially available solar cells based on encapsulated polycrystalline silicon manufactured by Leybold. Solar cells were irradiated with Pu-Be point neutron source. Samples were in direct contact with the source, and maximum dose was $dD/dt = 0.36$ mGy/h. Current-voltage data were used for the characterization of the properties of solar cells. Standard measurement equipment was used to measure I - V curve for two illumination levels of 32 W/m^2 and 58 W/m^2 after every irradiation step.

3. RESULTS AND DISCUSSION

3.1. Noise Level in Silicides

The results of structural (RBS) analysis have shown that ion implantation did not induce redistribution of components for lower implantation doses. The spectra indicate that the entire titanium layer has interdiffused with the silicon substrate. The presence of the TiSi_2 and TiSi_3 phase in the implanted samples was observed. In all cases top TiN layer remains unaffected, but for higher doses of implantation (1×10^{16} ions/cm²) a disordered structure was registered. This corresponds to the amorphization of silicon substrate, which is moving deeper with the ion dose, showing that the physical properties of TiN/Ti/Si are influenced by the implantation.

This observation was confirmed by noise level measurements. Spectra of the frequency noise ($1/f$) level for unimplanted sample and two different doses of implantation (5×10^{15} ions/cm² and 1×10^{16} ions/cm²) were presented in Fig. 1.

These spectra were taken at different time constant τ (frequency ranges $\tau \sim 1/f$) of the low noise amplifier. From Fig. 1, it could be seen that implantation doses have different effects for different frequency ranges. In the frequency range of 15–26 kHz (time constant of 6–10 μs) ion dose of 1×10^{16} ions/cm² gives better results than unimplanted sample, but in the range beyond 80 kHz (time constant lower than 2 μs), it produces greater noise compared to unimplanted sample. However, ion dose of 5×10^{15} ions/cm² shows the best results for the entire measuring range, suggesting that this dose of implantation induce a more homogeneous silicidation and the formation of Ti-Si phase with a lower concentration of crystal defects (after annealing). The lower concentration of point defects and dislocations and a more homogeneous silicide/silicon interface result in a lower frequency noise level of the analyzed structures. Also, previous results have shown [4] that the noise level was lowest for the samples implanted after annealing. This suggests that thermal treatment induce relaxation of crystal lattice and improvement of the crystal structure of the silicides. All of

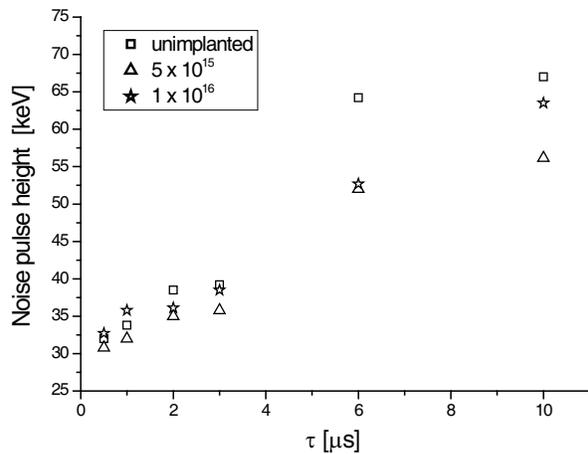


Figure 1: Frequency noise level for three TiN/Ti/Si samples.

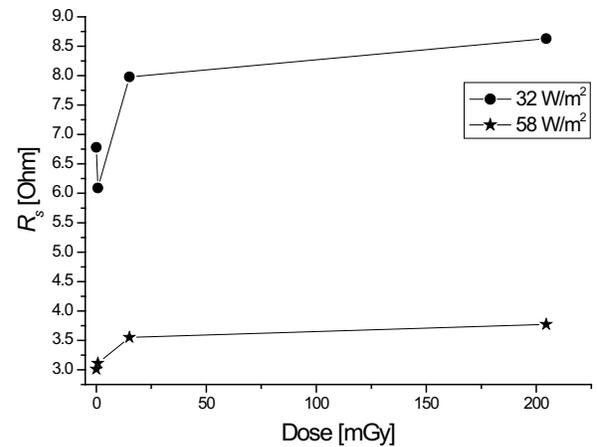


Figure 2: Dependence of the series resistance on doses for two illumination levels.

these methods lead to an improvement of electrical characteristics of silicides and devices based on silicides as contacts.

3.2. Irradiation of Solar Cells

Negative influence of radiation on electrical characteristics of the semiconducting devices is a well known and thoroughly investigated fact, especially when working in hostile conditions. Solar cells used in space are exposed to all kinds of radiation, both wave and particle. Their special design (surface to volume ratio) makes them susceptible to radiation damage, and improvement of electrical performance of such damaged cells is the aim of many experiments [8]. Radiation damage due to neutrons (heavy particles) is, as mentioned before, primarily connected to the displacement of silicon atoms from their lattice sites in the crystalline silicon solar cells, leading to destruction and distortion of local lattice structure and formation of defects. If, under the influence of neutrons, stable defects are made, they could, together with impurity atoms, donors and for example implanted atoms, form complex defects acting as recombination sites or traps, significantly decreasing minority carrier lifetime. This lifetime decrease produces degradation of electrical parameters of the cell, such as series resistance (R_s), output current and finally efficiency (η). High level of series resistance usually indicate the presence of impurity atoms and defects localized in the depletion region acting as traps for recombination or tunneling effects, increasing dark current of the cell. Moreover, shallow recombination centers in the vicinity of conducting zone enhance tunneling effect, further degrading output characteristics of the cell by increasing noise level (especially burst noise that is connected to the presence of excess current).

Such negative impact of neutron radiation was observed in this experiment for higher illumination level also, as could be seen in Fig. 2. But interesting phenomena — the decrease of series resistance, was observed for lower values of illumination. (Different behavior for different illumination level is due to the presence of finite series and parallel resistance in the cell.)

This decrease is very significant from the solar cell design standpoint because it indicates possible beneficial influence of low doses of irradiation, even with neutrons. It could be explained by the fact that during fabrication process of any semiconducting device, structural defects and impurities that were unavoidably made, produce tension in the crystal lattice. Low doses of radiation could act similarly to annealing, relaxing lattice structure and decreasing series resistance. Subsequently, this leads to lowering of noise level and an increase of the output current as shown in Fig. 3 (J_m — current in the maximum power point).

Finally, improvement of output characteristics after the first irradiation step for low illumination level is registered for the efficiency also, Fig. 4.

Although higher doses of neutron radiation undoubtedly have negative impact on the performance of solar cells, observed phenomena give possibilities for using radiation as a method for the improvement of solar cell characteristics.

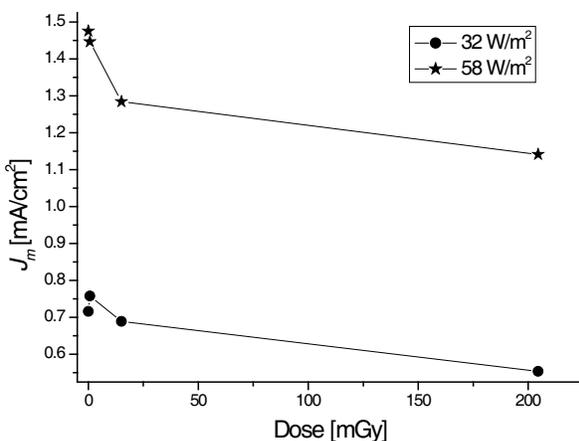
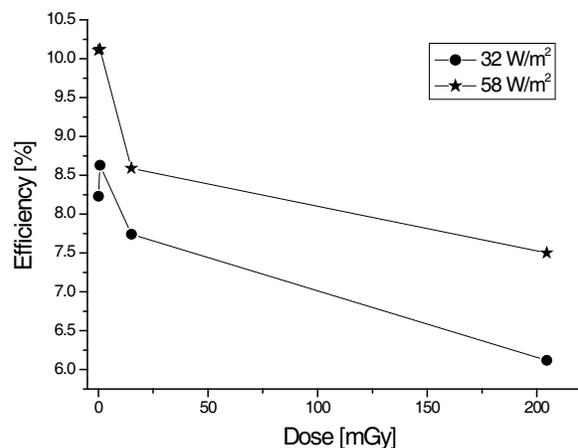
Figure 3: Dependence of J_m on the doses.

Figure 4: Dependence of the efficiency on doses.

4. CONCLUSION

Noise, as one of the most important limiting factor for energy resolution of detectors (in this case, solar cells) was investigated in this paper. Two different approaches for lowering noise in silicon solar cells were presented. First part of the paper was oriented to the frequency dependent $1/f$ noise in contacts, since recombination, trapping-detrapping of carriers and other interactions are more pronounced in the surface area. For that reason, silicides were proposed to be used as contacts, because of their good electrical characteristics and stability. Arsenic ion implantation of TiN-Ti-Si structure was used to further improve the performance of silicides as contacts, and to reduce their noise. It was established that both physical and electrical properties of used silicides are influenced by the implantation doses. As could be expected, higher doses result in degradation of electrical characteristics (via increasing noise level). But the results of frequency noise measurements indicate that ion implantation could successfully be applied in order to achieve a more homogeneous silicidation, if carefully optimized dose (in our case 5×10^{15} ions/cm²) was used. This dose of As⁺ ions was proved to be optimal for fabrication of low-resistivity and low-noise contacts.

Second approach investigated in this paper was lowering noise in the bulk of solar cells. It was shown that, though commonly referred to as a source of noise in semiconducting devices, radiation induced effects (interaction of neutrons with Si solar cells, in particular) could have in some cases positive effect on main electrical characteristics (R_s , J_m , η). Initial improvement of the characteristics observed for small doses of neutron radiation and low illumination level, indicates that there is a possibility of using irradiation for enhancement of the solar cells quality.

ACKNOWLEDGMENT

The Ministry of Science and Environmental Protection of the Republic of Serbia supported this work under contract 141046.

REFERENCES

1. Jayaweera, P. V. V., P. K. D. D. P. Pitigala, A. G. U. Perera, and K. Tennakone, "1/f noise and dye-sensitized solar cells," *Semiconductor Science Technology*, Vol. 20, L40–L42, 2005.
2. Jayaweera, P. V. V., P. K. D. D. P. Pitigala, M. K. I. Senevirante, A. G. U. Perera, and K. Tennakone, "1/f noise in dye-sensitized solar cells and NIR photon detectors," *Infrared Physics & Technology*, Vol. 50, 270–273, 2007.
3. Stojanović, M., C. Jeynes, N. Bibić, M. Milosavljević, A. Vasić, and Z. Milošević, "Frequency noise level of As ion implanted TiN-Ti-Si structures," *Nuclear Instruments and Methods B*, Vol. 115, 554–556, 1996.
4. Stojanović, M., A. Vasić, and C. Jeynes, "Ion implanted silicides studies by frequency noise level measurements," *Nuclear Instruments and Methods B*, Vol. 112, 192–195, 1996.
5. Hovel, J., *Semiconductors and Semimetals*, Academic Press, New York, 1975.
6. Holwes-Siedle, A. G. and L. Adams, *Handbook of Radiation*, Second Edition, Oxford University Press, 2002.

7. Messenger, G. C. and M. S. Ash, *The Effects of Radiation on Electronic Systems*, Van Nostrand Reinhold, New York, 1992.
8. Alexander, D. R., “Transient ionizing radiation effects in devices and circuits,” *IEEE Transaction on Nuclear Sciences*, Vol. 50, No. 3, 565–582, 2003.
9. Horiushi, N., T. Nozaki, and A. Chiba, “Improvement in electrical performance of radiation-damaged silicon solar cells by annealing,” *NIM A*, Vol. 443, 186–193, 2000.
10. Alurralde, M., M. J. L. Tamasi, C. J. Bruno, M. G. M. Bogado, J. Pla, J. F. Vasquez, J. Duran, J. Shuff, A. A. Burlon, P. Stoliar, and A. J. Kreiner, “Experimental and theoretical radiation damage studies on crystalline silicon solar cells,” *Solar Energy Materials & Solar Cells*, Vol. 82, 531–542, 2004.
11. Khan, A., M. Yamaguchi, Y. Ohshita, N. Dharmaraso, K. Araki, V. T. Khanh, H. Itoh, T. Ohshima, M. Imaizumi, and S. Matsuda, “Strategies for improving radiation tolerance of Si space solar cells,” *Solar Energy Materials & Solar Cells*, Vol. 75, 271–276, 2003.

Influence of Tube Volume on Measurement Uncertainty of GM Counter

Koviljka Stankovic, Predrag Osmokrovic, and Milos Vujisic
University of Belgrade, Serbia

Abstract— GM counter is used often in radiation detection since it generates strong signal which can be easily detected. Working principal of GM counter is based on ionization interaction of radiation with atoms and molecules of gaseous in counter's tube. Free electrons created as a result of the interaction, present in a counter's tube, become initial electrons, i.e., they start an avalanche process which is detected as a pulse of current. This current pulse is independent on energy imparted in gaseous which is the main difference between GM counter and majority of other radiation detectors. Dependence on incidence radiation energy, tube's orientation and reading system characteristics are labeled in the literature as the main sources of measurement uncertainty of GM counter. The aim of this paper is to determine influence of counter's tube volume on the measurement uncertainty of GM counter. Therefore, dependence of detecting pulse current formation on counter's tube size will be considered, both in radial and parallel geometry. Initiation and current pulse developing will be treated throughout elementary processes of electrical discharge as Markov processes, while the change of counter's tube volume will be treated throughout space-time enlargement law. Random variable "current pulse in the counter's tube" (i.e., electrical breakdown of the electrode configuration) will be considered statistically and based on it, appropriate theoretical distribution will be determined. Results obtained theoretically will be compared with appropriate experimental results obtained under well controlled laboratory conditions.

1. INTRODUCTION

According to the known characteristics of GM counter [1], by the process of detecting ionizing radiation, potential sources of uncertainty can be identified: detection dependence on energy and incident angle of radiation, counter dead time, reading system (by the resolution of instrument), instrument calibration errors, influence of background radiation, uncertainty arising from the measurement process (counting impulses), influence of overvoltage phenomenon in electronic devices (in their wire structures) generated by induction of overvoltage upon the electromagnetic rays which are a consequence of electromagnetic radiation in environment where the measurements are performed (this phenomenon is specially marked in urban environment). In essence, functioning of GM counter is based on self-sustained avalanche gaseous effect, and in that sense, energy of incident radiation determines the number of free, potentially initial electrons in counter's tube, so the energy contributes significantly to the stochastic response of the counter, in fact, to the statistical discharge time [2], and directly determines the nature of type A uncertainty. In a similar way, the angle of incident radiation contributes to type A uncertainty, because the number and position of free electrons depends on this angle, and that is specially marked in a tube with coaxial electric field [3]. Dead time of a counter is a source of type B or combined uncertainty, which depends on determining method. Determining dead time by recording pulses at counter's output can be conditionally arranged to type B uncertainty, and determining dead time by the two sources method has to be considered as a combined uncertainty, because the stochastic nature of radioactive decay has to be taken in consideration. Applied reading system is source of type B uncertainty and depends on resolution of counter's technological solution as well as the true value of measured variable (electrical discharge throughout counter's tube which is of analog nature) is symmetrically arranged through digital reading, in fact, is it uniformly distributed in interval between $n - 1/2$ and $n + 1/2$ digits [4]. Uncertainty due to instrument calibration is of type B and in almost all cases is a component of uncertainty arising from systematic effects. Background cosmic radiation is source of combined uncertainty and determination of this kind of uncertainty is the most difficult one, because of its fluctuation and energy structure. Contribution of background radiation to uncertainty can be decreased by applying anticoincidence protection and by background radiation correction, but it is never completely eliminated [5]. Uncertainty of counting impulses from radioactive source is of type A, because the deexcitation of a nucleus is completely random and nothing can determine the deexcitation moment, in other words, it is impossible to connect deexcitation with any measuring law — each process arising in that way is of stochastic

nature and is associated with Gaussian distribution (because of possible time balance of its occurrence) [6]. Minimization of electronic components, and even more the exposure of environment to electromagnetic radiation, lead to frequent occurrences of overvoltages within electronic devices of GM counters, which can, independently of the counter's tube, trigger the reading system and in that way cause type A uncertainty. Influence of this source of uncertainty can be decreased by applying the overvoltage protection of electronic devices (coordination of isolations at low voltage level) and/or by performing measurements in an area protected from electromagnetic radiation (over 100 dB protection).

2. THE TUBE VOLUME REDUCTION

Seeing that the probability of self sustained discharge occurrence is of geometrical nature, the tube volume reduction causes the reduction of Geiger discharge occurrence probability.

In order to describe this effect throughout mathematical modeling, it is suitable to present the tube volume reduction as a reduction of capacitor interelectrode gap, in two steps: 1 — n times reduction in the direction of z -axis and 2 — m times reduction in xy plane (electrode surface).

First step, i.e., determination of pulse number and its deviation, as a function of interelectrode gap reduction, is based on gas discharge development as Markov process [7]. Free electron becomes initial electron when it has been found within critical volume where it can take enough energy from electrical field to start an avalanche process.

Interelectrode gap reduction, followed by lowering of operational voltage and/or increasing the pressure, causes the critical volume reduction according to the similarity law [7, 8], where the number of electrons arised from primary ionization increases because of an increased gas density. The result of these two effects is less stochastic primary ionization position within a smaller interelectrode gap. Taking into account mentioned effects, the change of pulse number and its standard deviation, with n times reduction of interelectrode gap, can be expressed as [8, 9]:

$$I_{n,1} = \frac{I_1}{n^2} \quad (1)$$

$$\sigma_{n,1} = \frac{\sigma_1}{n^{2r}}, \quad (2)$$

where the parameter r is higher than 0.

Second step, i.e., the tube reduction in xy plane is based on presenting the starting system as a capacitor, in fact as a parallel connection of m identical capacitors which is equivalent to said capacitor. Non-breakdown probability, i.e., no Geiger discharge, in the starting system presupposes non-breakdown of all m subsystems, which means that the value of breakdown voltage U_m is higher than the value of operational voltage U . In this case of discrete random variable, with two possible events (breakdown and non-breakdown), the problem can be presented with constant voltage testing. Critical volume of each two-electrode system is proportional to the voltage applied, with no respect to the electrode configuration [10, 11]. Therefore the random number of breakdown, i.e., the number of pulses and its standard deviation of the reduced size tube in xy plane can be expressed as [8, 9]:

$$I_{1,m} = I_{1,1} + L(\beta) \left[1 - \frac{1}{m^\beta} \right] \frac{\sigma_{1,1}}{m^{\frac{1}{\beta}}} \quad (3)$$

$$\sigma_{1,m} = m^{\frac{1}{\beta}} \sigma_{1,1} \quad (4)$$

Summing the effects of a tube volume reduction n times in z -axes direction and m times in xy plane gives:

$$I_{n,m} = I_{1,1} \left(1 + \frac{1}{n^2} \right) + L(\beta) \left[1 - \frac{1}{m^\beta} \right] \frac{\sigma_{1,1}}{m^{\frac{1}{\beta}}} \quad (5)$$

$$\sigma_{n,m} = \sigma_{1,1} \left(\frac{1}{n^{2,2}} + m^{\frac{1}{\beta}} \right) \quad (6)$$

3. THE EXPERIMENT

Two type of tubes are used in the experiment. Type 1 tube was used for analyzing the influence of interelectrode gap on the measurement uncertainty of type A. This tube was made out of glass

and filled with He gas with pressure of 50 mbar where the cylindrical shape of electrodes ensured the homogeneous electrical field (plain-parallel configuration). The interelectrode gap had values of 0.1 mm, 0.3 mm, 0.5 mm, 1 mm, 3 mm, 5 mm i 10 mm where the ratio of interelectrode gap and cylindrical electrode radius kept constant. Type 2 tube was commercial tube made out in two variants (2a — coaxial geometry and 2b — plain-parallel geometry). This type was used for measuring the influence of size change in xy plane.

Experimental procedure for testing the influence of the tube reduction in z -axis direction on the measurement uncertainty consisted of the following steps:

1. Measuring the pulse number for 5 minutes.
2. Measuring the dead time of GM counter by two sources method.
3. Repeating steps 1 and 2 fifty times.
4. Measuring the background radiation for an hour.
5. One step of tube reduction in z axis direction.
6. Repeating the procedure.

Experimental procedure for testing the influence of the tube reduction in xy plane on the measurement uncertainty consisted of the following steps:

1. Measuring the pulse number for 5 minutes.
2. Measuring the dead time of GM counter by two sources method.
3. Repeating steps 1 and 2 fifty times.
4. Measuring the background radiation for an hour.
5. One step of tube reduction in xy plane.
6. Repeating the procedure.

Results of measurements were processed by a statistical calculation consisting of the following steps:

1. Pulse number correction due to background radiation and dead time of counter.
2. Formation of the statistical sample composed of 50 values of the random variable “mean value of pulse number” and appropriate statistical sample composed of 50 values of the random variable “counter’s dead time” for all series of measurements (for all sizes of the tube used).
3. Using of Chauvenet’s criterion for rejecting spurious measurement results.
4. U-test with a 5% significance level was employed to establish whether random variables belonged to the same random variable.
5. χ^2 -test and graphical test were applied for testing the random variables on belonging to Normal, Weibull and double-exponential distribution.
6. Determination of measurement uncertainty of type A for statistical samples: comparison of quantitatively and qualitatively obtained values of measurement uncertainty of type A, determined experimentally with theoretically expected values according to Enlargement Law [8, 9].

4. EXPERIMENTAL RESULTS

Figures 1 and 2 show results obtained for mean value and standard deviation of detected pulse number with respect to counter’s tube reduction, where an Americium source is used. Figs. 3 and 4 show results obtained for mean value and standard deviation of detected pulse number with respect to counter’s tube reduction, where a Caesium source is used.

The agreement of experimental and theoretical results has been achieved for mean value of pulse number (Figs. 1 and 3) where the applicability of the Enlargement Law (reduction, in our case) has been approved. The discrepancy between the theoretically expected effects of reduction (or enlargement) and the results obtained experimentally can be noticed in Figs. 2 and 4 where standard deviation of pulse number is shown.

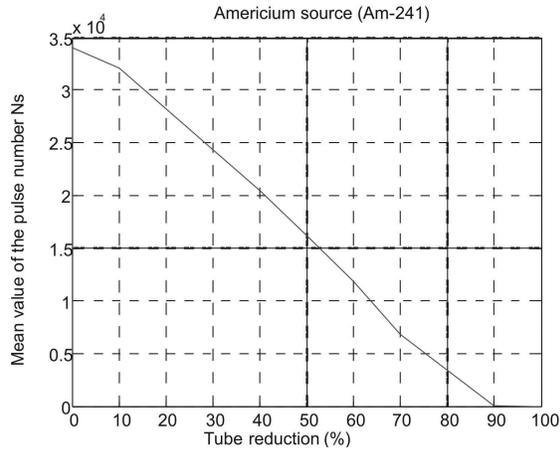


Figure 1: Detected mean value pulse number using an Americium source.

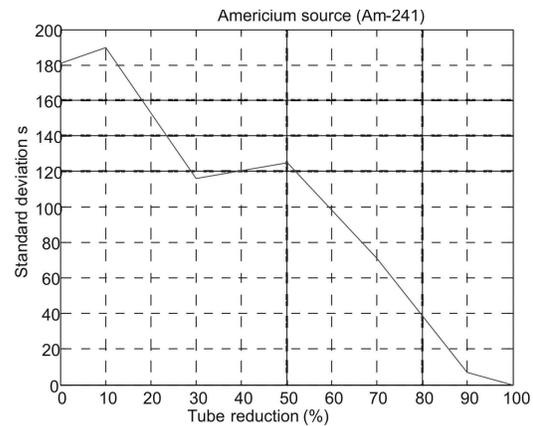


Figure 2: Standard deviation of the pulse number using an Americium source.

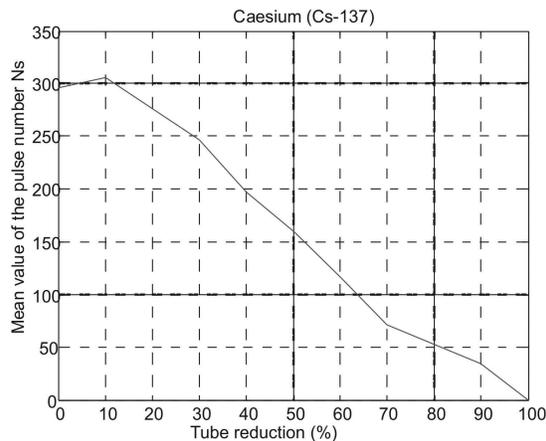


Figure 3: Detected mean value pulse number using a Caesium source.

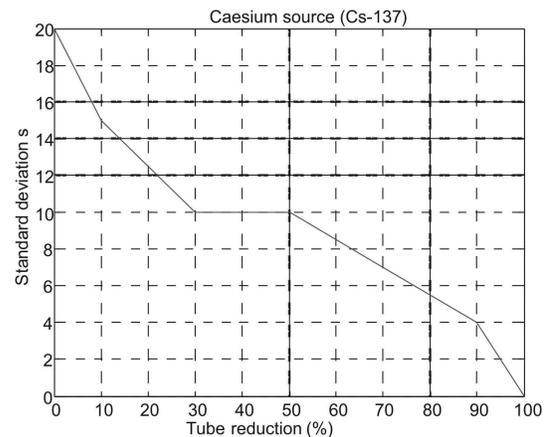


Figure 4: Standard deviation of the pulse number using a Caesium source.

5. CONCLUSION

As mentioned above, the agreement of experimental and theoretical results has been achieved for mean value of pulse number detected, where the discrepancy between the theoretically expected standard deviation and the experimentally obtained values has been found. This could be the consequence of the radiation energy emitted from the source used and the angle of the incidence energy which has been detected. In further investigations, the effects of incident radiation energy and angle on measurement uncertainty are to be included, especially in the vicinity of a counter's discrimination level.

REFERENCES

1. Knoll, G. F., *Radiation Detection and Measurement*, John Wiley and Sons, New York, USA, 1979.
2. Osmokrović, P., "Mechanism of electrical breakdown of gases at very low pressure and inter-electrode gap values," *IEEE, Mechanism of Electrical Breakdown Plasma Science*, Vol. 21, No. 6, 645–654, 1993.
3. Stanković, K. and M. Vujisić, "Influence of radiation energy and angle of incidence on the uncertainty measurements by GM counters," *Nuclear Technology & Radiation Protection*, Vol. XXIII, No. 1, 41–43, 2008.
4. IAEA Safety Reports Series, No. 16, Calibration of Radiation Protection Monitoring Instruments, www.ns.iaea.org/standards/documents/.

5. Stanković, K., D. Arandić, D. Lazarević, and P. Osmokrović, “Expanded and combined uncertainty in measurements by GM counters,” *Nuclear Technology & Radiation Protection*, Vol. XXII, No. 2, 64–70, 2007.
6. Burcham, W. E., *Nuclear Physics — An Instruction*, Longman Group Limited, 1972.
7. Osmokrović, P., T. Živić, B. Lončar, and A. Vasić, “The validity of the general similarity law for electrical breakdown of gases,” *Plasma Sources Science and Technology*, Vol. 15, No. 4, 703–713, 2006.
8. Stanković, K., P. Osmokrović, Č. Doličanin, M. Vujisić, and A. Vasić, “Time enlargement law for gas pulse breakdown,” *Plasma Source Science and Technology*, Vol. 18, No. 2, 025028, 2009.
9. Stanković, K., M. Pesic, P. Osmokrović, and M. Vujisić, “Surface-time enlargement law for gas breakdown,” *IEEE Transactions on Dielectrics and Electrical Insulation*, Vol. 15, No. 4, 994–1005, 2008.
10. Osmokrović, P., “Electrical breakdown of SF6 at small values of product pd ,” *IEEE Transactions on Power Delivery*, Vol. 4, No. 4, 2095–2100, 1989.
11. Osmokrović, P., N. Arsić, Z. Lazarević, and Z. Kušić, “Numerical and experimental design of three-electrode spark gap for synthetic test circuits,” *IEEE Transactions on Power Delivery*, Vol. 9, No. 3, 1444–1451, 1994.

Monte Carlo Simulations of Proton and Ion Beam Irradiation on Titanium Dioxide Memristors

Ćemal Dolićanin¹, Bratislav Irićanin², Miloš Vujisić², and Predrag Osmokrović²

¹University of Novi Pazar, Serbia

²Faculty of Electrical Engineering, University of Belgrade, Serbia

Abstract— Effects of titanium dioxide memristor exposure to proton and ion beams are investigated. A memristor model assuming ohmic electronic conduction and linear ionic drift is used for the analysis. Simulations of particle transport suggest that radiation induced oxygen ion/oxygen vacancy pairs can influence the device's operation by lowering both the mobility of the vacancies and the resistance of the stoichiometric oxide region. These radiation induced changes affect the current-voltage characteristic and state retention ability of the memristor.

1. INTRODUCTION

Memristor is a one-port circuit element that maintains a nonlinear relationship between time integrals of the voltage across its terminals and the current running through it. In mathematical terms, this defining property of a memristor can be expressed in a differential form as:

$$v(t) = M(w)i(t) \quad (1)$$

where M is called *memristance*, and w is a state variable that, in turn, depends on the time integral of the current, i.e., on the amount of charge q that has passed through the device. It is from this nonlinear relationship that the characteristic properties of memristors ensue: the hysteretic features of the i - v curve (namely the single and double loops, with segments of apparent negative differential resistance) and the ability to operate as a switch by holding or “remembering” the value of resistance (which gives memristor its name, as a portmanteau for “memory resistor”). Since, in general, there are many ways for such a nonlinear relationship to hold, there is no such thing as a generic memristor. The nonlinearity can in principal be achieved through quantities (state variables) specific of each separate realization of the memristor.

Equation (1) suggests that memristance can be regarded as the *effective resistance* of a memristor. Furthermore, a memristor does not introduce a phase shift between current and voltage at zero crossings, i.e., $i = 0$ if and only if $v = 0$. It is therefore a purely dissipative element, much like the resistor. In the trivial case of constant M , when nonlinearity is removed, memristor indeed reduces to a resistor. The physical unit for memristance is Ohm.

In the nontrivial case of M depending on q via w , memristor has such an i - v curve that no combination of nonlinear resistors, capacitors or inductors can reproduce. This inability to duplicate the properties of a memristor with the other passive circuit elements is what lead to it being seen as the forth fundamental passive two-terminal circuit element, when it was first hypothesized back in 1971 [1].

A memristor can remember a state by freezing the value of the state variable w , which is not necessarily zero when $v = 0$. The value of w at this point depends on the history of the voltage change, i.e., on the shape, amplitude and frequency of the driving voltage.

A physical component demonstrating memristance was not created until 2008, when a two-terminal (one-port) realization of a memristor, based on a thin film of titanium dioxide, was finally constructed by the Hewlett-Packard Laboratories [2]. Operation of this device relies on the distribution of oxygen vacancies in the oxide, which can be perturbed through atom displacement interactions when the component is exposed to proton or ion beams.

2. THE TITANIUM DIOXIDE MEMRISTOR

The constructed memristor is composed of a titanium dioxide thin film between two platinum electrodes. The oxide layer further consists of a high-resistance stoichiometric TiO_2 layer and a low-resistance oxygen deficient TiO_{2-x} layer, in which oxygen vacancies serve as electron donors. Even a small nonstoichiometry of 0.1% in TiO_{2-x} is equivalent to $5 \cdot 10^{19}$ donors/cm³ and has a very strong effect on the electronic conductivity. Memristance is achieved in this stacked TiO_2 - TiO_{2-x} structure through a specific coupling of electronic and ionic transport. Oxygen vacancies

act as mobile +2-charged ions, which can drift in the electric field created by a voltage applied to the device's terminals, shifting the boundary between the high-resistance and low-resistance layers. The charge-flow dependent state variable is the thickness w of the conducting oxygen-poor layer.

Since the concentration of oxygen vacancies in the TiO_{2-x} region is large, the transfer of some of the vacancies into the stoichiometric layer has very little effect on the electronic conductivity of the oxygen-poor region. On the other hand, the electronic conductivity of the initially stoichiometric material increases dramatically, since it is going from a state in which there are no vacancies to the one in which there are some.

The conduction mechanism in the titanium dioxide memristor involves electrons from the electrodes tunneling through the energy barrier existing at the TiO_2 /metal electrode interface. Both the height and the width of the tunnel barrier are dependent on the size of the oxygen deficient region within the memristor, diminishing as w becomes larger. Migration of oxygen vacancies, on the other hand, is best described as weak ionic conduction in a solid [3]. The coupling of these two drifts rests on the fact that both charged species involved in memristor operation (electrons and oxygen vacancies) move by virtue of the same electric field present across the low-resistance region [4].

An idealized physical model of the titanium dioxide memristor used in this paper assumes ohmic electronic conduction and linear ionic drift. Total resistance of the device is determined as a series connection of the highly resistive stoichiometric layer and the conducting oxygen-poor layer. Ohm's law relation between voltage and current is then:

$$v(t) = \left(R_{\text{ON}} \frac{w(t)}{D} + R_{\text{OFF}} \left(1 - \frac{w(t)}{D} \right) \right) i(t) \quad (2)$$

where $w(t)$ is the size of the oxygen-poor layer, while R_{ON} and R_{OFF} are the resistances of the oxygen-poor and the stoichiometric region respectfully, given for the full length D of the device. Based on the said assumptions, the following expression for doped region size is obtained:

$$w(t) = w_0 + \frac{\mu_{\text{OV}} R_{\text{ON}}}{D} \int_0^t i(\tau) d\tau = w_0 + \frac{\mu_{\text{OV}} R_{\text{ON}}}{D} q(t) \quad (3)$$

where $w(t=0) = w_0$ is the initial size of the oxygen deficient region, μ_{OV} is the mobility of oxygen vacancies in titanium dioxide, and $q(t)$ is the electronic charge that has passed through the device, for which $q(t=0) = 0$. Inserting Equation (2) into Equation (1), and comparing it to Equation (1), memristance is obtained as:

$$M(q) = R_0 - \Delta R \frac{\mu_{\text{OV}} R_{\text{ON}}}{D^2} q(t) \quad (4)$$

where $R_0 = R_{\text{ON}}(w_0/D) + R_{\text{OFF}}(1 - w_0/D)$ is the effective resistance at $t = 0$, and $\Delta R = R_{\text{OFF}} - R_{\text{ON}}$. Memristance is thus dependent directly only on the charge which has passed through the memristor. If the applied voltage is removed, the memristor "remembers" its last state, i.e., the value of total resistance at the moment of voltage suspension [2].

Since the mobility of oxygen vacancies in titanium dioxide is low ($\mu_{\text{OV}} \sim 10^{-10} \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$), memristive effects are appreciable only when the memristor size is nano-scale, which is illustrated by the inverse-square dependence on total film thickness ($M \sim 1/D^2$) in Equation (4). Only in a film this thin is the amount of time required to drift enough oxygen vacancies into or out of the TiO_2 region to substantially change its conductivity short enough. Moreover, for a given D , memristive behaviour is expressed only when $\Delta R \gg R_0$.

Equation (1) can now be rewritten as:

$$v(t) = M(q) \frac{dq}{dt} \quad (5)$$

Inserting Equation (4) into Equation (5), solving for $q(t)$ and substituting it back in (5), yields the i - v characteristic of an ideal titanium dioxide memristor:

$$i(t) = \frac{v(t)}{R_0 \sqrt{1 \mp 2\Delta R \frac{\mu_{\text{OV}} R_{\text{ON}}}{D^2 R_0^2} \int_0^t v(\tau) d\tau}} \quad (6)$$

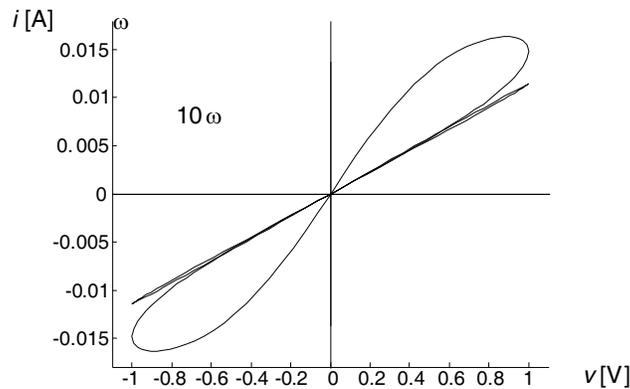


Figure 1: Current-voltage characteristic of a titanium dioxide memristor, for which ohmic electronic conduction and linear ionic drift of oxygen vacancies are assumed. The applied voltage is $v(t) = v_0 \sin(\omega t)$, with $v_0 = 1$ V and $\omega = \pi/10$ s $^{-1}$. Other parameters are: $R_{\text{ON}} = 100$ Ω , $R_{\text{OFF}} = 16$ k Ω , $D = 60$ nm, $w_0 = 30$ nm, $\mu_{\text{OV}} = 10^{-10}$ cm 2 s $^{-1}$ V $^{-1}$. The dotted plot is for a ten-time higher frequency of the sinusoidal voltage.

The minus sign in the denominator of Equation (6) applies when the oxygen-poor region is expanding, while the plus sign corresponds to the shrinking of this region.

Typical memristor i - v curve for a sinusoidal driving voltage, obtained theoretically from Equation (6) as a 2D parametric plot, is shown in Fig. 1. It has a form of a double-loop hysteresis, with segments of negative differential resistance corresponding to the intervals during which $w(t)$ is increasing while $v(t)$ is already in recess, but still of same polarity. A similar hysteresis is obtained for any symmetrical ac voltage applied to the memristor. The hysteresis is observed only for small-amplitude (~ 1 V) low-frequency (~ 1 Hz) voltages, for which w never reaches either of the limiting values (0 or D), i.e., the low resistance oxygen-poor layer never stretches across the length of the device, nor vanishes completely. For high-frequency low-amplitude ac voltages, the size of the oxygen-deficient layer barely changes for the duration of the voltage half-period, making the effective resistance of the memristor nearly constant and reducing the i - v hysteresis to a straight line, which is also demonstrated in Fig. 1.

The experimental i - v curve, obtained from measurements performed on real titanium dioxide memristors, does not differ much from the one derived for the adopted model. Nonlinearities in ionic transport and electron tunneling at the metal/oxide interfaces cause certain nonlinearities in the i - v curve, without changing its overall hysteretic character.

If w reaches either of the two boundaries, it remains constant until the voltage polarity is reversed. When boundary conditions such as these are introduced, the device is more appropriately referred to as a *memristive system*, that acts as a true memristor, one for which Equation (6) holds, only as long as w is within the interval (0, D) [5]. The value of w can be pushed to one of the limits either by large applied voltages or by long times under same polarity bias. Boundary states differ greatly in resistance, thus forming the basis of memristor bipolar switching. If the voltage across memristor terminals is suddenly suspended, the value of memristance is frozen and stays unchanged while there is no bias. Long state lifetimes and fast switching observed in titanium dioxide memristors make them potential candidates for future non-volatile RRAMs (Resistive Random-Access Memories), based on the so called crossbar architecture. This paper examines the influence of proton and ion beam irradiation on the shape of the titanium dioxide memristor i - v hysteresis, and on state retention when memristor is used as a switch.

3. RESULTS OF RADIATION TRANSPORT SIMULATION

Monte Carlo simulations of proton and ion beams traversing the Pt-TiO $_2$ -TiO $_{2-x}$ -Pt structure were performed in the TRIM part of the SRIM software package [6].

Figure 2(a) shows the trajectories of one thousand 10 keV protons traversing the 60 nm thick memristor structure. The incident proton beam is perpendicular to the surface of the left platinum electrode. The thicknesses of the layers along the horizontal axis are as follows: 5 nm platinum layer, 25 nm stoichiometric TiO $_2$ layer, 25 nm oxygen deficient TiO $_{2-x}$ layer ($x = 0.05$), and another 5 nm platinum layer. Total length of the titanium dioxide film is then $D = 50$ nm. Fig. 2(b) presents the distribution of oxygen ions displaced by the impinging protons. The default value of oxygen

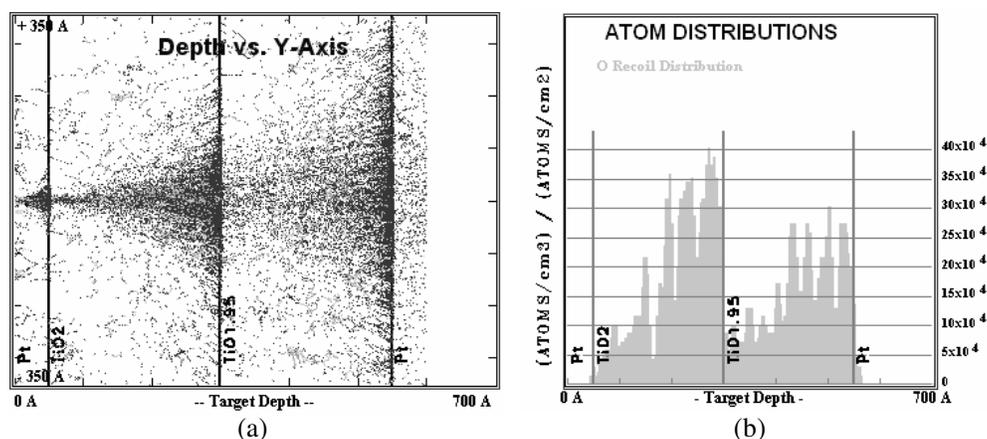


Figure 2: Simulation results for a 10 keV proton beam (1000 protons) incident perpendicularly on a Pt-TiO₂-TiO_{2-x}-Pt stack, with the total thickness of 60 nm. (a) Particle and ion tracks. (b) Distribution of displaced oxygen atoms. Target depth axis is in units of angstrom.

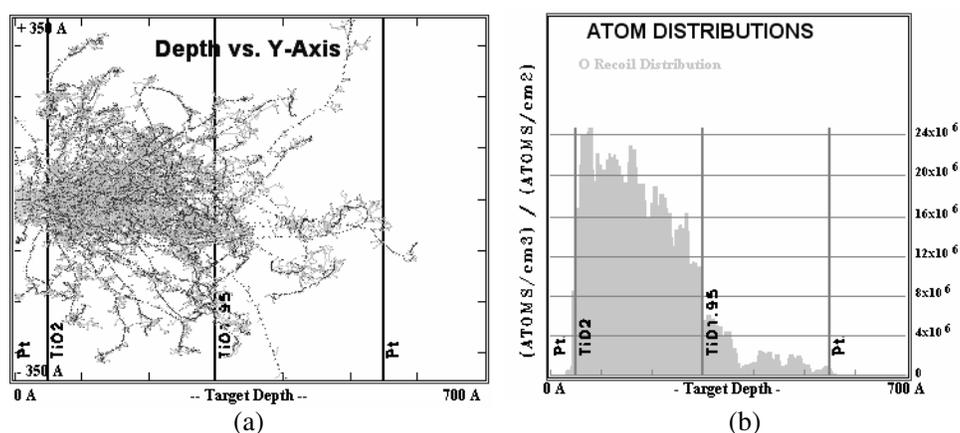


Figure 3: Simulation results for a 10 keV carbon ion beam (100 ions) incident perpendicularly on a stacked Pt-TiO₂-TiO_{2-x}-Pt structure, with the total thickness of 60 nm. (a) Ion tracks. (b) Distribution of displaced oxygen atoms. Target depth axis is in units of angstrom.

atom threshold displacement energy in titanium dioxide provided by SRIM was changed to a value obtained by a molecular dynamics simulation study for the rutile modification of TiO₂ [7].

Figure 3 provides an example of a 10 keV carbon ion beam traversing the same memristor structure, with equal thicknesses of the TiO₂ and the TiO_{1.95} layer.

4. DISCUSSION

As Monte Carlo simulations of proton and ion transport show, these radiations can cause the generation of a significant amount of oxygen ion/oxygen vacancy pairs in both the high- and the low-resistance layers of titanium dioxide. Whereas the electronic conductivity of the low-resistance oxygen-poor region is little affected by the appearance of additional vacancies, the effect on the conductance of the stoichiometric vacancy-free TiO₂ region can be considerable. Radiation induced emergence of oxygen vacancies in the stoichiometric region can cause its resistance R_{OFF} to drop, disrupting the R_{OFF}/R_{ON} ratio of the memristor. The change of the resistance ratio affects the memristor $i-v$ characteristic, through quantities R_0 and ΔR in Equation (6). The effect that the decrease of R_{OFF} has on the memristor $i-v$ curve is illustrated in Fig. 4.

Oxygen ions O²⁻ produced by radiation in the stoichiometric layer can become interstitial atoms, or migrate in the electric field. If the amplitude of the applied voltage is high enough, oxygen ions may reach one of the electrodes, where they can form O₂ gas and cause deformation of the oxide/metal interface, leading to permanent disruption of memristor operation [8]. While traversing the oxygen-deficient layer, some oxygen ions may recombine with the existing vacancies.

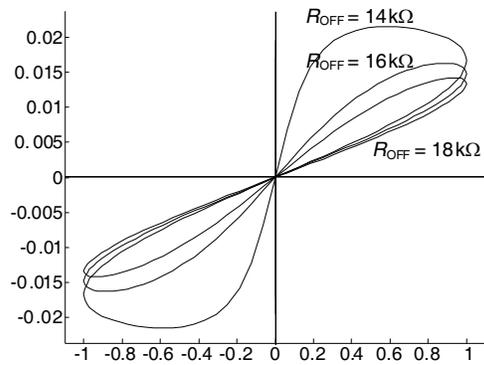


Figure 4: Current-voltage curves of a titanium dioxide memristor, plotted for three different values of the stoichiometric region resistance: $R_{\text{OFF}} = 18 \text{ k}\Omega$, $16 \text{ k}\Omega$, and $14 \text{ k}\Omega$. The applied sinusoidal voltage and all other parameters are the same as for Fig. 1. The decrease in R_{OFF} is caused by radiation induced emergence of oxygen vacancies in the stoichiometric TiO_2 region.

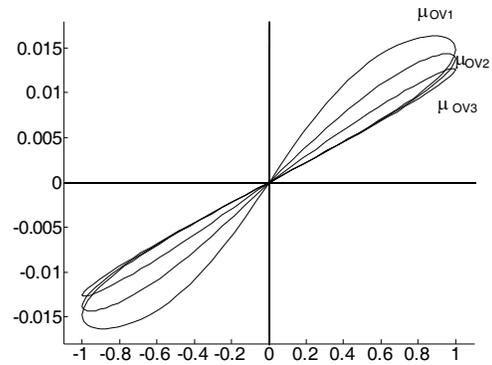


Figure 5: Current-voltage curves of a titanium dioxide memristor, plotted for three different values of the oxygen vacancy mobility: $\mu_{\text{OV}1} = 10^{-10} \text{ cm}^2\text{s}^{-1}\text{V}^{-1}$, $\mu_{\text{OV}2} = 0.8 \cdot 10^{-10} \text{ cm}^2\text{s}^{-1}\text{V}^{-1}$, and $\mu_{\text{OV}3} = 0.5 \cdot 10^{-10} \text{ cm}^2\text{s}^{-1}\text{V}^{-1}$. The applied sinusoidal voltage and all other parameters are the same as for Fig. 1. The decrease in μ_{OV} is caused by the radiation induced presence of oxygen ions and atoms in the oxide.

The presence of oxygen ions and atoms can also reduce the mobility of oxygen vacancies μ_{OV} [9]. According to Equation (6), a decrease of μ_{OV} affects the memristor i - v hysteresis, as shown by example plots in Fig. 5.

The specific switching functionality of a memristor rests on a high $R_{\text{OFF}}/R_{\text{ON}}$ ratio, which enables two boundary states to be unambiguously distinguishable by a read voltage signal, and on the ability to hold a state at zero bias. Since for the highly conducting boundary state, corresponding to $w = D$, the low-resistance region stretches across the whole of the oxide, the radiation produced change of R_{OFF} has no effect on state retention. The high-total-resistance state is, however, susceptible to change when exposed to proton or ion radiation. This state, corresponding to $w \approx 0$, is characterized by a diminished or non-existent oxygen-poor region, with the total memristor resistance approximately equal to R_{OFF} . The decrease of R_{OFF} caused by irradiation can therefore perturb this state, resulting in an error at readout.

5. CONCLUSION

Exposure of a titanium dioxide memristor to proton and ion beams can influence the device's operation in several ways. Significant generation of oxygen ion/oxygen vacancy pairs in the oxide is to be expected, as suggested by Monte Carlo simulations of particle transport. Radiation induced appearance of oxygen vacancies in the stoichiometric TiO_2 layer can cause its resistance to drop, producing counter-clockwise rotation of the memristor i - v curve and a larger swing in its double-loops. The presence of oxygen ions and atoms displaced by the radiation can reduce the mobility of oxygen vacancies, causing the memristor i - v curve to rotate clockwise. When memristor is operated as a switching element of a non-volatile memory, e.g., within a crossbar array, the high-total-resistance state, characterized by a diminished oxygen-poor region, can be perturbed by irradiation and result in an erroneous readout. Finally, if the displaced oxygen ions reach the platinum electrodes, they can form O_2 gas and cause permanent disruption of memristor functionality.

ACKNOWLEDGMENT

The ministry of Science and Environmental Protection of the Republic of Serbia supported this work under contract 141046.

REFERENCES

1. Chua, L. O., "Memristor — The missing circuit element," *IEEE Trans. Circuit Theory*, Vol. 18, 507–519, 1971.
2. Strukov, D. B., G. S. Snider, D. R. Stewart, and R. S. Williams, "The missing memristor found," *Nature*, Vol. 453, 80–83, 2008.

3. Strukov, D. B. and R. S. Williams, “Exponential ionic drift: fast switching and low volatility of thin-film memristors,” *Appl. Phys. A — Mater. Sci. Process.*, Vol. 94, 515–519, 2009.
4. Strukov, D. B., J. L. Borghetti, and R. S. Williams, “Coupled ionic and electronic transport model of thin-film semiconductor memristive behavior,” *Small*, Vol. 5, 1058–1063, 2009.
5. Chua, L. O. and S. M. Kang, “Memristive devices and systems,” *Proc. IEEE*, Vol. 64, 209–223, 1976.
6. Ziegler, J. F., J. P. Biersack, and M. D. Ziegler, *SRIM (The Stopping and Range of Ions in Matter)*, available online: <http://www.srim.org>.
7. Thomas, B. S., N. A. Marks, L. R. Corrales, and R. Devanathan, “Threshold displacement energies in rutile TiO₂: A molecular dynamics simulation study,” *Nucl. Instr. and Meth. B*, Vol. 239, 191–201, 2005.
8. Yang, J. J., M. D. Pickett, X. Li, D. A. A. Ohlberg, D. R. Stewart, and R. S. Williams, “Memristive switching mechanism for metal/oxide/metal nanodevices,” *Nat. Nanotechnol.*, Vol. 3, No. 7, 429–433, 2008.
9. Hendry, E., M. Koeberg, J. Pijpers, and M. Bonn, “Reduction of carrier mobility in semiconductors caused by charge-charge interactions,” *Phys. Rev. B*, Vol. 75, 233202-1–233202-4, 2007.

Influence of Irradiation on Semiconductor and Gas-filled Diodes for Over-voltage Protection

Radeta Maric¹, Miladin Jurosevic², Gvozden Ilic¹, and Predrag Osmokrović³

¹Electric Power Industry of Serbia (EPS), Serbia

²Alumina Factory, Birač, Zvornik, Republic of Srpska, Bosnia and Herzegovina

³University of Belgrade, Serbia

Abstract— The wide-spread use of semiconductor and gas-filled diodes for non-linear over-voltage protection results in a variety of possible working conditions. It is therefore essential to have a thorough insight into their reliability in exploitation environments which imply exposure to ionizing radiation. The aim of this paper is to investigate the influence of irradiation on over-voltage diode characteristics, by exposing the diodes to Californium-252 combined neutron/gamma radiation field. Irradiation of semiconductor over-voltage diodes causes severe degradation of their protection characteristics. On the other hand, gas-filled over-voltage diodes exhibit a temporal improvement of performance. Results are presented with the accompanying theoretical interpretations of the observed changes in over-voltage diode behavior, based on the interaction of radiation with materials constituting the diodes.

1. INTRODUCTION

Over-voltage is a rather common occurrence in all electronic circuits, which makes efficient over-voltage protection a primary design requirement. An efficient over-voltage protection has two aspects: Protection of integrity (no permanent damage of the protected device) and maintainance of operational functionality (operation reliability in the event of an over-voltage). Both power systems (energy generation, transmission and distribution) and low-voltage (electronic) systems, are susceptible to over-voltages. The extent to which an electronic component can withstand a temporary over-voltage without damage is reduced significantly as components are miniaturized [1, 2].

Over-voltage protection components can generally be divided into linear and nonlinear ones. The linear group includes capacitors, coils, resistors, or their combinations as filters. Semiconductor over-voltage diodes (known also as transient voltage suppression diodes) and gas-filled over-voltage diodes (also called gas-filled surge arresters) fall into the group of nonlinear over-voltage protection components. In practice, various hybrid schemes combining the linear and non-linear components are often used.

Semiconductor over-voltage diodes provide a way to increase immunity of a circuit to electromagnetic interference (EMI) and electrostatic discharge (ESD). Most integrated circuits (ICs) contain internal surge protection circuits that function well at preventing ESD failures that occur in assembly. However, they are often inadequate for protecting against surge events that occur in normal product usage. The surge ability of an over-voltage diode is directly related to its size, and external devices are typically ten times larger than the internal IC over-voltage devices. External over-voltage diodes provide a higher level of surge protection because it is not practical for an IC to incorporate large protection devices. In addition, the internal protection circuit of most ICs is designed to handle only a few ESD events, while an external over-voltage device provides immunity for an indefinite amount of surges.

Gas-filled over-voltage diodes are largely used for protecting circuits in telecommunications (where EMI events can be induced in many ways, including lightning), as well as in high-voltage engineering (where switching over-voltages may arise as a consequence of energy redirection within a power system).

Semiconductor and gas-filled over-voltage diodes are complementary with respect to surge capability and response time. While semiconductor diodes are effective at currents up to 50 A, gas-filled diodes can withstand currents as high as 20 kA. On the other hand, semiconductor over-voltage diodes have a subnanosecond response, while the response time of gas-filled diodes can go up to 5 μ s [3, 4].

Use of modern electronic and electrical devices in conditions which imply exposure to ionizing radiation, e.g., at nuclear plants, in military industry and space technology, brings up the issue of over-voltage diode radiation hardness. The aim of this paper is to investigate the influence of radiation exposure on over-voltage diode characteristics. Only permanent effects, which are manifested even upon cessation of irradiation, are considered.

2. THE EXPERIMENT

Measuring equipment for both semiconductor and gas-filled over-voltage diodes consisted of a current source (with a 3000 V maximum voltage), a digital oscilloscope, a *dc* high-voltage power supply and a personal computer. All measuring instruments were protected from electromagnetic interference by electromagnetic shielding. The experimental procedure was fully automated. This approach assured very high accuracy of measurement and good repeatability of results. Specialized PC-based control software (HP-IB or IEEE488 protocol) was developed to provide overall experiment sequencing, measurement and data acquisition.

Radiation induced changes of diodes' characteristics were investigated by exposing them to a combined neutron/gamma radiation field of the Californium-252 source. Californium ^{252}Cf isotope, encapsulated in the form of Cf_2O_3 , was used. The mass of the used ^{252}Cf radionuclide was 2.265 μg , its specific neutron and gamma emission rates $2.34 \times 10^6 (\mu\text{g s})^{-1}$ and $5.3 \times 10^9 (\mu\text{g s})^{-1}$, respectively. Average neutron energy of the ^{252}Cf source is 2.14 MeV, and average gamma photon energy is 0.88 MeV.

2.1. Semiconductor Over-voltage Diodes

Commercially available avalanche silicon diodes, with 250 V nominal turn-on voltage, were used for the experiments. Diodes were all produced by a single manufacturer, with identical nominal characteristics, which resulted in low statistical dispersion of the obtained results. Each series of measurements has been performed on a sample consisting of 50 diodes. Diode characteristics presented in the paper are based on sample mean values.

The influence of ionizing radiation on semiconductor over-voltage diode operation was investigated by monitoring the volt-ampere characteristic, the volt-ohm characteristic, breakdown voltage, and the nonlinearity coefficient α , defined as: $\alpha = \log(I_2/I_1)/\log(U_2/U_1)$, where (U_1, I_1) and (U_2, I_2) are points taken from the volt-ampere curve. Experiments consisted in applying double exponential current pulses (13 A, 8/20 μs) to the diodes and recording voltage response at the trailing edge of the current pulse, while varying the absorbed dose. Prior to each series of measurements, over-voltage diodes were conditioned with 25 breakdowns. A thirty-second pause between each two successive measurements was introduced. The volt-ohm characteristic was calculated from the volt-ampere curve.

Semiconductor diodes were exposed to three different levels of neutron absorbed dose in silicon (D_n), given in Table 1 in ascending order. Each dose level is marked by a number (n) shown in the first column. The pre-irradiation state is designated by $n = 0$.

2.2. Gas-filled Over-voltage Diodes

Commercially available argon-filled over-voltage diode, with 750 V nominal *dc* breakdown voltage, was used for the experiment. Radiation induced changes in the diode were examined by monitoring the “*dc* breakdown voltage” and “pulse breakdown voltage” random variables, as well as the pulse (volt-second) characteristic, as the absorbed dose increased. Measurement of the gas-filled diode volt-second characteristics was based on the Area law. The method consisted in applying a series of 50 double exponential voltage pulses (1.2/50 μs) before and immediately after exposure to radiation, with a thirty-second pause between consecutive pulses.

Gas-filled over-voltage diode was subjected to three neutron absorbed doses in argon: 3.79 Gy, 6.14 Gy, and 11.17 Gy. Gamma component of the radiation field influenced the electric properties of the gas-filled diode only in the course of irradiation. Moreover, for neutron and gamma ray energies of the used ^{252}Cf source, according to NIEL hypothesis, displacement damage cross section of the neutron component is much larger than for the corresponding gamma-ray component [5, 6]. These facts allowed only neutron fluence to be considered.

The influence on the “*dc* breakdown voltage” and “pulse breakdown voltage” random variables was tested measuring 1000 values of each. The discharge energy (current) was maintained constant

Table 1: Values of neutron absorbed dose (D_n) used for irradiation.

n	D_n [cGy]
1	24.85
2	49.7
3	74.62

during the measurement series. Results obtained in the measurement series were divided into ten groups of 50 successive values of the breakdown voltage. Thereafter, each group of the results was tested statistically. U-test with 5% significance level was used to determine whether measurements in a single group belonged to the same random variable. Applying chi-square and Kolmogorov tests, measured values of the breakdown voltage (within one group of measurements) were tested with respect to the type of statistical distribution (normal, exponential, double-exponential, and Weibull) [7].

The influence of radiation on the pulse characteristic was probed in the same manner, using the Area law. According to this law, there has to be a constant geometrical area formed in the voltage-time plane between the pulse-voltage time-shape and the *dc* breakdown voltage level, for pulse (dynamic) breakdown to occur. Since the area delimited by pulse and *dc* breakdown voltages is a specific property of a particular gas, and doesn't depend on the applied voltage, knowing the size of this area suffices for the calculation of pulse (volt-second) characteristic.

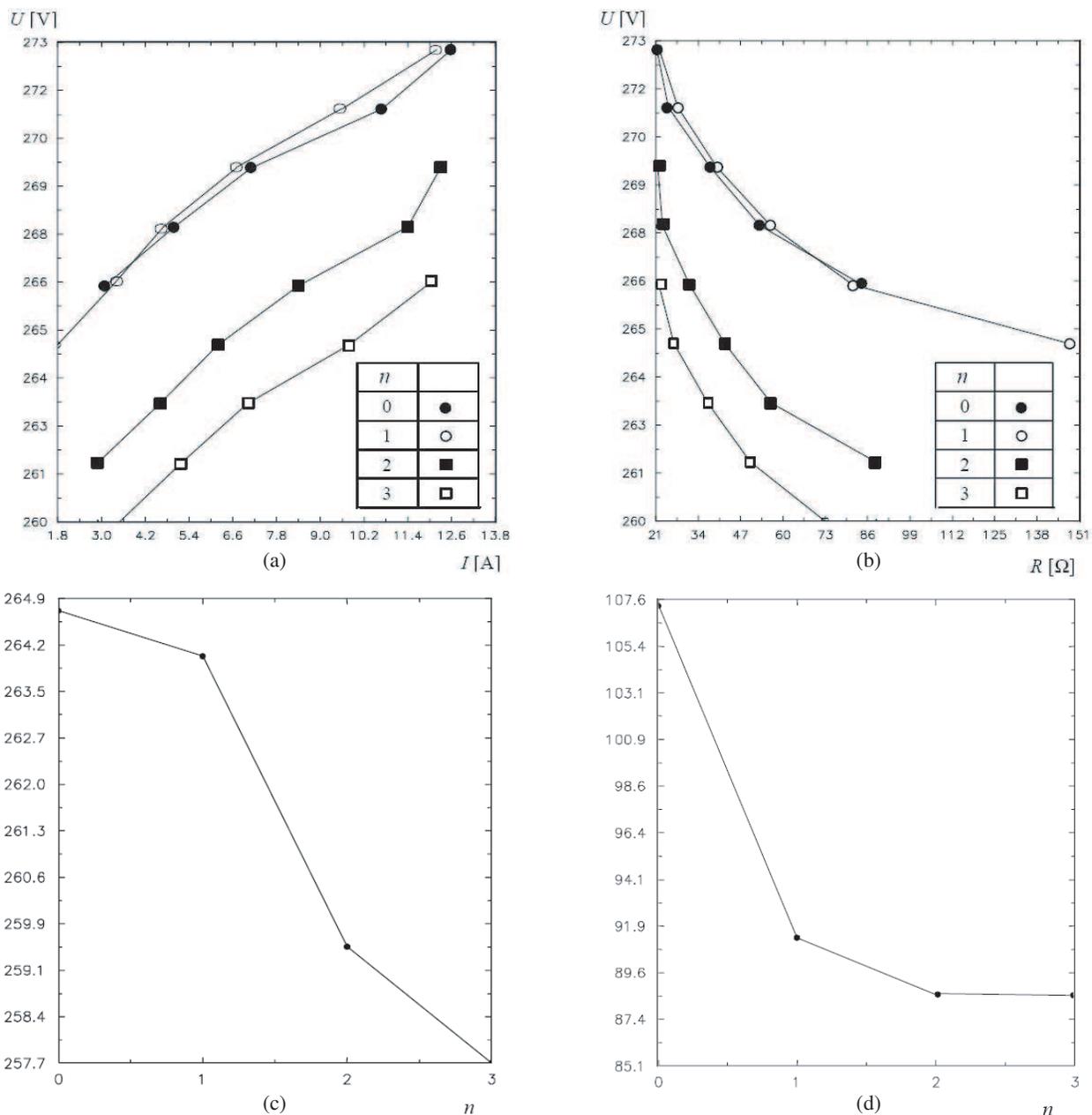


Figure 1: Post-irradiation characteristics of semiconductor over-voltage diodes: (a) Volt-ampere characteristic, (b) volt-ohm characteristic, (c) breakdown voltage, (d) nonlinearity coefficient.

3. RESULTS AND DISCUSSION

3.1. Semiconductor Over-voltage Diodes

Figure 1 shows experimentally observed changes in volt-ampere characteristic, volt-ohm characteristic, breakdown voltage, and the nonlinearity coefficient of semiconductor over-voltage diodes, as the absorbed dose increases. According to these plots, over-voltage diodes exhibit breakdown voltage drop, reflected in the increase of the volt-ampere plot slope, and consequently the decrease of the nonlinearity coefficient.

Each dose level is marked by a number (n), as shown in Table 1. The pre-irradiation state is designated by $n = 0$. Permanent changes noticed during post-irradiation inspection of semiconductor diodes' characteristics can be attributed to the so called displacement damage caused by neutron and gamma radiation. The basic radiation defect of this kind is the Frenkel pair, consisting of a displaced interstitial atom and a vacancy. Energy levels of these defects, as well as of the stable complexes which they form with atoms of impurities and dopants present in the semiconductor, are located within the energy gap. Some of these defects represent very efficient recombination centers. Recombination rate of minority carriers depends on the concentration of recombination centers, that becomes higher in an irradiated semiconductor material. This further decreases minority carrier lifetime, causing an increase in diode leaking current and a decrease of the breakdown voltage. The rise of charge carrier recombination rate produces a drop in their concentration and mobility, which causes an increase of semiconductor material specific resistance. Consequently, the nonlinear coefficient decreases [8, 9].

The influence of the neutron field component on the increase of bulk carrier recombination is much larger than the influence of the gamma component, since at energies characteristic of the Californium-252 source neutrons cause significantly more (approximately hundredfold more) displacements of atoms from the crystal lattice than do gamma rays.

Lattice discontinuities at the boundary surface of diode semiconductor material introduce surface states, with energy levels within the forbidden gap. These states act as recombination centers for charge carriers reaching the surface, in the same way as bulk defects and impurities. Areal density of such states for oxide-passivated diodes used in this paper is $\sim 10^{11} \text{ cm}^{-2}$. When semiconductor material is irradiated with gamma photons, both the surface recombination rate and the density of surface states increase [10, 11]. Increase of the surface recombination rate, as well as transient radiation effects (also called dose-rate effects, such as photocurrents produced by gamma ray ionization in the semiconductor), are manifested only during the irradiation process, and therefore cause no permanent damage to the diodes which could have been observed in the experiments. For this reason, gamma absorbed doses, measured during experiments, have not been stated in the paper [12, 13].

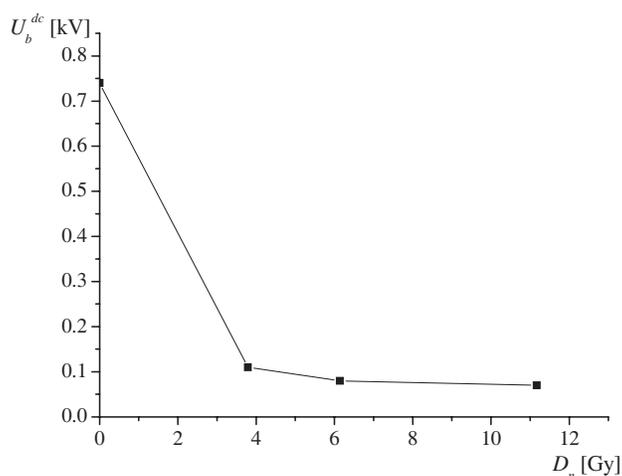


Figure 2: Gas-filled over-voltage diode dc breakdown voltage versus neutron absorbed dose in argon.

3.2. Gas-Filled Over-voltage Diodes

Figure 2 shows dc breakdown voltage U_b^{dc} of the gas-filled over-voltage diode versus the absorbed dose. In Figs. 3(a) and (b) the 99.99% and 0.01% quantiles of the gas-filled diode volt-second char-

acteristic before and after irradiation are presented. Experimental results show that by irradiating the gas-filled over-voltage diode the standard deviation of its *dc* breakdown voltage significantly decreases (quantile curves move closer to each other). The irradiated gas-filled diode exhibited a more rapid response and had a narrower volt-second characteristic (i.e., a smaller value of pulse breakdown voltage U_b^p). This means that its protective characteristics were improved.

Faster response of the gas-filled diode after irradiation is a consequence of a higher concentration of free electrons in the inter-electrode gap created by gas ionization. Ionization is induced through neutron activation of diode construction materials. The shortening of gas-filled diode response time can be attributed to higher availability of potentially initializing free electrons. Fig. 4 presents the gas-filled diode activation analysis diagrams a) immediately after exposure to the radioactive source and b) six hours after irradiation, obtained by a gamma spectrometer. The activity of radioactive isotopes consists of both γ and β components. Due to the induced radioactivity, gas ionization is intensified and the statistical time of a pulse breakdown voltage is reduced. Neutron radiation effects improve the pulse shape characteristic for a short period of time. This effect of neutron radiation disappears quickly, since the half-life time of induced activity varies from several minutes to several hours. This fact is confirmed clearly by the activation analysis plot recorded six hours after irradiation (Fig. 4(b)). It can be noticed from these plots that neutron activation products have decayed almost completely and that the gas-filled diode had recovered back to the pre-irradiation state.

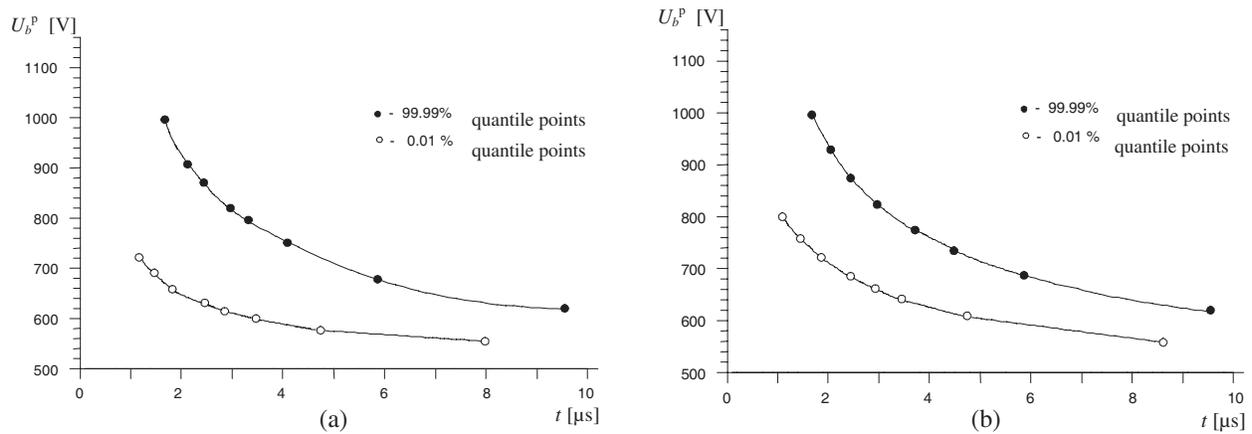


Figure 3: Quantiles (99.99% and 0.01%) of the gas-filled over-voltage diode volt-second characteristic (a) before and (b) after irradiation.

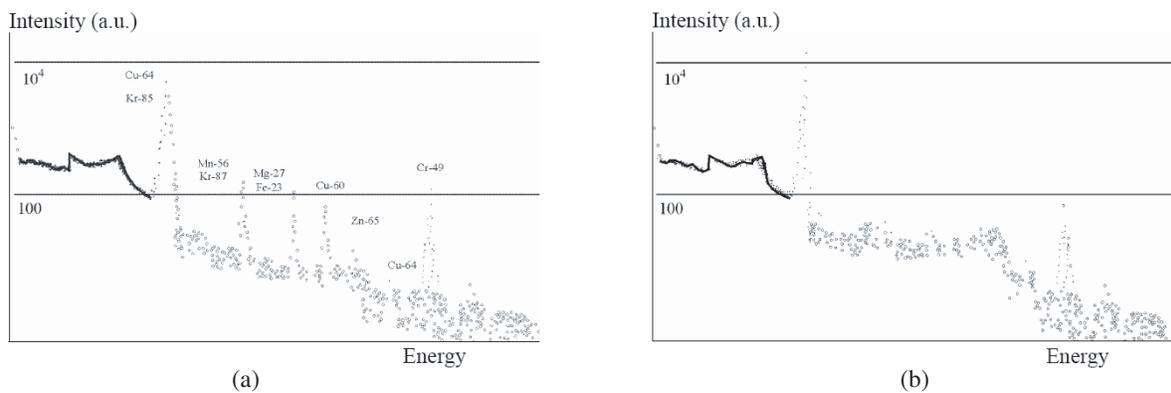


Figure 4: Activation analysis diagrams for the gas-filled over-voltage diode. (a) Immediately after irradiation and (b) six hours after irradiation.

4. CONCLUSION

Experimental results showed that irradiation of semiconductor over-voltage diodes by a combined neutron/gamma field causes a permanent degradation of their protection characteristics. On the

other hand, gas-filled over-voltage diodes exhibit a temporal improvement of performance. As was discussed in the paper, radiation induced changes in semiconductor over-voltage diode operation is attributed to the rise of bulk and surface carrier recombination rates, caused by both neutrons and gamma rays. In the case of gas-filled over-voltage diodes, radiation induced changes are mainly due to the effects of neutron field component, producing a higher concentration of potentially initializing free electrons in the diode inter-electrode gap.

ACKNOWLEDGMENT

The Ministry of Science and Environmental Protection of the Republic of Serbia supported this work under contract 141046.

REFERENCES

1. Standler, R. B., *Protection of Electronic Circuits from Overvoltages*, Dover Publications, 2002.
2. Beyer, M. and W. Boeck, *Hochspannungstechnik, Theoretische und Praktische Grundlagen*, Springer-Verlag, Berlin, 1986.
3. Howard, C., "New avalanche diode for transient protection," *Electronic Product Design*, 1983.
4. Osmokrović P., B. Loncar, and S. Stanković "The new method of determining characteristics of elements for overvoltage protection of low-voltage system," *IEEE Transactions on Instrumentation and Measurement*, Vol. 55, 257–265, 2006.
5. Gill, K., G. Hall, and B. MacEvoy, "Bulk damage effects in irradiated silicon detectors due to clustered divacancies," *J. Appl. Phys.*, Vol. 82, 1, July 1997.
6. Gill, K., G. Hall, S. Roe, S. Sotthibandhu, R. Wheadon, P. Giubellino, and L. Ramello, "Radiation damage by neutrons and photons to silicon detectors," *Nucl. Inst. and Meth.*, Vol. A322, 177–188, 1992.
7. Hauschild, W. and W. Mosch, *Statistical Techniques for High-voltage Engineering*, IEEE Power Series 13, Peter Peregrinus Ltd., 1992.
8. Vavilov, V. S. and N. A. Ukhin, *Radiation Effects in Semiconductors and Semiconductor Devices*, Consultants Bureau, New York, 1977.
9. Holmes-Siedle, A. G. and L. Adams, *Handbook of Radiation Effects*, 2nd Edition, Oxford University Press, 2002.
10. Osmokrović, P., M. Stojanović, B. Loncar, N. Kartalović, and I. Krivokapić, "Radioactive resistance of elements for over-voltage protection of low-voltage systems," *Nucl. Inst. and Meth. in Physics Research B*, Vol. 140, 143–151, 1998.
11. Alexander, D. R., "Transient ionizing radiation effects in devices and circuits," *IEEE Trans. Nucl. Sci.*, Vol. 50, 565–582, 2003.
12. Vasić, A., P. Osmokrović, B. Loncar, and S. Stanković, "Extraction of parameters from I-V data for nonideal photodetectors: A comparative study," *Materials Science Forum*, Vol. 494, 83–88, 2005.
13. Loncar, B., P. Osmokrović, M. Stojanović, and S. Stanković, "Radioactive reliability of programmable memories," *Japanese Journal of Applied Physics*, Vol. 40, Pt. 1, No. 2B, 1126–1129, 2001.

A Shape Display Method Based on Electromagnetic Localization and Actuation

K. Deng¹, E. T. Enikov¹, and P. Marek²

¹Department of Aerospace and Mechanical Engineering, University of Arizona, USA

²Institute of Applied Mechanics and Mechatronics, Slovak University of Technology in Bratislava, Slovakia

Abstract— A 3D shape display method based on electromagnetic localization and actuation is presented and the correspondent miniature device is developed. Evaluation of the single actuator device shows that majority of shape information can be delivered. The device containing multiple actuators is designed to improve the small shape and large curvature perception. The scaling analysis of EM force is performed to determine the maximal allowable actuators per unit area. The prototype device containing 4×4 actuators array is developed for further evaluation.

1. INTRODUCTION

The need to display graphics to visually impaired people has been growing with the emergence of World Wide Web (WWW). Although several types of computer-aided software were devised to transfer selective information into Braille language, the major obstacle of the communication for the visually impaired people is the deficient delivery of the graphic information. As a reason, the interest has arisen in developing shape display devices. Electronic devices aiming to fulfill this function can be categorized into two groups. One is to use active touch method where the profile of the shape is produced by using a tactile matrix [1, 2] or virtual environment through force feedback [3]. These devices, however, either demand a relatively large platform to incorporate a touching interface, or require large driving element to produce enough force to users. The other group of devices is to use passive sensing method where the pattern is locally produced onto human skin by using spatially discriminated stimulation [4], which is mainly successful for two dimensional graphs.

This paper presents an approach that combines both active and passive tactile method to display three dimensional virtual objects by using the miniature vibrator(s). The presented device incorporates an electromagnetic (EM) position detector and EM actuator(s) to selectively stimulate user's finger based on its location. The working principle is shown by the block diagram in Figure 1. Human brain interacts with the EM actuator as receiving stimulation and guides the finger to move following user's estimate of the reference shape. The position comparison results in a decision to engage the actuator in selective locations. The iterations of receiving stimulation and modifying shape estimates produce to users an optimal estimation of the shape.

2. SINGLE ACTUATOR DEVICE

The device containing a single EM actuator is shown in Figure 2. The three position sensor modules, each containing a high sensitivity solenoid and a developed signal processing circuit board,

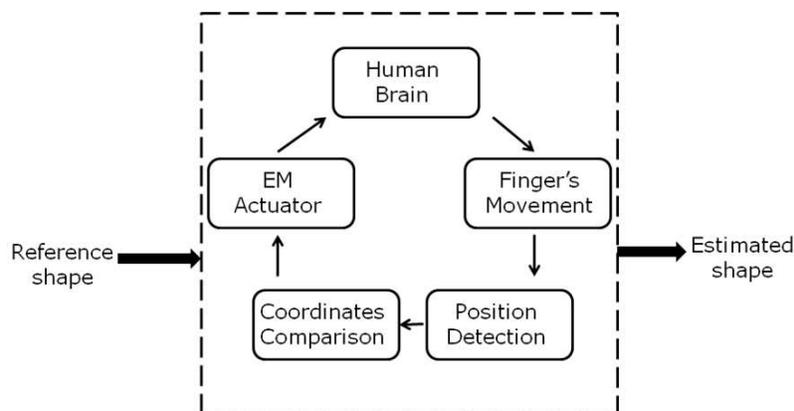


Figure 1: Block diagram of shape display method.

are mounted perpendicularly on the sides of the plastic holder. An EM actuator containing a permanent magnet and a driving coil is inserted into center of the holder. The permanent magnet is attached to a thin plastic film to hold the vibration. The two small cylinders besides the actuator are used to support the fingertip in order to prevent users from pressing the magnet too tightly, thus allowing a less resistance of the vibration. The sensors' outputs are connected to the analog to digital converter (ADC) and are directed to microcontroller for signal processing.

The principle of the EM localization derived by [5] is shown in Equation (1),

$$|B_n| = \frac{k}{R_n^4} \sqrt{3(z - z_n)^2 + R_n^2} \quad (1)$$

where $|B_n|$ is the magnitude of flux density caused by the n th powered beacon coil where $n = 1, 2, 3$; $k = \mu_0 I_0 N A$, where μ_0 is the free space permeability, N , A , I_0 are the beacon coil's number of turns, area, and applied current respectively; z and z_n are the z direction coordinate of the location and center of n th beacon coil, where z axis is defined to be perpendicular to beacon coils' plane; R_n is the distance between location and center of the n th beacon coil. Mathematically, a unique proper solution of three dimensional position vector can be obtained by solving Equation (1) for three beacon coils.

We have presented our preliminary shape display results in [6] and [7]. It has been shown that a simple on and off mode for actuator that works at a constant vibration frequency provided only limited cues to users. In this study, a variable vibratory frequency based on stimulator's location is

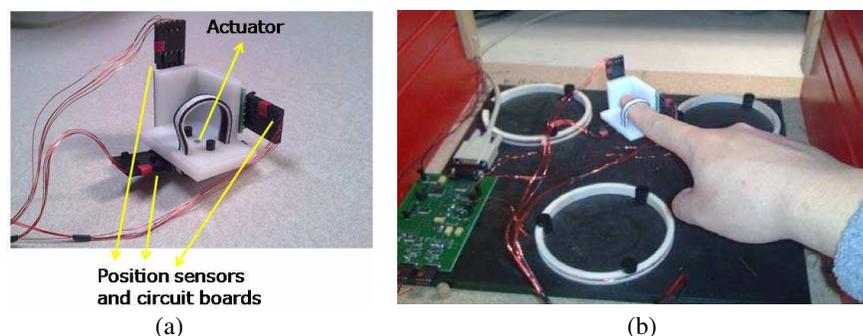


Figure 2: (a) Photograph of the device and (b) shape display spatially defined by 3 beacon coils, and the master circuit board for signal transmission.

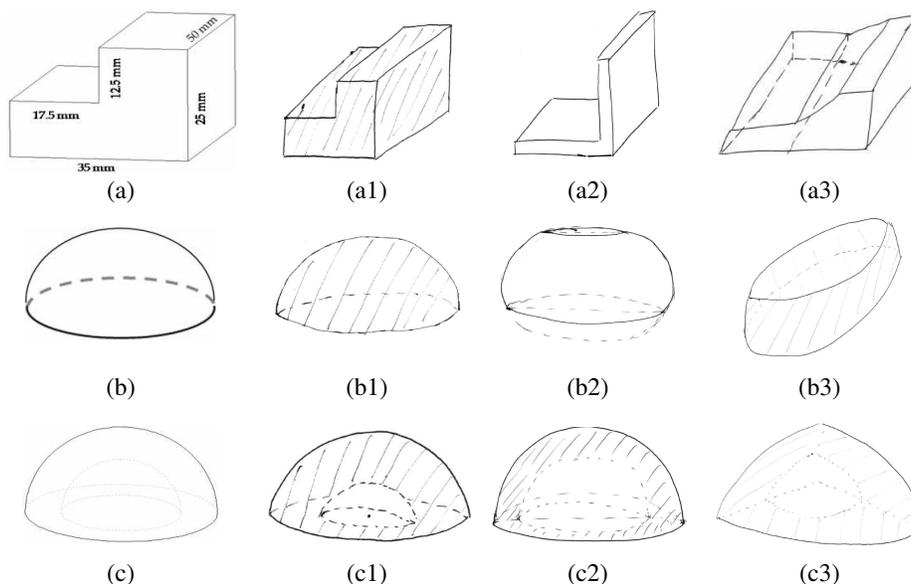


Figure 3: Shape (a) L structure; (b) half sphere with 65 mm radius; (c) half hollow sphere with outer and inner radius of 60 mm and 37 mm respectively; drawings from 3 volunteers based on their perception.

experimented to enhance the detection of the curved surface. The perception test of three shapes participated by three volunteers was performed. The volunteers are sighted, between 20–30 years old. Two of them are male and one is female. The displayed shapes were not disclosed except that the shapes were three dimensional. The perception time was restricted to three minutes for each shape perception. For each perceived shape, the volunteer was told to sketch the shape to one’s best knowledge. However, the brief description was allowed in case the volunteer is not able to sketch the drawing.

The test results are shown in Figure 3. The shape (a) was an L structure, which was perceived well except for a minor incorrect estimation of join plane by user 3 (Figure 3(a3)). For shapes (b) and (c), we applied a fast vibration (120 Hz) at the surfaces. The vibration frequency linearly decreased along inside direction of the object and reached a minimal value of 15 Hz at the innermost location of the shape. The actuator remained off outside of the shapes. As a result, user 1 correctly perceived shape (b) and (c). User 2 had slight incorrect curvature estimation in shape (b) but well perceived shape (c). User 3 perceived the shape (b) as a tilted partial sphere and shape (c) as a major part of the upper half of the hollow sphere.

3. MULTIPLE ACTUATORS DESIGN

In principle, users can get the complete shape information from a single actuator by fully scanning the area provided sufficient time. However, the users’ perception performance usually decreases after the initial one minute’s perception. Small scale and highly curved shapes are therefore more difficult to identify in a short time. One possible improvement is to use multiple such actuators to simultaneously display different part of the shape. It is easier for the brain to interpret the integration of the individual display other than performing the same interpretation after equivalent times of scanning activities.

To allow the maximal number of actuators mounted onto fingertip, the minimal diameter of the actuator needs to be determined. The EM force as a function of the actuator’s scale is then analyzed by using finite element software. The actuator model is illustrated in Figure 4(a) while the simulation result is shown in Figure 4(b). In the analysis, the residual flux density of the permanent magnet is 0.75 Tesla and its coercive force is 5750 Oersteds. The magnet has the same side length as the outer diameter of the driving coil. The inner diameter of the driving coil is half of the outer diameter. The inner core material of the coil has a relative permeability of 200. Height of the coil is fixed at 5 mm. The perception limit of the peak EM force was found to be 51 mN [8]. The thermal limit was experimentally observed at a peak current of 0.9 A. From Figure 4(b), 1.5 mm is observed to be a suitable size for the miniaturized actuators. The distance between two adjacent actuators can be set as 2 mm which is the two point discrimination distance of human finger [4]. The prototype device containing 4×4 actuators array is shown in Figure 5.

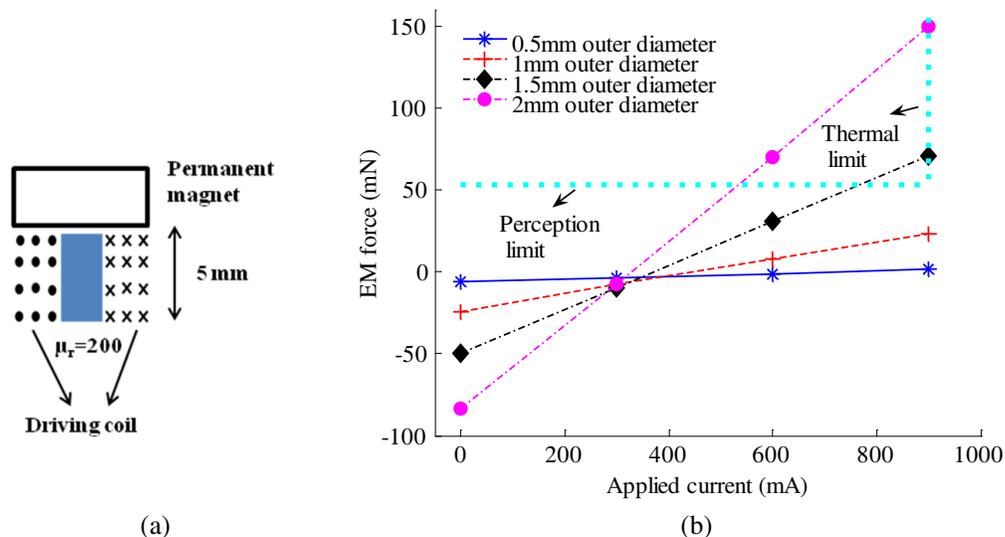


Figure 4: (a) Simulation model and (b) EM force scaling analysis.

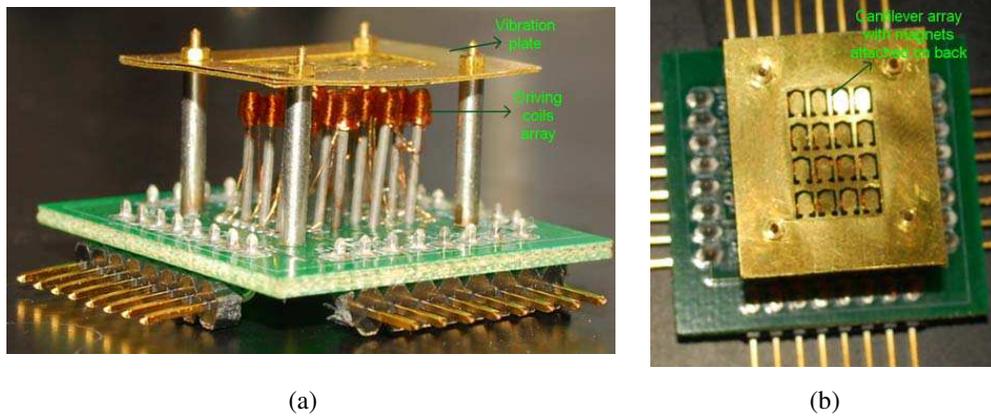


Figure 5: Photograph of (a) 3D view of 4×4 actuators array (b) top view of vibration plate containing cantilever array with permanent magnets attached on back side.

4. CONCLUSIONS

A shape display method based on EM positioning and vibratory actuation is presented. The device containing single EM actuator is developed and tested. Variable vibration frequency is implemented to enhance the surface detection. The test results show that majority of information of the designed shapes can be delivered to users without voice aid within three minutes. The accuracy of curvature perception may be further improved by using multiple actuators to simultaneously display individual part of the shape. The EM force as of the size of the actuator is therefore analyzed to determine the minimal diameter of the actuator. It is shown that a minimum of 1.5 mm outer diameter is suitable to meet both perception and thermal requirement. The device containing 4×4 actuators is developed for future evaluation.

ACKNOWLEDGMENT

This work was supported by the National Science Foundation under Grant No. BES-0603198. The authors would like to also acknowledge the help from Mr. Vasco Polyzojev and Mr. Edward White in experimental aspect of this work.

REFERENCES

1. Iwata, H., H. Yano, and F. Nakaizumi, "Project FEELEX: Adding haptic surface to graphics," *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques*, 469–476, 2001.
2. Kawai, Y. and F. Tomita, "Interactive tactile display system — A support system for the visually disabled to recognized 3D objects," *ASSETS'96*, 45–50, 1996.
3. Iwamoto, T., M. Tatezono, and H. Shinoda, "Non-contact method for producing tactile sensation using airborne ultrasound," *EuroHaptics 2008*, 504–513, 2008.
4. Fukuda, T., H. Morita, F. Arai, et al., "Micro resonator using electromagnetic actuator for tactile display," *1997 International Symposium on Micromechatronics and Human Science*, 143–148, 1997.
5. Prigge, E. A. and J. P. How, "Signal architecture for a distributed magnetic local positioning system," *IEEE Sensors Journal*, Vol. 4, No. 6, 864–873, 2004.
6. Deng, K., E. T. Enikov, and H. Zhang, "Development of a pulsed micro-actuator for 3D tactile display," *2007 IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, 1–5, Zurich, Sep. 2007.
7. Deng, K., E. T. Enikov, and V. Polyzojev, "Development of a shape conveying interface based on tactile feedback," *2009 IEEE/ICME International Conference on Complex Medical Engineering*, 1–5, Tempe, USA, Apr. 2009.
8. Deng, K., "Development of virtual 3D tactile display based on electromagnetic localization," Ph.D. dissertation, University of Arizona, 2009.

Coplanar-fed UWB Elliptical Patch Antenna with Notched Band Characteristics

R. A. Sadeghzadeh¹, M. A. Honarvar², and A. R. Eskandari³

¹Dept. of Eng., Khajenasirtoosi University, Tehran, Iran

²Islamic Azad University, Najafabad Branch, Isfahan, Iran

³Islamic Azad University, East Tehran Branch, Tehran, Iran

Abstract— This paper presents a novel band notched compact size antenna for ultra-wideband (UWB) communication. The antenna consists of a elliptical metal patch and a $50\ \Omega$ coplanar waveguide (CPW) transmission line. By etching two U and inverted U-shaped slots in the metal patch, band-stopped filtering properties are achieved. By attaching two parasitic U-shaped strip to the top layer of antenna stronger notch is obtained. The proposed antenna yields an impedance band width of 3.1–10.5 GHz with VSWR < 2 except the band width of 5–6 GHz for IEEE 802.11a. The stable omnidirectional patterns in H -plane and symmetric in E -plane are also obtained.

1. INTRODUCTION

In recent years, ultra wideband (UWB) system has required for many applications. The feasible design and implementation of UWB system has become a highly competitive topic in both academy and industry communities of telecommunications. Recently, the Federal Communication Commission (FCC)'s allocation of the frequency band 3.1–10.6 GHz for commercial use has sparked attention on ultra-wideband (UWB) antenna technology in the industry and academia [1]. According to the Federal Communications Commission (FCC), UWB system is defined as any radio system that has a 10-dB bandwidth larger than 25 percent of its centre frequency, or has a 10-dB bandwidth equal to or larger than 1.5 GHz if the centre frequency is greater than 6 GHz. For these frequency band applications several antenna configurations have been studied [2–4]. However, in the frequency band of UWB communication systems there might potentially exist several narrowband band interferences caused by other wireless communication systems, such as IEEE 802.11a wireless local area network (WLAN) in the frequency band of 5.15–5.825 GHz. Therefore, it is necessary for UWB antennas performing band-notched characteristic in this frequency band to avoid electromagnetic interference (EMI) between UWB and WLAN systems. Lately, a number of antennas with band-notched property have been discussed [5–8].

In this paper, we propose a simple and compact CPW-fed UWB antenna with band-notched characteristics in 5–6 GHz. To achieve wide band characteristic, tapered ground plane is used. The band-notched operation is achieved by etching two U and inverted U-shaped slots in the metal patch and attaching two parasitic U-shaped strip to the top layer of antenna. In this case the total length of two strips is equal to the total length of two slots. By attaching two parasitic U-shaped strip stronger notch is obtained. Effects of varying the parameters of the U and inverted U-shaped slot and parasitic U-shaped strip on the performance of antenna have also been investigated. The simulation of antenna structure was performed with Ansoft HFSS software.

The remaining of this paper organized as follows. Section 2 presents the configuration of the antenna. Simulation results accompanied with some discussions are presented in Section 3. Finally, Section 4 concludes the paper.

2. ANTENNA CONFIGURATION

The geometry and configuration of the proposed UWB antenna with notched-band characteristic is illustrated in Figure 1. The antenna was fabricated on $h = 1.6$ mm FR4 epoxy substrate with the size of $L \times W = 30 \times 34$ mm and dielectric constant $\epsilon_r = 4.4$ and loss tangent $\tan \delta = 0.02$. As shown in the Figure 1 an elliptical patch is fed by a $50\ \Omega$ coplanar waveguide transmission line with $W_f = 3.8$ mm. Since both the antenna and the feeding are implemented on the same plane, only one layer of substrate with single-sided metallization is used, and the manufacturing of the antenna is very easy and low cost. To achieve wide band characteristic, tapered ground plane is used. For this antenna, the radiation element consists of an elliptical patch with axes of A and B .

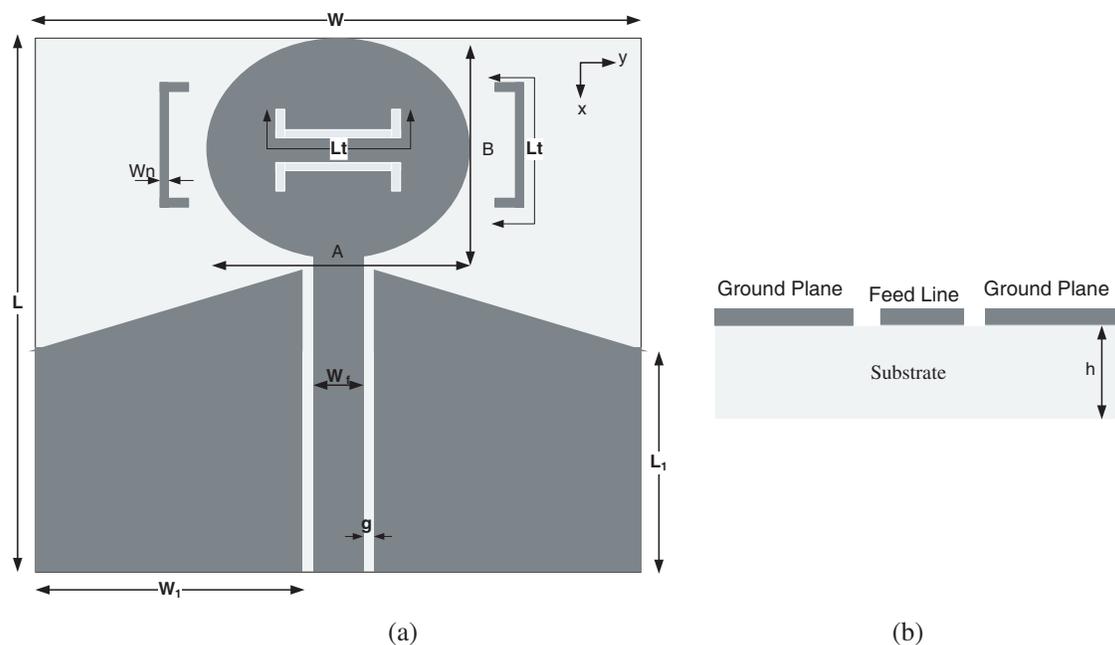


Figure 1: Configuration of proposed antenna. (a) Top view. (b) Side view.

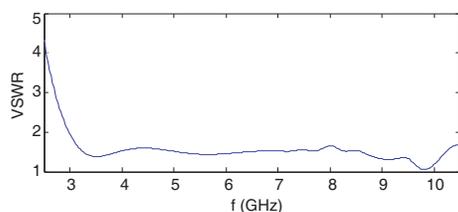


Figure 2: VSWR of antenna without two U-shaped slots in the metal patch and two parasitic strip on the top layer of substrate.

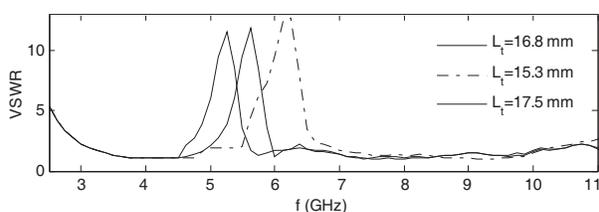


Figure 3: Effects of the total length L_t on the VSWR of proposed antenna.

3. SIMULATION AND RESULTS

To achieve the UWB characteristics of the proposed antenna after optimization process the final parameters are $W = 30$ mm, $W_1 = 12.6$ mm, $W_n = 0.5$ mm, $W_f = 3.8$ mm, $L = 34$ mm, $L_1 = 13$ mm, $A = 9$ mm, $B = 6.5$ mm and $g = 0.5$ mm. Figure 2 shows the characteristic of simulated VSWR of antenna without two U and inverted U-shaped slots in the metal patch and two parasitic U-shaped strip on the top layer of substrate. It is found that the input impedance of the antenna is well matched as the bandwidth covers the entire UWB (3.1–10.6) with $VSWR < 2$. The band notched characteristics of antenna is achieved by attaching two parasitic strip on the top layer of substrate and etching U-shaped slots in the metal patch with the same total length of parasitic strip. In order to achieve the desired band-notched characteristics, the effects of the total length of the U-shaped slot and parasitic strip on the antenna behavior are studied here. The design concept of the notch function is to adjust the total length of the U-shaped slot and parasitic strip to be approximately half wavelength at the desired notched frequency.

The notched frequency f_r can be empirically approximated by:

$$f_r \cong \frac{c}{2L_t \sqrt{\frac{\epsilon_r + 1}{2}}} \quad (1)$$

where L_t is the total length of the U-shaped slots and parasitic strip and c is the speed of light. We can take (1) into account in obtaining the total length of L_t at the very beginning of the design and then adjust the geometry for the final design. Figure 3 shows the simulated VSWR of proposed antenna for different L_t . As shown in Figure 3 it is investigated that by increasing the length of L_t the notched frequency is shifted toward lower frequencies.

The simulated gain of the proposed antenna is illustrated in Figure 4. As shown in Figure 4, gain decreases drastically at the notched frequency band of 5.5 GHz. The simulated radiation patterns

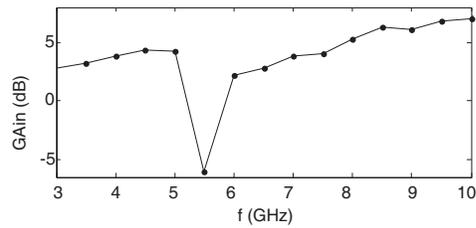


Figure 4: Simulated gain of antenna.

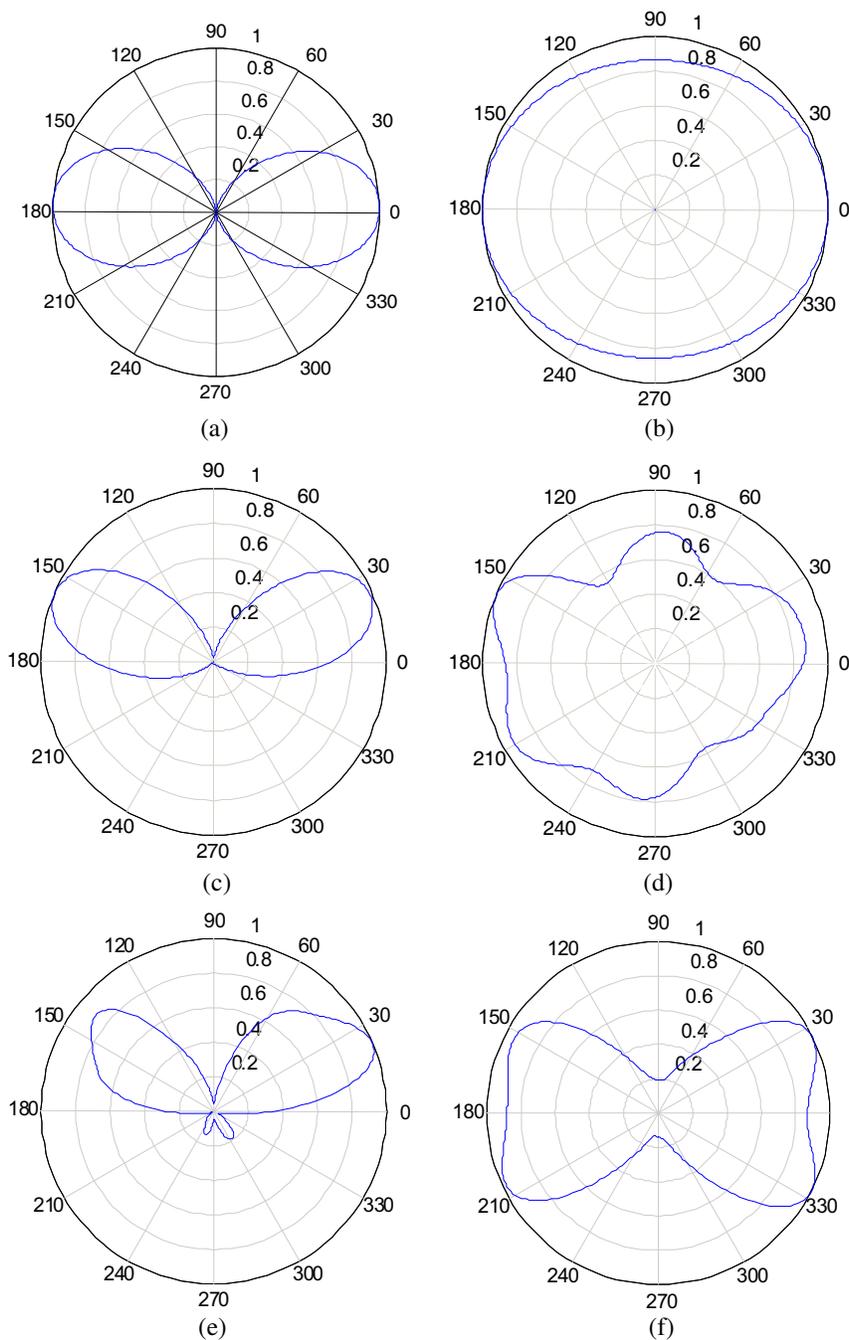


Figure 5: Normalized radiation pattern. (a) *E*-plane at 3 GHz, (b) *H*-plane at 3 GHz, (c) *E*-plane at 7 GHz, (d) *H*-plane at 7 GHz, (e) *E*-plane at 10.5 GHz, (f) *H*-plane at 10.5 GHz.

of antenna in the E -plane (xz -plane) and H -plane (yz -plane) for two different frequencies 3, 7 GHz are shown in Figure 5. The pattern in the H -plane are quite omnidirectional and in the E -plane, the radiation patterns like a small dipole.

4. CONCLUSIONS

A novel band-notched CPW-fed UWB antenna with rejection bands at WLAN frequencies has been proposed. The band-notched operation is achieved by etching two U and inverted U-shaped slots in the metal patch and attaching two parasitic U-shaped strip to the top layer of antenna. By attaching two parasitic U-shaped strip stronger notch is obtained. The relation between the total length of U-shaped slots and parasitic strip and the band-rejected operation has been discussed. Stable radiation pattern are obtained.

REFERENCES

1. FCC report and order for part 15 acceptance of ultra wideband (UWB) systems from 3.1–10 GHz, Washington, DC, 2002.
2. Angelopoulos, S., A. Z. Anastopoulos, D. I. Kaklamani, A. A. Alexandridis, F. Lazarakis, and K. Dangakis, "Circular and elliptical CPW-fed slot and microstrip-fed antennas for ultra wideband applications," *IEEE Antennas and Wireless Propagation Letters*, Vol. 5, 294–297, 2006.
3. Ray, K. P. and Y. Rang, "Ultra wideband printed elliptical monopole antennas," *IEEE Transactions on Antennas and Propagation*, Vol. 55, No. 4, 1189–1192, April 2007.
4. Zhou, H. J., Q. Z. Liu, J. F. Li, and J. L. Guo, "A swallow-tailed wideband planar monopole antenna with semi-elliptical base," *Journal of Electromagnetic Waves and Applications*, Vol. 21, No. 9, 1257–1264, 2007.
5. Zhou, H. J., Q. Z. Liu, Y. Z. Yin, and W. B. Wei, "Study of the band-notched function for swallow-tailed planar monopole antennas," *Progress In Electromagnetics Research*, PIER 77, 55–65, 2007.
6. Akhoondzadeh-Asl, L., M. Fardis, A. Abolghasemi, and G. Dadashzadeh, "Frequency and time domain characteristic of a novel notched frequency UWB antenna," *Progress In Electromagnetics Research*, PIER 80, 237–348, 2008.
7. Kim, K.-H. and S.-O. Park, "Analysis of the small band-rejected antenna with the parasitic strip for UWB," *IEEE Transactions on Antennas and Propagation*, Vol. 54, No. 6, June 2006.
8. Zhang, G.-M, J.-S. Hong, and B.-Z. Wang, "Two novel bandnotched UWB slot antennas fed by microstrip line," *Progress In Electromagnetics Research*, PIER 78, 209–218, 2008.

Near Field Antenna Investigation and Evaluation for UHF RFID Systems

Zijian Xing, Ling Wang, Changying Wu, and Dengshan Huang
School of Electronics and Information, Northwestern Polytechnical University
Xi'an 710072, China

Abstract— The reader antenna is an important part and has great influence on the performance of the whole radio-frequency identification (RFID) systems. This paper will investigate and evaluate two near field antennas for UHF RFID readers with the emphasis on the near field performance. Modeling, parameters optimization, and performance of these two near field antenna will be illustrated in detail. Furthermore, performance comparison of these two antennas to show the advantages and disadvantages will be given.

1. INTRODUCTION

The radio-frequency identification (RFID) technology is of growing interest to commerce, industry, and academia. The reader antenna is an important part and has great influence on the performance of the whole RFID systems. Currently, near field UHF RFID receives a lot of attention as a possible solution for item level tagging in pharmaceutical and retailing industry [1].

Two kinds of PCB near field reader antennas for RFID systems will be investigated and evaluated in this paper. The presented antennas are suitable for RFID application since it is smaller and lighter than some other kinds of antennas, easily produced, and low price. This paper is structured as follows: Section 2 describes the modeling and structure of the presented antennas. Performance analysis of the antennas, e.g., return loss, far field gain, current and Magnetic field distribution, are presented in Section 3. Finally, the main conclusions are summarized in Section 4.

2. MODELING AND STRUCTURE OF ANTENNAS

Both of the presented antennas are made of PCB board. This kind of antenna is smaller and lighter than some other kinds of antennas, and easily produced. The two antennas are both loop structure. With the loop structure, the current of loop circuit antenna can be designed to keep consistent, thus the magnetic field distribution around the loop's axis will be concentrated and enhanced, which is easier for tag to receive more energy from reader's antenna. Typical closed-path magnetic induction strength formula in (1) can verify this magnetic field distribution. Current direction is loop's tangent bearing and if it is consistent, the magnetic field direction is also consistent and superimposed. Therefore, the magnetic field strength can be enhanced by the current of every segment on the loop antenna.

$$\mathbf{B} = \frac{\mu_0}{4\pi r} \oint_l \frac{I d\mathbf{l}' \times \mathbf{a}_R}{R^2} \quad (1)$$

Antenna A is shown in Fig. 1, where the radius of PCB board is 50 mm, relative permittivity, and the dielectric loss tangent of the substrate is 3.4 and 0.02, respectively. The width of the loop circuit on PCB board is 2 mm and the internal radius is 36 mm. There is great influence of the loop size on antenna performance. Generally, smaller size will result in higher resonance frequency [2]. The marker "1" on the antenna is the feed point. "2" to "7" are six capacitors of 1 pf. "8" is resistance of 50 ohm. There is a cavity around the PCB board and internal radius of the cavity is 60 mm, the internal height is 2 cm and the thickness of PCB board is 2 mm.

Antenna B shown in Fig. 2 is similar to Antenna A, where the width of the loop circuit on PCB board is 2 mm and the internal radius is 38 mm. "1" is feed point. "2" and "3" are two slots with width of 1 mm. The function of the slot is same as the capacitors in Antenna A. "4" is 50 ohm resistance.

The purpose of the capacitors on the circle is to reduce the electrical length of circle and adjust the matching network. Loop's terminal load is a resistance. It will broaden bandwidth and reduce far field gain. The purpose of using cavity is to concentrate the electromagnetism energy in the region of interest.

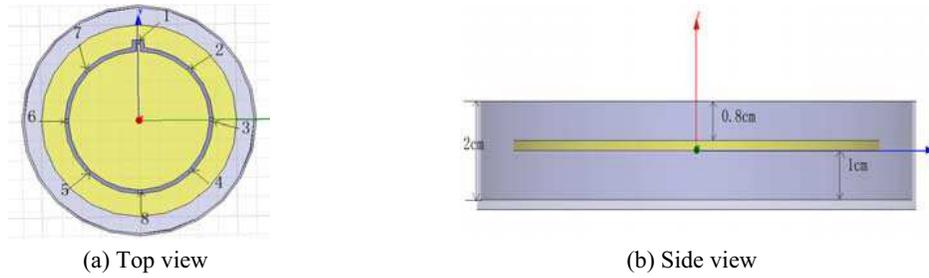


Figure 1: Antenna A.

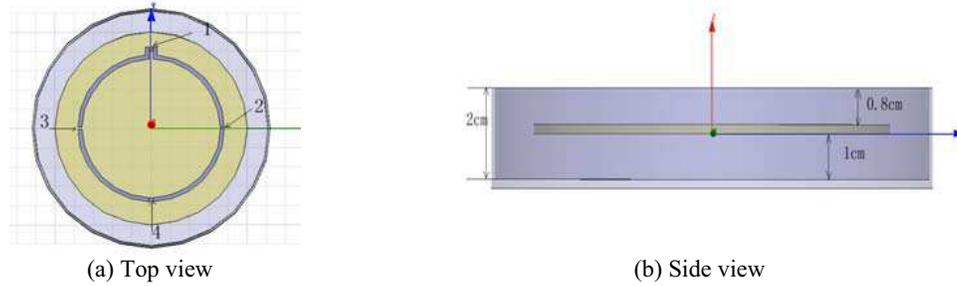
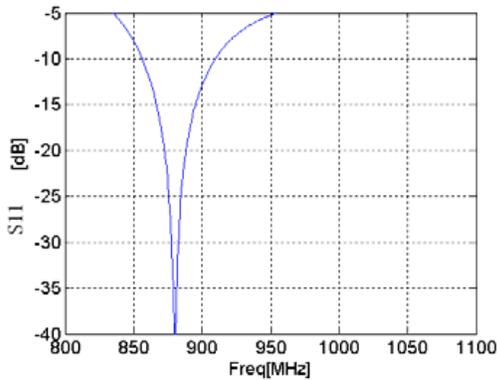
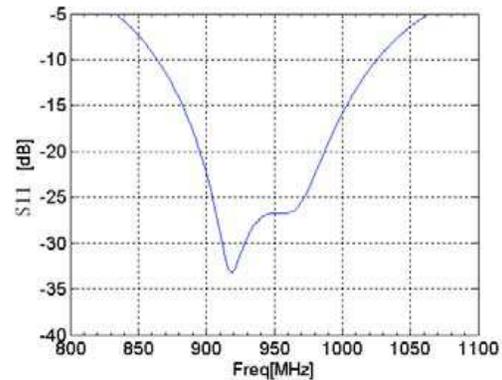


Figure 2: Antenna B.

Figure 3: S_{11} of Antenna A.Figure 4: S_{11} of Antenna B.

3. PERFORMANCE ANALYSIS

3.1. Return Loss

Performance of antenna reflects the electric circuit characteristic and radiation characteristic. Return loss is the important electric circuit characteristic. The S_{11} parameter of the presented antennas is shown in Fig. 3 and Fig. 4.

From Fig. 3 and Fig. 4, it is shown that the two antennas work at UHF band and the bandwidth is very wide. The bandwidth of Antenna B is 116 MHz at -15 dB. Since both of the antennas have matched resistance at the terminal, the current distribution on the loop is traveling wave, and hence the antennas show the nice impedance matching characteristic at the working band [3]. Resonance network is an important aspect for antenna design. The resonant network of Antenna A consists of 6 capacitors and that of Antenna B consists of slots. The width and space of the slot has to be designed for capacitance characteristic according to (2).

$$C = \frac{\varepsilon S}{4\pi k d} \quad (2)$$

Metal loop may regard as the resistance and inductance. Therefore the entire antenna consists of many series of connected LC resonators. The series resonator resonance point can be calculated by

$$\omega = \omega_0 \frac{1}{\sqrt{LC}} \quad (3)$$

By adjusting capacitance of C or changing L by adjusting the size of the loop circuit, the resonant point of the antenna would be changed correspondingly.

3.2. Far Field Gain

The designed antennas work in the near field areas and far field gain should be as low as possible. If the excess the far field gain is strong, there will be more radiation energy to adjacent tag and cause mis-registration [4]. The far field gain of Antenna A and Antenna B are given in Fig. 5 and Fig. 6, respectively, where different color stands for the gain when Phi is 0, 90, 180, 270 degree. The gain of Antenna A and Antenna B under different direction is less than -17 dB and -12 dB. So low gain is induced by antenna's circle structure and matching resistance, which consumes the majority of input power and result in weak current, and hence reduced Q value.

3.3. Current Distribution

The current is a very important aspect of near field antenna, because the current's intensity and direction is inseparable with near magnetic field distribution. If the direction of current on the

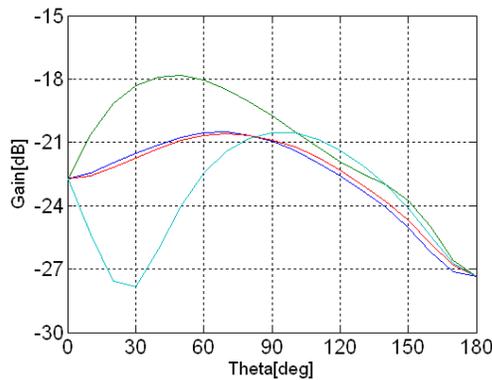


Figure 5: Far field gain of antenna A.

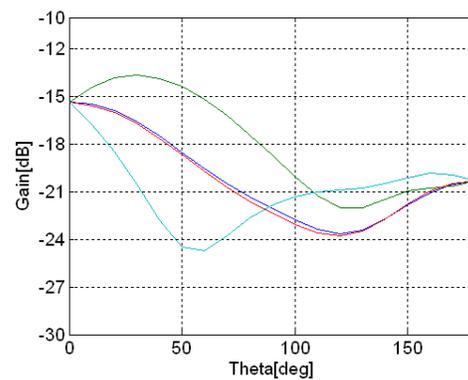


Figure 6: Far field gain of antenna B.

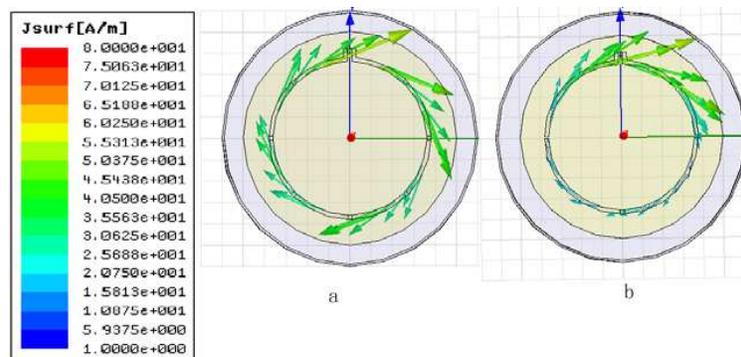


Figure 7: Current distribution of antenna A and antenna B.

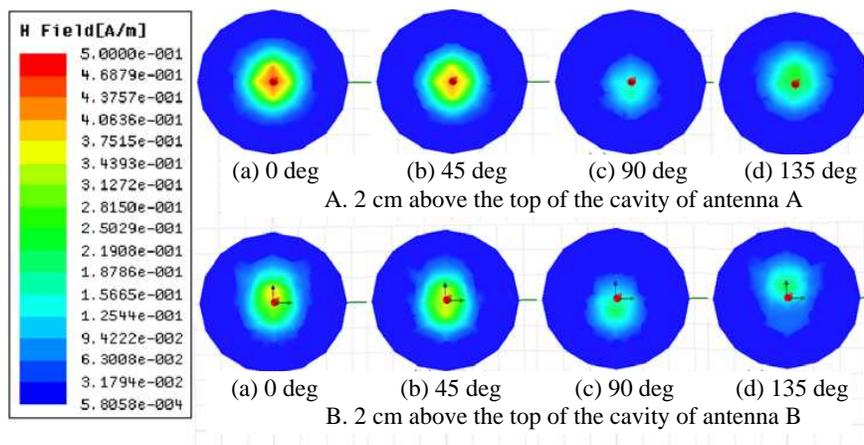


Figure 8: Magnetic distribution.

loop is inconsistent, the direction of magnetic field is also inconsistent. It will counter-balance very large part of power, causes the weaker magnetic field strength. In the presented antennas, half-ring electrical length has to be less than $\lambda/4$ to avoid current reversal.

The current distribution of Antenna A and Antenna B is shown in Fig. 7. Perimeter of these two antennas is 22.6 cm and 23.8 cm, respectively, and more than half of the wave length of about 16.5 cm. Therefore, it is not conform to the current consistent principle. By adding capacitors in the circle can reduce the electrical length of the loop without changing the physical length [5]. In Antenna A, there are 6 capacitors of 1pf to change the electrical length. In Antenna B, two slots have been used to change the electrical length. By this way, the obvious nodal point and anti-nodal point can be seen on different position of the antenna loop. The objective is to keep the current consistent.

3.4. Magnetic Field Distribution

Near field strength is the most important performance of the NF antenna. Most of near field tags are ring-like. The magnetic field radiated by the antenna induces the electromotive force on the loop of tag and the induced strength can be calculated by

$$E_{in} = - \int_S \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{S} \quad (4)$$

Electromotive force is not only related with field intensity, but also related with loop's size and direction of tag. In order to read tag better, magnetic field's direction and intensity should be simultaneously considered in designing reader antenna.

Figure 8 shows the magnetic distribution of Antenna A and Antenna B for different phase. In Fig. 8(a), magnetic field strength changed greatly with phase, but the peak point of magnetic field is always nearby central axis since Antenna A's current phase of the loop is nearly the same. In Fig. 8(b), peak point of magnetic field is constantly changes along with phase since antenna B has obvious traveling wave characteristic and magnetic field's peak point generally nearby current anti-nodal point.

4. CONCLUSIONS

Two near field antenna for RFID systems and the performance have been presented. Bandwidth of antenna B is greatly wider than antenna A. Wide bandwidth brings two advantages. Firstly, antenna's anti-jamming ability can be stronger. Secondly, wide bandwidth antenna can work in many frequency bands. But the wide bandwidth also brings shortcoming, for instance, there will be more noises can pass through the wide bandwidth. Antenna A and Antenna B has very low far field gain and is suitable for NF RFID application. Merely, with the reduction of far field gain, the near magnetic is inevitably reduced. Therefore a good trade-off has to be balanced. As for magnetic field aspect, antenna A is better than antenna B since antenna A's peak point of magnetic field is always located nearby the axis and it is not easily to appear "empty hole" while reading tags. On the other hand, antenna B is more stable than antenna A since antenna A has 6 capacitances, but on antenna B only has two slots. Capacitance's instability and error will cause unexpected problem. It would be better to use less lumped elements in designing antenna.

ACKNOWLEDGMENT

This work is supported by the National Natural Science Foundation of China (No. 60802083) and by NPU Foundation for Fundamental Research (NPU-FFR-JC200817).

REFERENCES

1. Nikitin, P. V., K. V. S. Rao, and S. Lazar, "An overview of near field UHF RFID," *2007 IEEE International Conf. on RFID*, 166–174, Mar. 26–28, 2007.
2. Pal, A., S. Behera, and K. J. Vinoy, "Design of multi-frequency microstrip antennas using multiple rings," *IET Microwaves, Antennas & Propagation*, Apr. 9, 2008.
3. Ling, C., *Antenna Project Handbook*, 306–308, Publishing House of Electronics Industry, June 2002.
4. Liu, Z.-M. and R. R. Hillegass, "A 3 patch near field antenna for conveyor bottom read in RFID sortation application," Accu-Sort System Inc., Telford, PA, 2006.
5. Ryu, H.-K. and J.-M. Woo, "Size reduction in UHF band RFID tag antenna based on circular loop antenna," Department of Radio Science & Engineering, Chungnam National University, Daejeon, Korea, 2005.

The Design and Simulation of an S-band Circularly Polarized Microstrip Antenna Array

Ying Jiang, Hongchun Yang, and Xiong Wang

School of Physical Electronics

University of Electronic Science and Technology of China

Chengdu 610054, China

Abstract— It is important to use microstrip antenna in aircraft for stealth and movement. Based on the principle of rectangle microstrip antenna, the S-band circularly polarized microstrip antenna array is designed according to the quota requirement. By using a software for 3D electromagnetic field analysis (Ansoft HFSS), the S-band microstrip antenna array is simulated and the optimization of the parameters of design is obtained. The material object of antenna is made and measured in microwave anechoic chamber. The design of the S-band microstrip antenna array meets the requirement well comparing with HFSS's simulation result and has practical engineering value.

1. INTRODUCTION

Since the concept of microstrip antenna was proposed, it has been widely used in satellite communications, navigation and other fields because of its thin profile, light weight, conformal with the carrier, easy integration with active devices, etc. In particular the circular polarization microstrip antenna [1] has some significant advantages: incident electromagnetic wave with any polarization can be received by circular polarized antenna; the circular polarized wave can also be received by antenna with any kind of polarization; circular polarized antenna is orthogonal and if the incident wave is circular polarized wave, the handed direction of reflected wave will be inversed and so on. Because of these characteristics, circular polarization antenna has strong anti-interference ability. It has been widely used in electronic reconnaissance and jamming, communications and radar polarization diversity as well as in areas such as electronic counterwork.

In this paper, based on the design of microstrip antenna feed networks and related issues, a circular polarization microstrip antenna array was designed in S-band, simulated by High Frequency Simulation Software (HFSS), fabricated and measured. Its design specifications are as follows:

Table 1: Design requirements.

Frequency range (S-band)	Antenna Gain (at 2491.75 MHz)	Polarization	Horizontal 3 dB beam width
2491 ± 5 MHz	≥ 15 dB	right circular	≥ 30°
Vertical 3 dB beam width	Polarization axial ratio	VSWR	size of panel
≥ 12°	≤ 6 dB	≤ 1.5 : 1	440 mm × 140 mm × 5 mm

2. DESIGN OF UNIT ANTENNA

Typically, microstrip antenna design is the overall objective of the work in the designated frequency range to perform a particular characteristic. To achieve this overall objective, the primary task is to choose a suitable geometry of the patch for the microstrip antenna, such as rectangular, circular patch, trigon and pentagon. If there are no special requirements, rectangular patch is preferred not only because that the rectangular microstrip antenna is simple to design and easy to manufacture. Besides, there is a series of more mature theory as a basis for analysis of rectangular patch antenna, such as the transmission line method, cavity model method [2] and so on. Rectangular microstrip patch antenna structure as shown in Fig. 1.

Suppose patch length is L , patch width is W , dielectric thickness is h , dielectric length is L_s , dielectric width is W_s , relative dielectric constant is ϵ_r , effective dielectric constant is ϵ_e , light speed is c , center frequency is f_r , wavelength in medium is λ_g , free-space wavelength is λ_0 , extension of volume is ΔL .

Patch width is

$$W = \frac{c}{2f_r} \left(\frac{\epsilon_r + 1}{2} \right)^{-\frac{1}{2}} \quad (1)$$

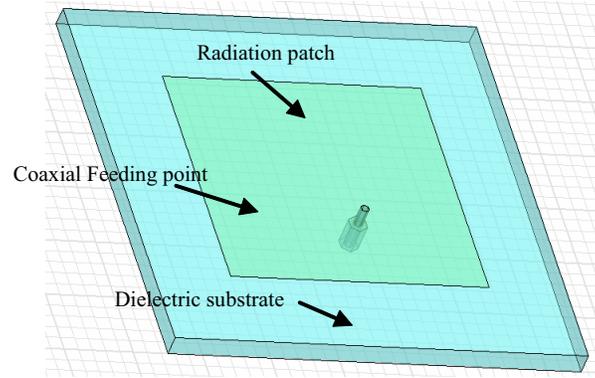


Figure 1: Rectangular microstrip patch antenna structure.

f_r, r, ε_r are known parameter.
Patch length is

$$L = \frac{c}{2f_r\sqrt{\varepsilon_e}} - 2\Delta L \quad (2)$$

$$\varepsilon_e = \frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r - 1}{2} \left(1 + \frac{12h}{w}\right)^{-\frac{1}{2}} \quad (3)$$

$$\frac{\Delta L}{h} = 0.412 \frac{(\varepsilon_e + 0.3)(w/h + 0.264)}{(\varepsilon_e - 0.258)(w/h + 0.8)} \quad (4)$$

Dielectric plate width

$$W_s = W + 0.2\lambda_g \quad (5)$$

$$\lambda_g = \lambda_0 / \sqrt{\varepsilon_e} \quad (6)$$

Higher gain can be attained by choosing lower dielectric constant material such as polyethylene, whose relative dielectric constant $\varepsilon_r = 2.25$, thickness $h = 3$ mm at $f_r = 2.4917$ GHz. Unknown parameters can be calculated from the Equations (1) ~ (6) and known parameters: radioactive element length $L = 38.3$ mm, the width of the radiation element $W = 47.2$ mm, $\varepsilon_e = 2.1$, $\Delta L = 1.6$ mm. Medium wavelength $\lambda_g = 80.3$ mm, the width of medium plate $W_s = 60$ mm, $L_s = 60$ mm.

Microstrip antenna is easy to achieve circular polarization with a single microstrip patch. Currently, the following several ways have been used to achieve circular polarization microstrip antenna radiation: 1. Orthogonal single-patch feed circular polarization microstrip antenna; 2. Single-point feed circular polarization microstrip antenna; 3. Curvilinear wideband microstrip circular polarization microstrip antenna; 4. microstrip antenna array consisting of circular polarization microstrip antenna.

This paper chose single-point feed method [4], whose most significant advantage is additional phase-shifting network and power splitter are unnecessary to achieve the circular polarization. Also, mutual coupling [5] interference brought by the feed network can be significantly reduced if the distance between the antenna unit in array is small enough. Based on cavity mode theory, the trajectory of feeding point is diagonal of rectangular patch. To facilitate determining the location of feeding point, the radiation element can be approximated a square patch, whose size is $W \times L = 38$ mm \times 38 mm, and medium-size is 60 mm \times 60 mm \times 3 mm. The size of the air box is 200 mm \times 200 mm \times 80 mm, about $\lambda/4$ at the lowest frequency. This paper setup the coaxial-fed model using Ansoft HFSS10.0. Its thickness, width of dielectric, feeding point and ratio of width to length are $h = 3$ mm, $w = 37.6$ mm and $t = 0.96$ respectively. It can be clearly seen the antenna performed not so well at 2491.75 MHz from simulation Smith chart (Fig. 2), indicating the antenna generated mode separation at 2491.75 MHz, and thus circular polarization wave radiation presented.

The antenna standing-wave ratio, directivity and axis ratio are shown in Fig. 3, Fig. 4, and Fig. 5.

As can be seen from Fig. 3, the standing wave ratio at the resonant frequency is less than 1.5, which meets the project targets and performs well on impedance match.

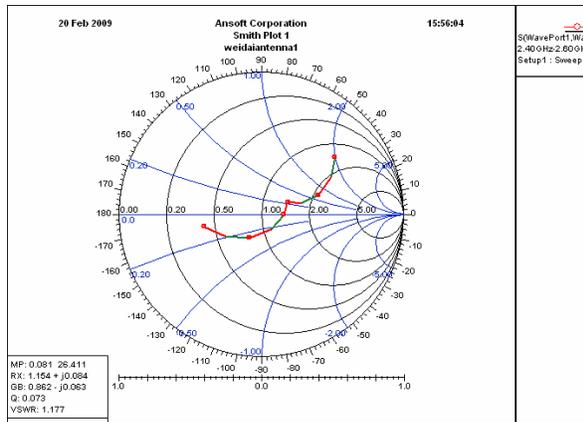


Figure 2: Simulation antenna Smith Chart.

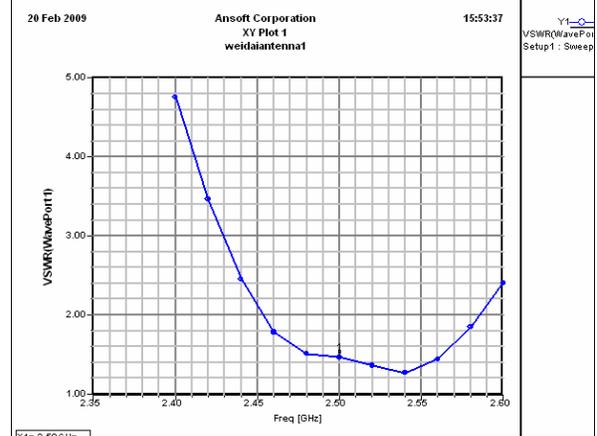


Figure 3: VSWR of the antenna.

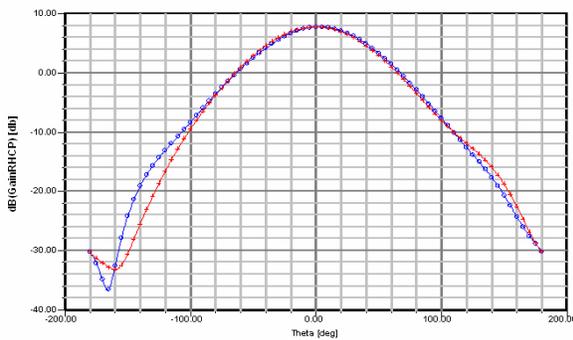


Figure 4: Antenna radiation pattern.

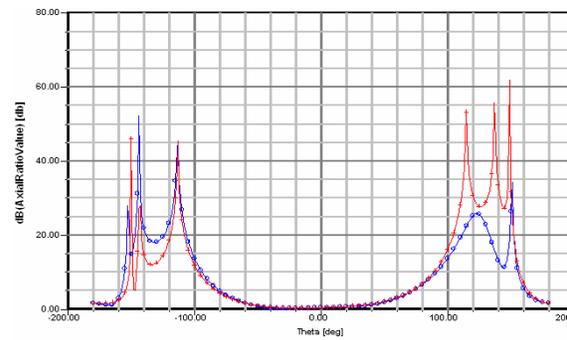


Figure 5: Antenna axial ratio.

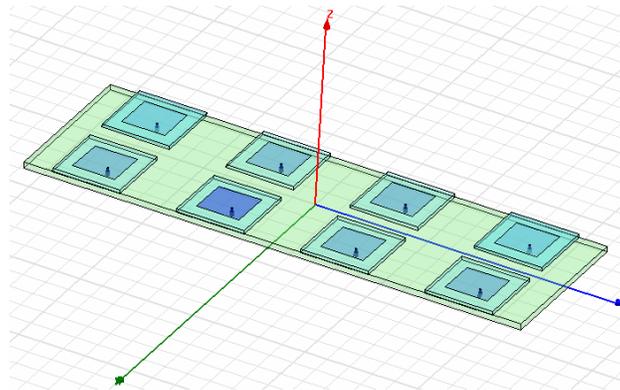


Figure 6: 2 × 4 circular polarization antenna array simulation model.

As can be seen from Fig. 4, the maximum radiation direction is 0° achieving a maximum gain of 8 dB at the operating frequency. The main lobe width is greater than 90° with good directivity. Fig. 5 shows axial ratio characteristic and it can be seen the antenna axial ratio is less than 6 dB in $-70^\circ \sim 70^\circ$ cone angle of the space which meets the requirements.

3. CIRCULAR POLARIZATION MICROSTRIP ARRAY DESIGN [6]

Usually a single microstrip patch antenna array attains 7 dB gain and it will increase by about 3 dB if the number of units doubles. Here, this paper designed an array of 2×4 to meet the requirements of project, including the horizontal 3 dB beam width larger than 30° and vertical 3 dB beam width larger than 12° . 8 patch antenna units are placed on the plate $440 \text{ mm} \times 140 \text{ mm} \times 5 \text{ mm}$. The array inter-element space is $0.5\lambda \sim 0.75\lambda$ in general, which means 115 mm in x axis and 45 mm in y axis. The optimized horizontal and vertical space are 110 mm and 40 mm respectively. Simulation model is shown in Fig. 6.

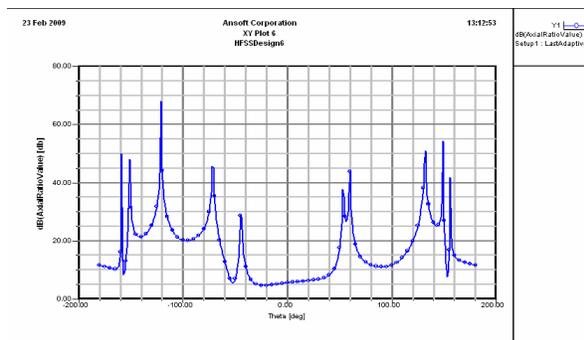


Figure 7: Polarized axial ratio.

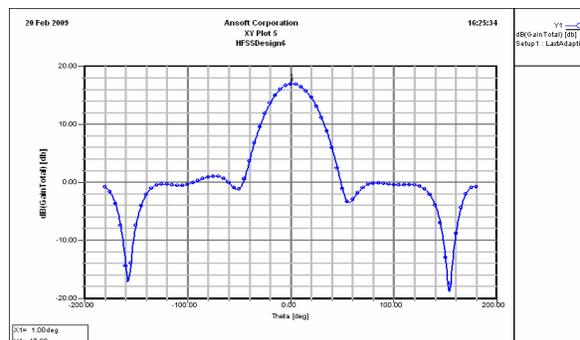


Figure 8: Radiation pattern of antenna array.

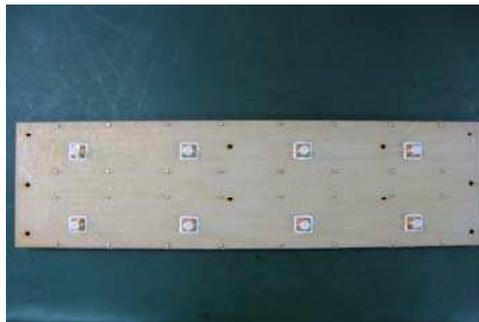
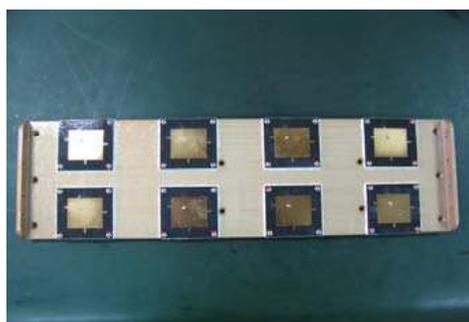
Figure 9: Top and bottom view of 2×4 circular polarization antenna array.

Figure 10: Measurements of array.

From this model, the axis ratio and the directivity of antenna array are shown in Fig. 7, Fig. 8.

Figure 7 shows the axial ratio of antenna array is better than the requirement by 6 dB. Fig. 8 shows radiation pattern of antenna array. In the direction of maximum radiation, the gain is close to 17 dB with good directivity which fully meets the design requirements.

4. FABRICATION AND MEASUREMENTS OF ARRAY

Based on the simulated results of the antenna array, it was fabricated and tested. Fig. 9 shows the top and bottom view of the array.

The antenna array is placed inside a cylindrical radome used as a receiving antenna, as shown in Fig. 10. The circular polarization antenna array was placed in the microwave chamber, measured radiation pattern with E8363B vector network analyzer [7].

After measurement, the maximum gain of antenna array is about 16 dB, as shown in Fig. 11. And the main lobe as well as the side lobe conform well with the simulation results of software HFSS in Fig. 8. The error of gain is about 1 dB, caused by a variety of factors, such as the accuracy of simulation software, mutual coupling between modules, the accuracy of fabrication, measurement instruments error and so on.

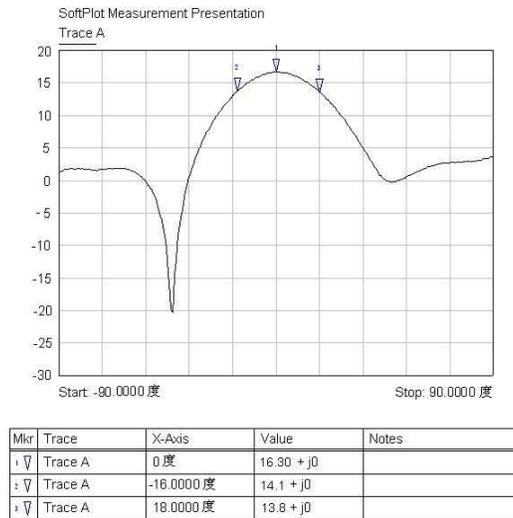


Figure 11: Measured *E*-plane pattern.

5. CONCLUSION

In this paper, based on empirical formula and Ansoft HFSS10.0, a circular polarization microstrip antenna array in the S-band is designed, simulated, fabricated and measured. The test results consist well with simulation results of Ansoft HFSS and meet the design requirements. The antenna array attains a simple compact structure and high gain of 16 dB at 2491 MHz.

REFERENCES

1. Abdin, A. M., "12 GHz planar array antenna for satellite communication," *PIERS Online*, Vol. 4, No. 4, 413–416, 2008.
2. Wong, K.-L., *Compact and Broadband Microstrip Antennas*, John Wiley & Sons, Inc., 2002.
3. Targonski, S. D. and D. M. Pozar, "Design of wideband circularly polarized aperture coupled microstrip antennas," *IEEE Trans. Antennas Propagat.*, Vol. 41, 214–220, Feb. 1993.
4. Nasimuddin, K. P. Esselle, and A. K. Verma, "Wideband high-gain circularly polarized stacked microstrip antennas with an optimized C-type feed anda short horn," *IEEE Transactions on Antennas and Propagation*, Vol. 56, No. 2, Feb. 2008
5. Fang, D. G., C. Z. Luan, and Y. P. Xi, "Mutual coupling in microstrip antenna array: Evaluation, reduction, correction or compensation," *Progress In Electromagnetics Research Symposium*, 37–40, Mar. 7–9, 2005.
6. Tao, D., "Design of circularly polarized microstrip patch antenna array," *Progress In Electromagnetics Research Symposium (PIERS)*, 636–638, Aug. 2007.
7. Nystrom, I. and D. Karlsson, "Reduction of back radiation and cross-coupling in dual polarized aperture coupled patch antennas," *IEEE Antennas Propagat. Soc. Int. Symp. Dig.*, 2222–2225, 1997.

A Design of Reconfigurable Patch Array Antenna with Dual Circular Polarizations

Chung-Hsun Weng, Hsien-Wen Liu, Sheng-Yu Lin, and Chang-Fa Yang

Department of Electrical Engineering, National Taiwan University of Science and Technology
Taipei, Taiwan

Abstract— In this paper, we present a 2×2 microstrip array antenna with dual circular polarizations. To switch the polarization states, the radiating patch is fed by a branch-line coupler and two T-junction power dividers. The array antenna can be reconfigured to radiate electromagnetic waves with either the right-handed or left-handed circular polarizations by controlling the branch-line coupler. A prototype is fabricated with a 0.4 mm thick FR4 top layer for radiating elements and a 0.8 mm thick FR4 bottom layer for a feeding network with a 5 mm air gap between the two layers. Simulated and measured results are compared, which show good agreements. Also, the properties of the reconfigurable array antenna with dual circular polarizations and high radiation efficiency are demonstrated.

1. INTRODUCTION

Microstrip array antennas have been widely and rapidly developed in numerous wireless communication systems. Those array antennas may achieve some advantages such as low-profile easy fabrication and high radiation performance. Compared with ordinary linear polarization (LP) arrays, the circular polarization (CP) arrays are capable of solving multipath and fading issues owing to their polarization diversity. However, lower performance will be expected when those CP arrays radiate with narrow operating bandwidth and half-power beamwidth. Therefore many designs for CP arrays have been reported [1–6] to achieve better performance. By utilizing a ring structure with two paths of 90° difference, the single-feed microstrip CP arrays [1, 2] were achieved to reduce array dimensions. Some studies have also been investigated to widen the operating bandwidth [3–5] and half-power beamwidth [6] suitable for CP applications.

In this paper, we propose a single-feed and reconfigurable 2×2 microstrip array antenna for dual circular polarizations. A feeding network consisted of a 90° branch-line coupler and two T-junction power dividers was employed to make good impedance matching and suitable axial-ratio across the operating band. Thus this broadband design covers not only the WLAN band of 2400 ~ 2483 MHz, but also the licensed WiMAX band of 2500 ~ 2690 MHz. By properly exciting the 90° branch-line coupler, the patch array is able to generate with either the right-handed (RH) or left handed (LP) circular polarizations. Hence the polarization states of the proposed array can be switched flexibly. Moreover, the array antenna is fabricated on an FR4 substrate so that a low-cost design can be implemented and achieved. Details of antenna design and electrical features were presented with the simulated and measured results.

2. ARRAY ANTENNA DESIGN

Figure 1 illustrates the whole geometry of the proposed reconfigurable microstrip array antenna. This array was fabricated with a 0.4 mm thick FR4 top layer for 2×2 radiating elements and a 0.8 mm thick FR4 bottom layer for a feeding network, where a 5 mm air gap was inserted between the two layers. Also, the FR4 substrate has a dielectric constant of 4.4 and a loss tangent of 0.02. For the radiator structure, d_1 and d_2 were 106 mm and 72 mm, respectively corresponding to the horizontal length along the y -axis and vertical length along the x -axis among the 2×2 radiating elements. Also, each element of the patch array has a radius equal to $d_0 = 30$ mm. To achieve better radiation performance in the far-field region, the upper two elements are fed out of phase, as shown in Fig. 1(a). Thus the linear radiation directions for those array elements are determined with the x - and y -axes for the lower and upper elements, respectively. Furthermore, for satisfying the requirement of the CP operation, a compact branch-line coupler is designed, which provides a 90° differential feed to the T-junction power dividers. Details of the feeding network are plotted in Fig. 1(b). Some design parameters indicated in Fig. 1(c) for the branch-line coupler are optimized for $Z_c = 47 \Omega$, $Z_{c1} = 105 \Omega$, $Z_{c2} = 66.8 \Omega$, and $\varphi = 90^\circ$. In this design structure, the RHCP and LHCP may be generated by exciting the port 1 and port 2, respectively. An EM solver, HFSSTM has been used to analyze the electrical characteristics and radiation performance of the patch array.

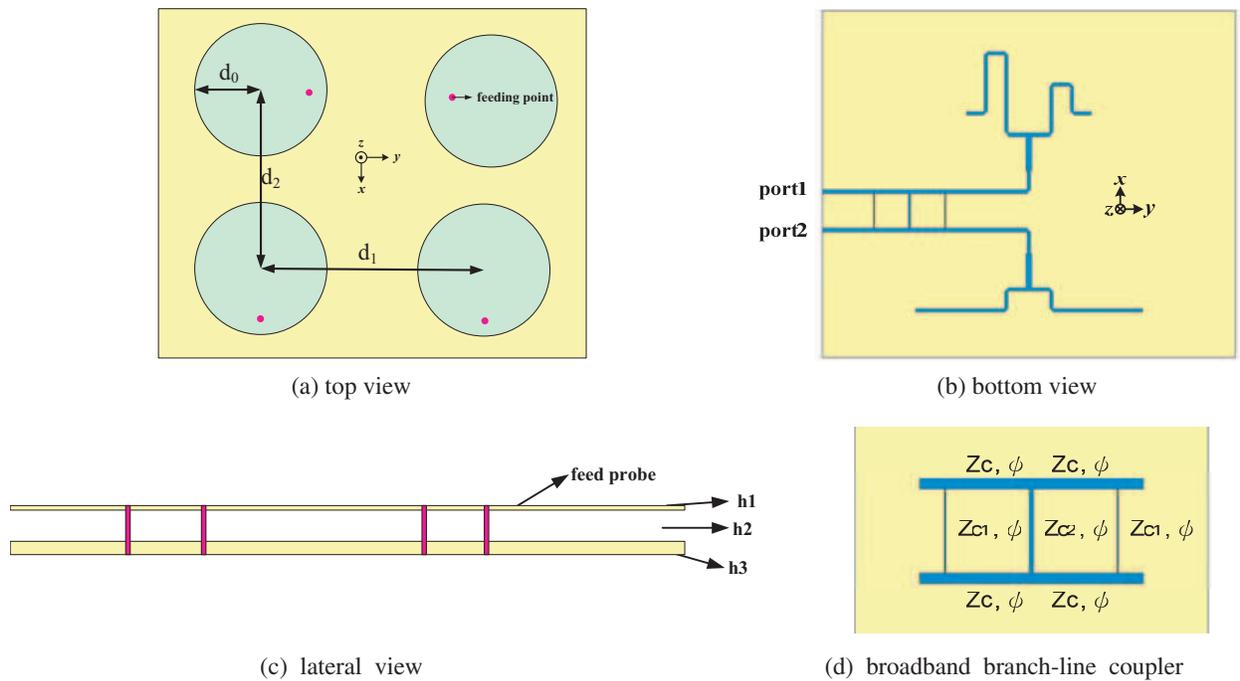


Figure 1: Geometry of the proposed 2×2 microstrip array antenna (designed parameters are $d = 30$ mm, $d_1 = 106$ mm, $d_2 = 72$ mm, $h_1 = 0.4$ mm, $h_2 = 5$ mm, $h_3 = 0.8$ mm).

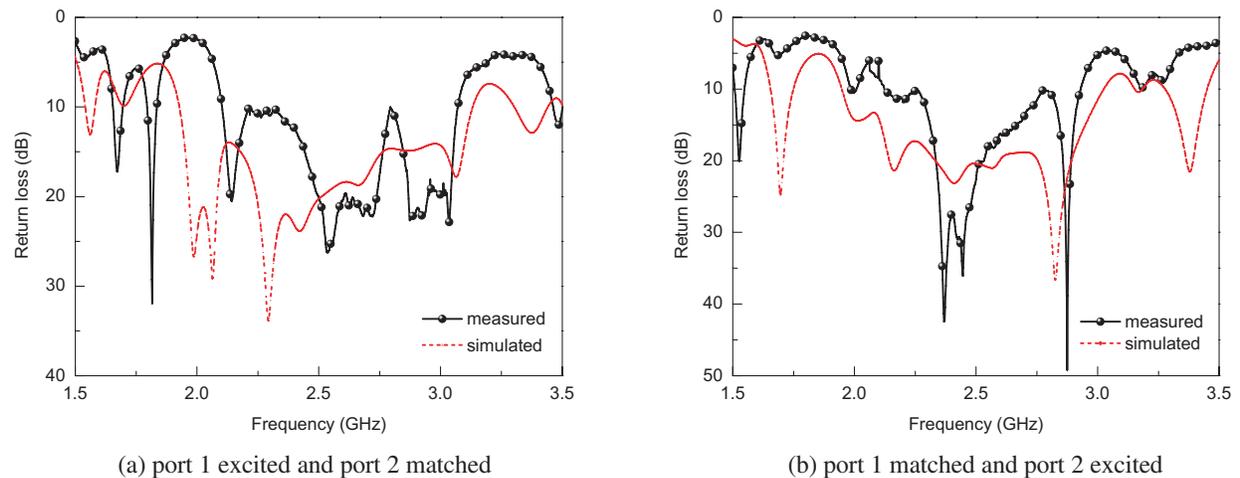


Figure 2: Simulated and measured return losses of the proposed microstrip array antenna.

3. RESULTS

Figure 2 shows the simulated and measured results of the return loss for the proposed reconfigurable patch array antenna. Small differences between the simulations and measurements may be mainly due to the cable connecting the antenna to the network analyzer and also fabrication inaccuracy. According to Figs. 2(a) and (b), the measured impedance bandwidths with 10 dB return loss for the port 1 and port 2 excitations are 2.1 ~ 3.1 GHz and 2.1 ~ 2.9 GHz, corresponding to 38.5% and 32%, respectively. As can be seen, the proposed patch array includes not only the WLAN band of 2400 ~ 2483 MHz, but also the authorized WiMAX band of 2500 ~ 2690 MHz. Moreover, with the broadband feeding network some undesired effects such as frequency shifting and mutual coupling can be reduced too Figure 3 plots the measured radiation patterns in the xz -plane at 2.45 GHz Good RHCP and LHCP patterns have been generated by exciting port 1 and port 2, where the 3 dB beamwidth are 44° and 47° for the RHCP and LHCP states, respectively. Bi-directional radiation patterns shown in the cross polarization are due to the non-orthogonal feed.

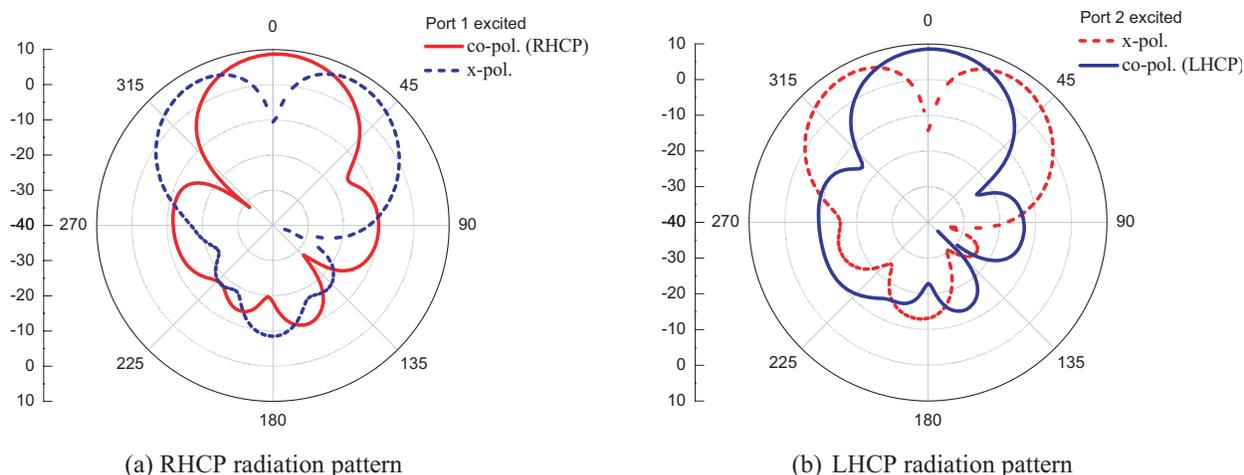


Figure 3: Measured radiation patterns at 2.45 GHz of the microstrip array.

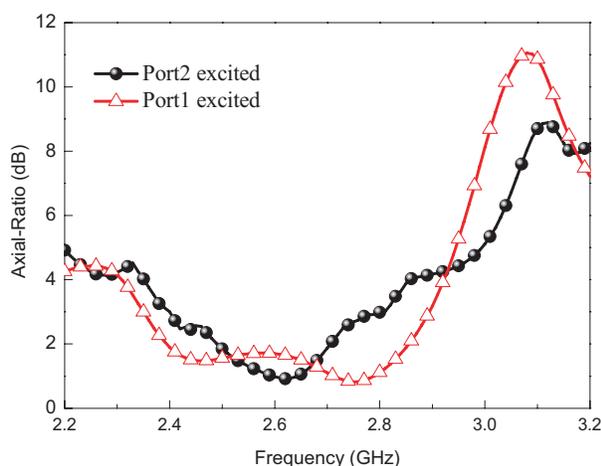


Figure 4: Axial-ratio of the microstrip array.

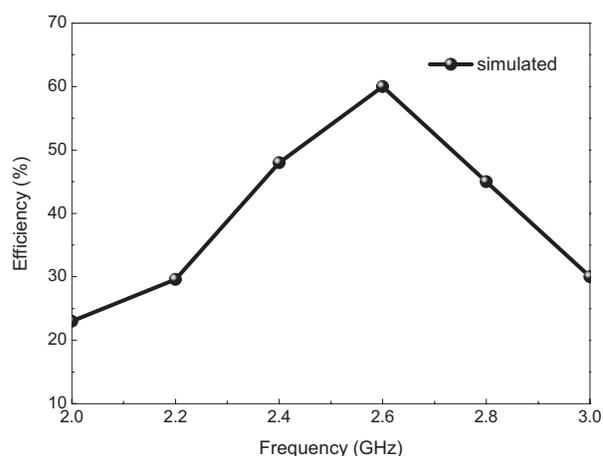


Figure 5: Simulated efficiency of the microstrip array.

According to the measured patterns, the axial-ratio of the array antenna can be computed, as shown in Fig. 4. For the WLAN and WiMAX operations, the axial-ratios in the frequency ranges of 2400~2483 MHz and 2500~2690 MHz are smaller than 3 dB for both the RHCP and LHCP states. Figure 5 simulates the radiation efficiency versus frequency for the proposed array antenna. A significant efficiency drop is mainly attributed to the dielectric loss of the feeding network. To achieve better performance, a low-loss substrate such as Rogers can be used.

4. CONCLUSIONS

A 2×2 microstrip array antenna with dual circular polarizations has been proposed in this paper. This array antenna can be reconfigured to radiate electromagnetic waves with either the right-handed or left-handed circular polarizations by controlling the branch-line coupler. Compared with conventional CP arrays, the proposed array achieves the broadband operation and polarization diversity. Good radiation patterns with an axial ratio less than 3 dB have been achieved. Also, a large coverage can be obtained with the wide beamwidth for both the RHCP and LHCP states. Thus, the proposed microstrip array antenna with dual circular polarizations is very suitable for WLAN and WiMAX applications.

ACKNOWLEDGMENT

This work was supported in parts by National Science Council of R.O.C. under Grant NSC 97-2221-E-011-024.

REFERENCES

1. Chen, H.-M. and K.-L. Wong, "On the circular polarization operation of annular-ring microstrip antennas," *IEEE Trans. Antennas Propag.*, Vol. 47, No. 8, 1289–1292, Nov. 1999.
2. Ramirez, R. R., F. D. Flaviis, and N. G. Alexopoulos, "Single-feed circularly polarized microstrip ring antenna and arrays," *IEEE Trans. Antennas Propag.*, Vol. 48, No. 7, 1040–1047, Jul. 2000.
3. Huang, C. Y., J. Y. Wu, and K. L. Wong, "Slot-coupled microstrip antenna for broadband circular polarisation," *Electron. Lett.*, Vol. 34, No. 9, 835–836, 1998.
4. Hong, Y.-P., et al., "S-band dual-path dual-polarized antenna system for satellite digital audio radio service (SDARS) application," *IEEE Trans. Microwave Theory Tech.*, Vol. 54, No. 4, 1569–1575, Apr. 2006.
5. Janapsatya, J. and K. P. Esselle, "Jawidening the bandwidth of single-fed circularly polarized microstrip patch antenna using sequential array," *Proc. IEEE AP-S Int. Symp. Dig.*, 1–4, Jul. 2008.
6. Guo, Y. X., K. M. Luk, and K. F. Lee, "Broadband dual-polarization patch element for cellular-phone base stations," *IEEE Trans. Antennas Propag.*, Vol. 50, No. 3, 346–351, 2002.

Moment-method Analysis of Planar Archimedean Spiral Antenna with Dielectric Superstrate

Yajian Wu, Huiling Zhao, Dan Jiang, and Nakun Jing
Northwestern Polytechnic University, Xi'an, Shanxi 710129, China

Abstract— A planar Archimedean spiral antenna with a dielectric superstrate and substrate is analyzed using curved segment moment method with curved piecewise triangle sub-domain basis and pulse testing functions. The numerical results of input impedance, current distribution and axial ratio of the Archimedean spiral antenna are presented. Numerical results have good agreement with the published results in [3] and commercial software HFSS. Input impedances as a function of arm length is presented. Affection of permittivity and thickness of antenna substrate and superstrate on the input impedance or bandwidth are discussed.

1. INTRODUCTION

In the past two decades, spiral antennas have received increasing interest because of their wide bandwidths, circular polarization and low profile etc. Different spiral antenna shape and structure have been analyzed. The two arm Archimedean spiral antenna is analyzed theoretical in the free space [1]. And the plane Archimedean spiral antenna printed on grounded substrates is analyzed with moment method, which uses piecewise curved segmentation along the spiral arm that reduces the used segments and the computation time [2]. Then the parameters of a printed spiral antenna with a dielectric superstrate are optimized by optimization using marginal distributions (OMD) algorithm to produce the best axial ratio [3].

In this paper, the curved segments moment method is used to analyze the Archimedean spiral antenna with a dielectric superstrate. The complex image method [4] is used to formulate the Green's functions to solve the Sommerfeld type potential functions with the generalized pencil of functions (GPOF) technique [5]. Then the effecting of the superstrate and substrate to the input impedance and bandwidth of the spiral antenna is discussed.

2. THEORY

The geometry of a two arm microstrip Archimedean spiral antenna with superstrate is shown in Fig. 1 for its top view and side view. Spiral is generated with function $\rho = a\phi$, where a is the spiral constant, ϕ is the winding angle. Typically, a wire with radius of one quarter of strip width is approximately equivalent to the strip. When the radius of the wire is very small compared to the free space wavelength, the thin wire can be used to approximate the spiral wire. Since for a thin wire only, the radiation and impedance characteristics are mainly determined by the axial current component, so the circumferential current can be ignored without any affection.

Based on Pocklington's integral equation, the tangential electric-field of the spiral arm can be written as

$$E_l = \frac{1}{j\omega\epsilon_2} \left[\int I_l k^2 G \cdot \hat{l} \cdot \hat{l}' dl' + \int \frac{\partial I_l}{\partial l'} \frac{\partial (G + \Pi)}{\partial l} dl' \right] \quad (1)$$

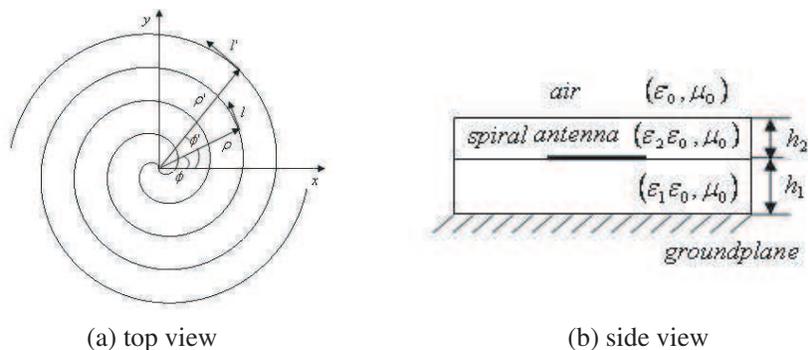


Figure 1: Geometry of two arm microstrip Archimedean spiral antenna with superstrate.

where G and Π are the closed-form Green's functions [4,6], and $i = x$ or y . In order to get the current distribution on the spiral, each arm is discretized into N segments. Expansion the current with curved piecewise triangle basis function, and testing with pulse testing function, Equation (1) can be solved by moment method.

When the current distribution along the spiral arm is achieved by MoM, the input impedance, radiation pattern and axial ratio of the spiral antenna can be obtained too. The far-zone radiation field is express by [3]

$$E_{\theta} = -j30k(\cos(\theta)G_x - \sin(\theta)G_z) \frac{e^{-jkR}}{R} \int_{l_{n-1}}^{l_{n+1}} (I_V \hat{l}_{x'} \cos \phi + I_V \hat{l}_{y'} \sin \phi) e^{jk \cdot \rho} dl' \quad (2)$$

$$E_{\phi} = -j30kG_x \frac{e^{-jkR}}{R} \int_{l_{n-1}}^{l_{n+1}} (-I_V \hat{l}_{x'} \sin \phi + I_V \hat{l}_{y'} \cos \phi) e^{jk \cdot \rho} dl' \quad (3)$$

3. RESULTS

The parameters of the Archimedean spiral antenna are given as follows: the spiral constant $a = 0.97$ mm/rad, the ending winding angle $\Phi_L = 14.68$ rad, and each arm length $L = 106.4$ mm. The parameters of the substrate and superstrate are given as: $\epsilon_1 = 2.2$, $h_1 = 8$ mm, $\epsilon_2 = 4.2$, $h_2 = 6$ mm. The input impedance of the spiral antenna that calculates by the curved segment MoM is presented in Fig. 2, which agrees well with the published result [3] and the result that obtained by HFSS (the feed wire length is set to 2 mm). The deviation of the input impedance between the three results is probably caused by the different feed wire length. The axial ratio of the spiral antenna is shown in Fig. 3. It can be seen from Fig. 2 and Fig. 3 that the spiral antenna has good circular polarization characters over a wide bandwidth.

Figure 4 shows the current distribution along one arm of the spiral at 6 GHz. The input impedance as a function of arm length is shown in Fig. 5. The input impedance behaves well with the different arm length of the spiral at 6 GHz. It suggests that the spiral antenna will have good performance with these configurations.

Figure 6 shows the S_{11} of the spiral antenna with different thickness of the superstrate with the input impedance at 180Ω . It is apparently that the superstrate improves the performance of the spiral antenna considerable. Fig. 7 shows the input impedance as a function of permittivity of the superstrate at 6 GHz, while the other parameters of the spiral antenna are kept the same. Compare to the result presented in Figs. 8 and 9, the substrate has more influence on the input

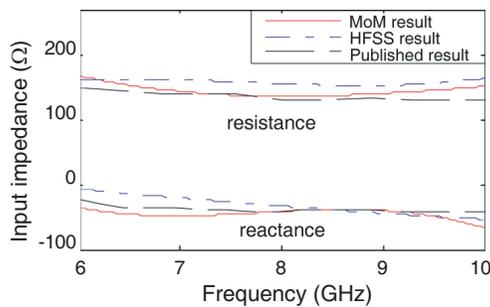


Figure 2: The input impedance of the spiral antenna.

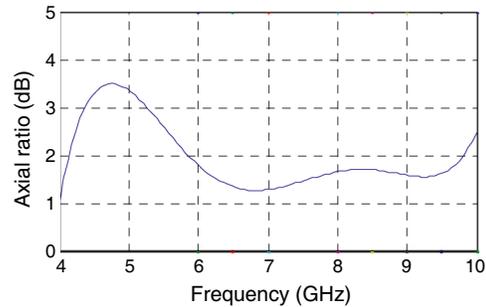


Figure 3: The axial ratio of the spiral antenna.

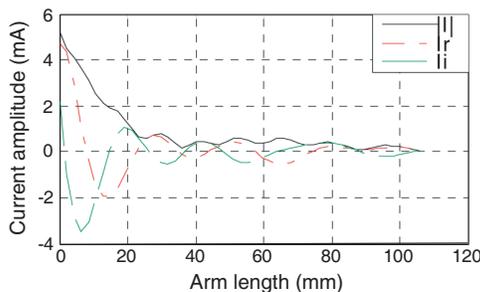


Figure 4: The current distribution along the spiral arm.

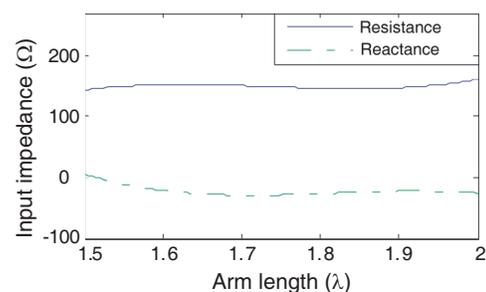


Figure 5: The input impedance as a function of arm length at 6 GHz.

impedance than the superstrate. Although the input impedance is not so depended on the thickness of the substrate and superstrate, it should be considered with the permittivity of the substrate and superstrate together to make the antenna have best performance over a wide bandwidth and be mechanical intension.

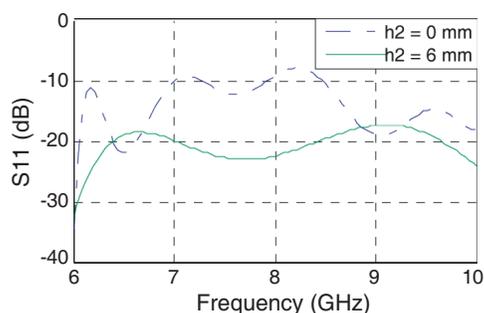


Figure 6: The S_{11} with different thickness of the superstrate.

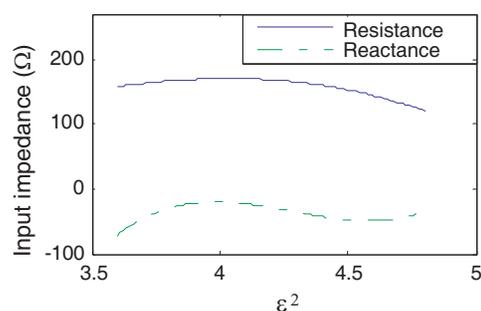


Figure 7: The input impedance as a function of ϵ_2 at 6 GHz.

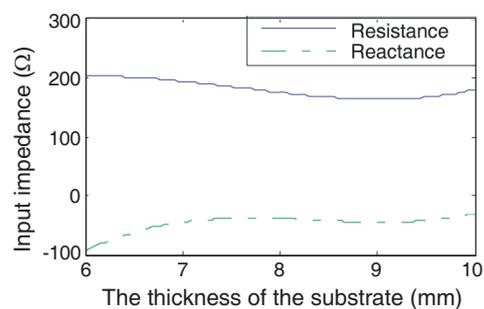


Figure 8: The input impedance as a function of thickness of the substrate at 6 GHz.

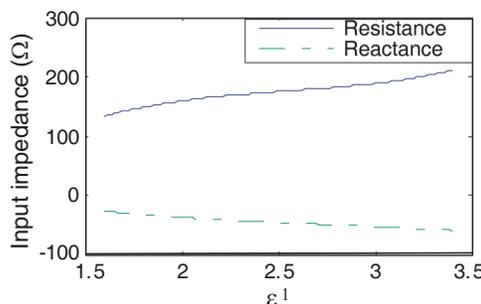


Figure 9: The input impedance as a function of ϵ_1 at 6 GHz.

4. CONCLUSION

The moment method with curved piecewise segments has been used to analyze the Archimedean spiral antenna with a dielectric superstrate, which reduces the computation time with good precision. The factors that impact the input impedance of the spiral antenna are considered. And with a dielectric superstrate, the bandwidth of the spiral antenna will be increased. Meanwhile the thickness and permittivity value of the substrate and superstrate should be chosen to make the spiral antenna have best ratio and gain over the designed bandwidth.

REFERENCES

1. Nakano, H. and J. Yamauchi, "A theoretical investigation of the two-wire round spiral antenna," *IEEE APS International Symposium*, 387–390, 1979.
2. Khamas, S. K. and G. G. Cook, "Moment-method analysis of printed wire spirals using curved piecewise sinusoidal subdomain basis and testing functions," *IEEE Trans. Antennas Propag.*, Vol. 45, 1016–1022, 1997.
3. Khamas, S. K., P. L. Starke, and G. G. Cook, "Design of a printed spiral antenna with a dielectric superstrate using an efficient curved segment moment method with optimization using marginal distributions," *IEE Proc. — Microw. Antennas Propag.*, Vol. 151, 315–320, 2004.
4. Chow, Y. L., J. J. Yang, D. G. Fang, and G. E. Howard, "A closed-form spatial Green's function for the thick microstrip substrate," *IEEE Trans. Microwave Theory Tech.*, Vol. 39, 588–592, 1991.
5. Hua, Y. and T. K. Sarkar, "Generalized pencil-of-function method for extracting poles of an EM system from its transient response," *IEEE Trans. Antennas Propag.*, Vol. 37, 229–234, 1989.
6. Aksun, M. I. and R. Mittra, "Derivation of closed-form spatial Green's functions for printed circuit structures with substrates and superstrates," *IEEE APS-Int. Symposium*, Vol. 3, 1356–1359, 1991.

Elasticity-stochastic Description on the Adhesion of Elastic Media via Molecular Bond Clusters

Jizeng Wang

College of Civil Engineering and Mechanics, Lanzhou University, Lanzhou, Gansu 730000, China

Abstract— This paper aims to study the size dependent steady-state pull-off load in molecular adhesion between two soft elastic materials. The adhesion consists of a patch of noncovalent bonds formed between ligand and receptor molecules on opposing adhesion surfaces. Classical contact mechanics is used to model the deformation of elastic materials, while Bell's model is adopted to describe stochastic breaking/reforming of molecular bonds. A coupled elastic-stochastic model is utilized to show that there exists a critical adhesion size beyond which stress concentration near the edge of contact region causes a crack like failure of the adhesion patch governed by Griffith's criterion and below which the pull-off stress is saturated at a constant value governed by Bell's model of molecular adhesion.

1. INTRODUCTION

Adhesion contact between elastic solids has been an active research topic in contact mechanics for several decades. In this field, the JKR [1] and DMT [2] models have successfully modeled adhesive contact at two opposite limits of relative length scales associated with interactive surface forces versus elastic deformation [3], while the transition between JKR and DMT is well described by the Maugis-Dugdale model [4].

To maintain a stable adhesion state, one usually needs to know the strength of a particular adhesion between two elastic solids. In the case that a pulling load is applied to the adhered elastic bodies, stress concentration is expected to occur at the edge of the contacting region. Increase of the load then enlarges the intensity of stress concentration and eventually drives the edge cracks to propagate and break the joint. In this case, the carrying capacity of the joint is not used most efficiently because only a small fraction of material is highly stressed at any instant of loading, and failure occurs by incremental crack propagation. The maximum strength should correspond to an optimal adhesion state that at pull-off the interfacial stress is uniformly distributed over the contact region with magnitude equal to the theoretical adhesion strength. Gao and Yao [5] suggest that a robust, shape-insensitive optimal adhesion becomes possible when the adhesion size is small enough. Below a critical structural size, the material fails no longer by propagation of a pre-existing crack, but by uniform rupture at the limiting strength of the binding molecules.

On the other hand, from a statistical mechanics point of view, a single molecular bond only has a finite life time. The time scale associated with individual binding/dissociation events take minutes under small stretching forces, say below $5pN$ for biotin-streptavidin bonds [6, 7]. Recent single molecule experiments for activated $\alpha_5\beta_1$ -integrin binding to fibronectin under low loading rates have indicated that the characteristic time scale for ligand-receptor binding/dissociation is on the order of 100s [8]. Although a single bond only has a limited lifetime, a cluster of bonds can survive for much longer time due to collective effects in a stochastic ensemble. A common assumption of the existing models on molecular cluster adhesion is equally sharing of applied load among all closed bonds. Based on this assumption, Erdmann and Schwarz [9] predict that the cluster lifetime monotonically increases as the cluster size grows: the larger the cluster, the more stable it is. When the cluster becomes large enough, constant adhesion strength will be reached.

It seems that the theories in classical contact mechanics and in stochastic adhesion of molecular clusters give two opposite trends for the dependence of adhesion strength on adhesion size. When coupling between elasticity and stochastic behavior is taken into account, we expect that optimal state can be achieved at a finite adhesion size. Adhesion for both large and small adhesion sizes will subject to failures due to crack spreading and stochastic deviation in bond breaking. Our aim in the present study is to develop a stochastic-elasticity model to investigate how the adhesion strength depends on the size of the molecular clusters.

2. THE MODEL

To illustrate the model problem considered in this study, Figure 1 shows the plane strain problem of two elastic half-spaces joined together by a cluster of ligand-receptor bonds forming an adhesion

patch of size $2a$ under the action of an applied force P (per unit out-of-plane thickness). Here, only specific adhesion via ligand-receptor linkages is considered, and secondary nonspecific interactions such as van der Waals forces are neglected. A set of Cartesian coordinates (x, z) are placed at the center of the adhesion region so that the two opposing surfaces are located at $z = h/2$ and $z = -h/2$. We assume that the adhesion region accommodates ρ_t pairs of ligand-receptor bonds per unit area at position x , of which ρ_b bonds are closed.

The Young's modulus and Poisson's ratio are E_1, E_2 and ν_1, ν_2 , for the lower and upper elastic half-space, respectively. The surface displacement u_1, u_2 of the lower and upper elastic half-space can be expressed in terms of the normal traction $p(x)$ along the interface as [10–13]

$$\frac{\partial u_1}{\partial x} = -\frac{2(1-\nu_1^2)}{\pi E_1} \int_{-a}^a \frac{p(s)}{x-s} ds, \quad (1)$$

$$\frac{\partial u_2}{\partial x} = \frac{2(1-\nu_2^2)}{\pi E_2} \int_{-a}^a \frac{p(s)}{x-s} ds. \quad (2)$$

For simplicity, we treat the closed bonds as Hookean spring with stiffness ξ . The interfacial traction $p(x)$ also causes bonds around x to elongate by

$$u = u_2 - u_1 \quad (3)$$

where

$$p(x) = -\xi \rho_b u. \quad (4)$$

From Eqs. (1)–(4), we have [11–13]

$$\frac{\partial u}{\partial x} = -\frac{2\xi}{\pi E^*} \int_{-a}^a \frac{\rho_b u}{x-s} ds \quad (5)$$

where E^* is the usual reduced elastic modulus

$$\frac{1}{E^*} = \frac{1-\nu_1^2}{E_1} + \frac{1-\nu_2^2}{E_2}. \quad (6)$$

The ligand-receptor bonds are assumed to undergo reversible transitions between open and closed states:



Following [14], the kinetics of bond association/dissociation can be expressed as a function of the bond force ξu as [11–14]

$$\frac{d\rho_b}{d\tau} = \gamma[\rho_t(x) - \rho_b(x, \tau)] - e^{\xi u/F_b} \rho_b(x, \tau) \quad (8)$$

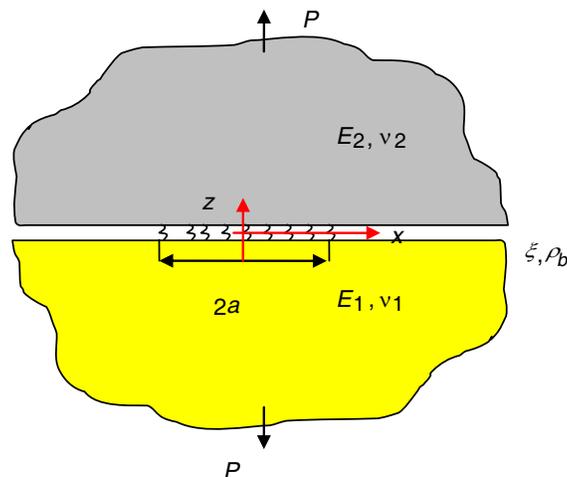


Figure 1: Adhesion of elastic solids via ligand-receptor bonds.

where k_0 denotes the dissociation rate of a bond in the absence of an applied force, $\tau = k_0 t$ is a normalized time, γ is the dimensionless rebinding rate of a pair of open bonds, and $F_b = k_B T/x_b$ is a reference force scale on the order of $4pN$, x_b being a measure of the distance between the minimum and the escape barrier of the binding potential. For the molecular patch, the initial bond density distribution is assumed as

$$\rho_t(x) = \rho_0 \Theta(x - a_0) \tag{9}$$

where ρ_0 and a_0 are constants, and

$$\Theta(x - a_0) = \begin{cases} 1, & a \geq |x| \geq a_0 \\ 0, & a_0 > |x| \end{cases} \tag{10}$$

Introduce the following dimensionless variables

$$w = \frac{u\xi}{F_b}, \quad r = \frac{x}{a}, \quad \rho = \frac{\rho_b}{\rho_0}, \quad \alpha = \frac{\xi a \rho_0}{E^*} \tag{11}$$

so that Eqs. (5) and (8) can be reduced to

$$\frac{\partial w}{\partial r} = -\frac{2\alpha}{\pi} \int_{-1}^1 \frac{\rho w}{r-s} ds, \tag{12}$$

$$\frac{d\rho}{d\tau} = \gamma \left[\Theta\left(r - \frac{a_0}{a}\right) - \rho \right] - e^w \rho. \tag{13}$$

In the steady state, Eq. (13) becomes

$$\rho_{eq} = \frac{\gamma}{\gamma + e^w} \Theta\left(r - \frac{a_0}{a}\right) \tag{14}$$

On the other hand, Eq. (12) can be transformed into [10]

$$\rho_{eq} w = -\frac{1}{2\pi\alpha(1-r^2)^{1/2}} \int_{-1}^1 \frac{(1-s^2)^{1/2} w'}{r-s} ds + \frac{Q}{(1-r^2)^{1/2}} \tag{15}$$

where $Q = P/\pi a \rho_0 F_b$ is the normalized applied load.

In order to solve the above coupled singular differential-integral equations with strong nonlinearity, we approximate w as

$$w \approx \sum_{j=0}^N w_j r^{2j} \tag{16}$$

Inserting Eq. (16) into Eq. (15) yields

$$\rho_{eq} w_0 + \sum_{k=1}^N w_k \left[\rho_{eq} r^{2k} + \frac{k}{\pi\alpha(1-r^2)^{1/2}} I_{2k-1}(r) \right] \approx \frac{Q}{(1-r^2)^{1/2}} \tag{17}$$

where

$$I_n = \int_{-1}^1 \frac{(1-s^2)^{1/2} s^n}{r-s} ds \tag{18}$$

By using the Galerkin method, Eq. (17) can be reduced to an algebraic equation

$$\mathbf{AW} = \mathbf{B} \tag{19}$$

where $\mathbf{A} = \{A_{ij}\}_{N+1, N+1}$, $\mathbf{W} = \{w_j\}_{N+1, 1}$, $\mathbf{B} = \{B_j\}_{N+1, 1}$, and

$$\begin{aligned} A_{i1} &= \int_{-1}^1 \rho_{eq} r^{2(i-1)} dr \\ A_{ik+1} &= \int_{-1}^1 \left[\rho_{eq} r^{2k} + \frac{k I_{2k-1}(r)}{\pi\alpha(1-r^2)^{1/2}} \right] r^{2(i-1)} dr, \quad N \geq k \geq 1 \\ B_j &= Q \int_{-1}^1 \frac{r^{2j-2}}{(1-r^2)^{1/2}} dr = Q \frac{\pi^{1/2} \Gamma(j - \frac{1}{2})}{(j-1) \cdot (j-1)!} \end{aligned} \tag{20}$$

From fracture mechanics, the critical condition for crack initiation in a perfectly brittle material is governed by the Griffith condition [5] $G = 2\gamma_{ad}$, where γ_{ad} is the fracture surface energy and G is the energy release rate, which in the present case can be expressed as

$$G = \frac{2K_I^2}{E^*} = \frac{2P^2}{\pi a E^*}, \quad (21)$$

At the critical point of onset of crack motion, the energy released per unit area of crack growth must be equal to the energy necessary to create a unit area of two new surfaces. For adhesive contact problems, the Griffith condition is often written as $G \geq W_{ad}$, where W_{ad} is the work of adhesion, and the critical pull-off stress can be determined as:

$$\sigma_{\text{Pull-off}} = \frac{P_{\text{Pull-off}}}{2a} = \sqrt{\frac{\pi E^* W_{ad}}{8a}}, \quad (22)$$

For the adhesion via molecular bonds, the work of adhesion can be expressed as

$$W_{ad} \sim \frac{\gamma \rho_0}{1 + \gamma} \left(\frac{1}{2} \xi u_{\text{max}}^2 \right) = \frac{\gamma \rho_0}{1 + \gamma} \left[\frac{1}{2} \xi \left(\frac{F_b}{\xi} \right)^2 \right] = \frac{\gamma \rho_0 F_b^2}{2(1 + \gamma) \xi} \quad (23)$$

where u_{max} is the maximum allowed deformation of each bond, $\gamma \rho_0 / (1 + \gamma)$ is the equilibrium bond density in the absence of an external load, and F_b / ξ is the typical length scale for the maximum bond deformation. From Eqs. (22) and (23), we can estimate the dimensionless pull-off stress as

$$\sigma_{\text{Pull-off}} \sim \frac{1}{4} \sqrt{\frac{\pi \gamma \rho_0 F_b^2 E^*}{(1 + \gamma) \xi a}} \propto a^{-\frac{1}{2}} \quad (24)$$

It can be seen from Eq. (24) that, according to the Griffith theory, the pull-off stress is proportional to $1/\sqrt{a}$. As the adhesion size a is reduced, Eq. (24) predicts an increasing pull-off stress, approaching infinity as a goes to zero. However, this cannot be true as the stress cannot exceed the theoretical strength of each molecular bond. This immediately yields a critical adhesion size below which adhesion strength can no longer be described by the Griffith theory. On the other hand, the strength of the molecular cluster can be determined as

$$\sigma_{\text{Pull-off}} \sim \frac{\gamma \rho_0}{1 + \gamma} \cdot \xi u_{\text{max}} \sim \frac{\gamma \rho_0}{1 + \gamma} \cdot \xi \frac{F_b}{\xi} = \frac{\gamma \rho_0 F_b}{1 + \gamma} \quad (25)$$

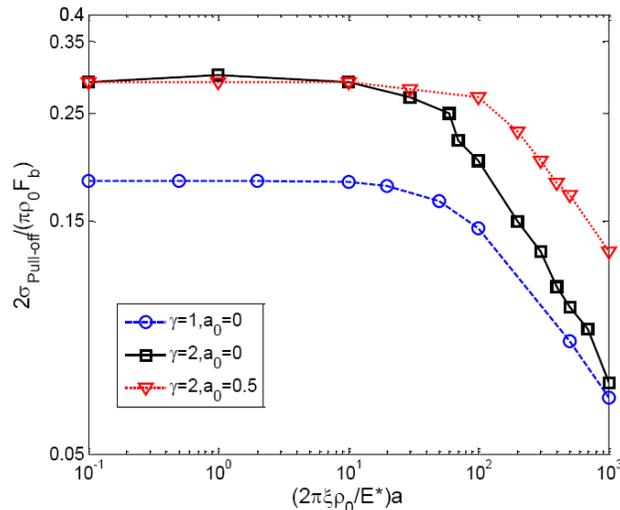


Figure 2: The normalized pull-off stress as a function of the normalized adhesion size for different rebinding rate and cluster pattern.

We thus expect that Eq. (25) gives the upper bound of adhesion strength of the molecular cluster.

Here we use a coupled elastic-stochastic model to investigate whether the continuum theory based on the Griffith concept is applicable or not in the case of molecular adhesion. By solving the coupled singular integral-differential equations, Figure 2 shows the normalized pull-off stress as a function of the normalized cluster size for different rebinding rate ($\gamma = 1$, $\gamma = 2$) and different patterns of initial bond distribution ($a_0/a = 0$, $a_0/a = 0.5$). It can be seen that there indeed exists a critical adhesion size, below which the normalized pull-off stress approaches a constant. This observation suggests a change in behavior once the size of the molecular cluster is reduced to below a critical length scale, the Griffith theory is no longer valid. Figure 2 also demonstrates that the larger rebinding rate corresponds to larger adhesion strength.

3. CONCLUSION

We have used continuum and stochastic concepts to investigate how the adhesion strength of a molecular cluster depends on cluster size. Numerical results show a smooth transition from the Griffith mode of failure via crack propagation to uniform bond rupture below a critical size.

REFERENCES

1. Johnson, K. L., K. Kendall, and A. D. Roberts, "Surface energy and contact of elastic solids," *Proc. R. Soc. Lond. A*, Vol. 324, 301–313, 1971.
2. Derjaguin, B. V., V. M. Muller, and Y. P. Toporov, "Effect of contact deformations on the adhesion of particle," *J. Colloid. Interf. Sci.*, Vol. 53, 314–326, 1975.
3. Greenwood, J. A., "Adhesion of elastic spheres," *Proc. R. Soc. Lond. A*, Vol. 453, 1277–1297, 1997.
4. Maugis, D., "Adhesion of spheres: The JKR-DMT transition using dugdale model," *J. Colloid. Interf. Sci.*, Vol. 150, 243–269, 1992.
5. Gao, H. and H. Yao, "Shape insensitive optimal adhesion of nanoscale fibrillar structures," *Proc. Natl. Acad. Sci.*, Vol. 101, 7851–7856, 2004.
6. Evans, E., "Probing the relation between force-lifetime-and chemistry in single molecular bonds," *Annu. Rev. Biophys. Biomol. Struct.*, Vol. 30, 105–128, 2001.
7. Merkel, R., P. Nassoy, A. Leung, K. Ritchie, and E. Evans, "Energy landscapes of receptor-ligand bonds explored with dynamic force spectroscopy," *Nature*, Vol. 397, 50–53, 1999.
8. Li, F., S. D. Redick, H. P. Erickson, and V. T. Moy, "Force measurements of the integrin-fibronectin interaction," *Biophys. J.*, Vol. 84, 1252–1262, 2003.
9. Erdmann, T. and U. S. Schwarz, "Stochastic dynamics of adhesion clusters under shared constant force and with rebinding," *J. Chem. Phys.*, Vol. 121, 8997–9017, 2004.
10. Johnson, K. L., *Contact Mechanics*, Cambridge University Press, 1985.
11. Qian, J., J. Wang, Y. Lin, and H. Gao, "Lifetime and strength of periodic bond clusters between elastic media under inclined loading," *Biophys. J.*, Vol. 97, 2438–2445, 2009.
12. Wang, J., J. Qian, and H. Gao, "Stability of molecular adhesion mediated by confined polymer repellers and ligand-receptor bonds," *Mol. Cell. Biomech.*, Vol. 5, 19–26, 2008.
13. Wang, J. and H. Gao, "Clustering instability in adhesive contact between elastic solids via diffusive molecular bonds," *J. Mech. Phys. Solids*, Vol. 56, 251–266, 2008.
14. Bell, G. I., "Models for the specific adhesion of cells to cells," *Science*, Vol. 200, 618–627, 1978.

Electromagnetic Elasto-plastic Dynamic Behaviors of Conductive Circular Plate

Yuanwen Gao

Key Laboratory of Mechanics on Western Disaster and Environment
College of Civil Engineering and Mechanics, Lanzhou University
Lanzhou 730000, China

Abstract— Electromagnetic elasto-plastic dynamic response of a conductive circular plate in a magnetic pulse field is studied in this paper, the influence of the strain rate effect is investigated for the electromagnetic elasto-plastic deformation of the conductive plate. Basic governing equations are derived for electromagnetic field (eddy current), the elasto-plastic transient dynamic response of a conductive circular plate, and then an appropriate numerical code is developed based on the finite element method to quantitatively simulate the magneto-elasto-plastic mechanical behaviors of the conductive circular plate. The Johnson-Cook model is employed to study the strain rate effect on the deformation of the plate. The dynamic response is explained with some characteristic curves of deformation, the eddy current, and the configurations of the conductive plate. The numerical results indicate that the strain rate effect has to be considered for the conductive plates, especially for those with high strain rate sensitivity.

1. INTRODUCTION

The magnetic-mechanical behaviors of electromagnetic materials and structures play a significant role in many practical applications. Due to the complicated multi-field coupling and nonlinear characteristics of these problems, most of researches conducted on the dynamic behaviors of conductive structures are always focused in the elastic deformation [1, 2]. There are a few researches conducted on the electromagnetic elasto-plastic behaviors on the conductive structures.

In this paper, we study the elastic-plastic dynamic response of the conductive plate under a magnetic pulse and investigate the influence of the strain rate effect on the deformation of the plate. Governing essential equations are introduced for the Maxwell electromagnetic equations, the dynamic equations of conductive plate and the plastic constitutive equations in Section 2, respectively. And then a numerical code based on the FEM is developed to quantitatively simulate the characteristics of electromagnetic dynamic response of conductive plate under a magnetic pulse in Section 3. The explicit time integration scheme and the backward-Euler algorithm for the yield surface return mapping are employed to effectively obtain the elastic-plastic results of the plate. Some numerical results on the electromagnetic elasto-plastic deformation of a plate, the influence of the strain rate effect on the deformation, the configuration of plate, the distribution of eddy current and the plastic strain at the surface of plate are illustrated in Section 4 and some conclusion are given in the Section 5.

2. GOVERNING EQUATIONS

Consider a thin conductive rectangular plate with the length of a , the width of b , and the thickness of h , which located in an applied magnetic pulse field $\mathbf{B}_0 = B_{0r}\mathbf{i} + B_{0z}\mathbf{k}$. Here \mathbf{i} , \mathbf{k} are the unit vectors in the polar coordinates (r, θ, z) . \mathbf{k} is the normal to the mid-plane of a plate. Based on the Maxwell Equation and using the Helmholtz's formula and the Biot-Savart's law for the thin plate, one can easily obtain a governing equation for the eddy current vector potential T as follows [2]

$$\nabla^2 T - \eta\mu_0 \frac{\partial T}{\partial t} - \frac{\eta\mu_0}{4\pi} \int_s \frac{\partial T'_n}{\partial t} \frac{\partial}{\partial z} \left(\frac{1}{R} \right) ds' = \eta \left[\frac{\partial B_{0z}(t)}{\partial t} - \left(B_{0r} \frac{\partial^2 w}{\partial r \partial t} \right) \right] \quad (1)$$

where σ and μ_0 are the electrical conductivity of the conductive plate and the magnetic permeability in the vacuum state, respectively. The boundary conditions and the initial conditions are respectively described as

$$r = 0, \quad |T| < +\infty, \quad r = R, \quad T = 0 \quad (2a)$$

$$t = 0, \quad T = 0 \quad (2b)$$

Thus the eddy current induced by the applied magnetic field can be described as

$$\mathbf{J}_e = \nabla \times \mathbf{T} \quad (3)$$

where $\mathbf{T} = T(r, \theta, t)\mathbf{k}$ called the eddy current vector potential. The electromagnetic force \mathbf{F} of the conductive plate, given by the Lorentz force formula, becomes

$$\mathbf{F} = F_r \mathbf{i} + F_z \mathbf{k} = \int_{-\frac{h}{2}}^{\frac{h}{2}} (\mathbf{J}_e \times \mathbf{B}) dz = h(\mathbf{J}_e \times \mathbf{B}) \quad (4)$$

According to the theory of vibration of thin plate, the dynamic governing equations of the conductive circular plate in the ax symmetric case can be written as [3]

$$\begin{aligned} r \frac{\partial N_r}{\partial r} + N_r - N_\theta + F_r(r, \theta, t) &= 0 \\ \frac{d^2}{dr^2}(rMr) - \frac{d}{dr}(M_\theta) + \left[N_r \frac{\partial^2 w}{\partial r^2} + N_\theta \left(\frac{1}{r} \frac{\partial w}{\partial r} \right) \right] + F_z(x, y, t) &= \rho h \frac{\partial^2 w}{\partial t^2} \\ \varepsilon_r - \frac{d}{dr}(r\varepsilon_\theta) - \frac{1}{2} \left(\frac{dw}{dr} \right)^2 &= 0 \end{aligned} \quad (5)$$

where ρ is the mass density, M_r and M_θ are the bending moments, respectively, N_r , N_θ are the internal forces respectively. We assume that the boundary is clamped, which can be described as

$$r = a : \quad w = \frac{dw}{dr} = 0, \quad r \frac{dN_r}{dr} + (1 - \nu)N_r = 0 \quad (6)$$

The initial conditions of vibration of a plate can be written as

$$\frac{\partial w(r, \theta, t)}{\partial t} = 0, \quad w(r, \theta, 0) = 0 \quad \text{at} \quad t = 0 \quad (7)$$

In order to investigate the influence of strain rate on the deformation of metal sheet, the Johnson-Cook material model is employed here, which can be described as [4]

$$\sigma = (A + B\varepsilon^n)(1 + C \ln \dot{\varepsilon})(1 - T^{*m}) \quad (8)$$

where T^* is the homologous temperature given by

$$T^* = \frac{T - T_{room}}{T_{melt} - T_{room}} \quad (9)$$

In which T is the temperature of workpieces, and T_{melt} is the melting temperature of workpieces. The Johnson-Cook model contains five material constants, A, B, n, C, m which are to be determined.

The Von Mises yield function with consideration of the strain hardening and the strain rate hardening is employed, which can be described as follow:

$$f = \frac{1}{\sqrt{2}} [(\sigma_r - \sigma_\theta)^2 + \sigma_r^2 + \sigma_\theta^2 + 6\tau_{r\theta}^2]^{1/2} - \sigma_Y(\varepsilon^p, \dot{\varepsilon}^p) \quad (10)$$

where ε^p , $\dot{\varepsilon}^p$ are the effective strain and the strain rate of plate respectively.

3. NUMERICAL APPROACHES

With the aid of the Galerkin finite element method, the Eqs. (1)–(3) of the eddy current of the plate, can be stated in a discrete form as [2]:

$$[\mathbf{K}^{ec}][\mathbf{T}] + [\mathbf{P}] \left[\frac{\partial \mathbf{T}}{\partial t} \right] = [F^{ec}(\mathbf{B}_0, \dot{\mathbf{w}})] \quad (11)$$

in which, $[\mathbf{N}^{ec}]$ is the shape function matrix for an eight-nodal quadrilateral element, $[\mathbf{T}]$ is the column matrix with an unknown vector potential of the eddy current, $[\mathbf{K}^{ec}]$ is the total magnetic

stiffness matrix; $[\mathbf{P}]$ is the stiffness matrix related to magnetic flux induced by the eddy current, $[\mathbf{F}^{ec}(\mathbf{B}_0, \dot{\mathbf{w}})]$ is the column matrix of force which is related to the applied magnetic field. After calculation of the eddy current vector potential component $[\mathbf{T}]$ for all discrete points, the magnetic force F_x, F_y, F_z can be calculated by Eq. (4).

In addition, from the principle of virtue work, one can easily get the finite element formulation of the dynamic equation of the plates as

$$\mathbf{M}\ddot{\mathbf{u}} + \mathbf{C}\dot{\mathbf{u}} + \mathbf{P} = \mathbf{F}([\mathbf{T}]) \quad (12)$$

where $\mathbf{u} = \{u, v, w\}$ is the displacement vector, $\dot{\mathbf{u}}$, $\ddot{\mathbf{u}}$ are the velocity and acceleration vectors, respectively. \mathbf{C} is the damping matrix. \mathbf{P} , \mathbf{F} are the matrices of internal force and external force, respectively.

In order to effectively achieve integrating of the flow equation of plastic deformation, a backward-Euler scheme [5] is adopted to update the stresses. The increment of the strain and stresses at the respective iteration are described as

$$d\varepsilon_{i+1}^{pl} = \frac{f_i - \mathbf{a}^T \mathbf{Q}^{-1} \mathbf{r}_i}{\mathbf{a}^T \mathbf{Q}^{-1} \mathbf{Y} \mathbf{a} + H'}, \quad d\sigma_{i+1} = -\mathbf{Q}^{-1} \mathbf{r}_i - d\varepsilon_{i+1}^{pl} \mathbf{Q}^{-1} \mathbf{Y} \mathbf{a} \quad (13)$$

where $\mathbf{a} = \frac{\partial f}{\partial \sigma}$, \mathbf{Y} is the elastic constant matrix,

$$\mathbf{Q} = \left(\mathbf{I} + \Delta\varepsilon_i^{pl} \mathbf{Y} \frac{\partial \mathbf{a}}{\partial \sigma} \right), \quad H' = \frac{\partial f}{\partial \varepsilon^p} + \frac{\partial f}{\Delta t \partial \varepsilon^p} + \frac{\partial f}{\partial T^{temp}} \frac{\beta \sigma^{eff}}{C\rho} \quad (14)$$

4. NUMERICAL RESULTS

Based on the finite element formulae introduced above, we conduct some calculation on the electromagnetic elasto-plastic dynamic responses of the conductive rectangular plate. In the simulation, the radius and the thickness of plate are assigned as $r = 0.2\text{m}$, $h = 0.003\text{m}$. The magnetic pulse is produced by the coils with the electric current, the applied current in the coil is given as $I = I_0 \exp(-t/\tau)$. Here I_0 , τ are the pulse parameters. In following simulations, we set $\tau = 0.06$. The material employed is SPCC. The material constants of SPCC are assigned as $A = 321.03$, $B = 453$, $n = 0.691$, $m = 0$.

Figure 1 displays the elastic-plastic transient dynamic response at the location ($r = 0$) of plate with the different applied electric current. It can be found the influence of the strain rate on the deformation of plate is remarkable and thus the strain rate effect should be considered in the quantitatively simulation or experiments, especially for those plates whose material is sensitive to the strain rate, such as SGACD, SPCC.

Figure 2 displays the distribution of the eddy current at $t = 0.06\text{s}$, which shows in the neighborhood of the center of the plate, the eddy current is very small.

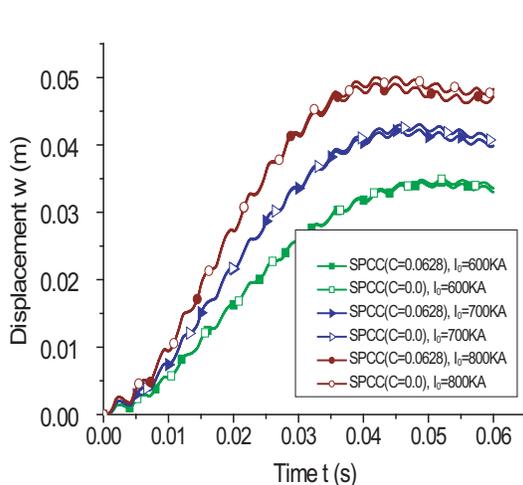


Figure 1: The elastic-plastic dynamics response with the different applied current.

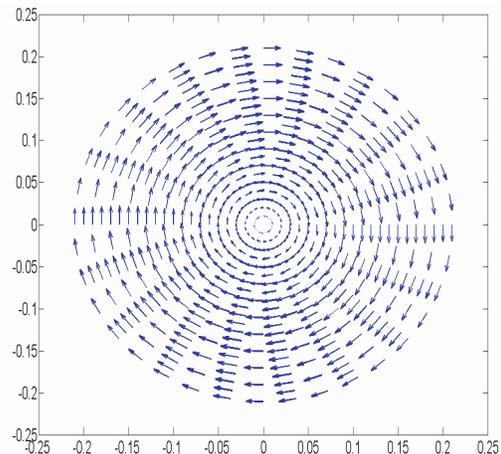


Figure 2: The plot of eddy current in the conductive plate.

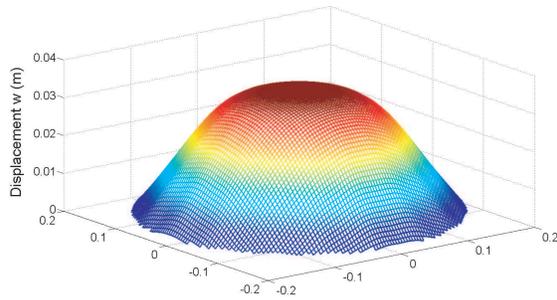


Figure 3: The configuration of the conductive plate at $t = 0.06$ s.

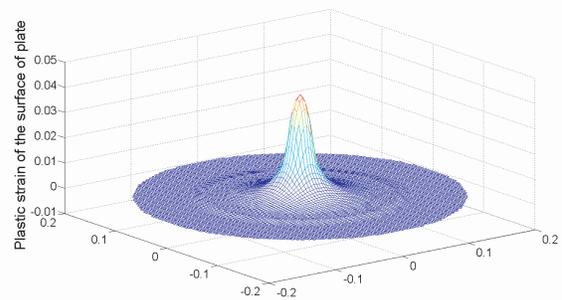


Figure 4: The effective plastic strain at the surface of plate at $t = 0.06$ s.

Figure 3 displays the configuration of the conductive plate under the magnetic pulse at $t = 0.06$ s which shows rather smooth deformation.

Figure 4 displays the effective plastic strain at the surface of the conductive plate at $t = 0.06$ s, which shows the effective plastic strain at the center of the plate surface is the largest of the plate, and the effective plastic strain at the boundary of the coils ($r = 0.13$ m), which is used to produced the magnetic pulse has an obviously change.

5. CONCLUSIONS

The influences of the plastic strain rate on the transient dynamic response/or the deformation of conductive plates under a magnetic field produced by the coils with electric current has been investigated. Numerical simulations of the conductive plates provide the dynamic response curves, the configuration, and the distribution of the eddy current, the effective plastic strain at the surface of plate in this paper. The numerical results demonstrated that the strain rate effect should be considered in the electromagnetic forming process

ACKNOWLEDGMENT

This paper was (partially) supported by the Research Grants from the Natural Science Fund of China (10672070), Program for New Century Talented (NCET-06-0896).

REFERENCES

1. Lee, J. S., J. H. Prevost, and P. C. Y. Lee, "Finite element analysis of magnetically induced vibration of conductive plates," *Fusion Eng. Des.*, Vol. 13, 125–141, 1997.
2. Zheng, X. J., J. P. Zhang, and Y. H. Zhou, "Dynamic stability of a cantilever conductive plate in transverse impulsive magnetic field," *Int. J. Solids and Struct.*, Vol. 42, 2417–2430, 2005.
3. Timoshenko, S., *Theory of Plates and Shells*, McGraw-Hill, New York, 1959.
4. Huh, H., W. J. Kang, and S. S. Han, "A tension split hopkinson bar for investigating the dynamic behavior of sheet metals," *Exp. Mech.*, Vol. 42, No. 1, 8–17, 2002.
5. Crisfield, M. A., *Non-linear Finite Element Analysis of Solids and Structures*, John Wiley & Sons, Chichester, 1991.

Rearrangement of Martensitic Variants and Mechanical-magneto-thermal Behavior of a Ferromagnetic Shape Memory Alloy Rod

Xingzhe Wang, Fang Li, and Xuebing Han

Key Laboratory of Mechanics on Western Disasters and Environment, Ministry of Education of China
College of Civil Engineering and Mechanics, Lanzhou University, Lanzhou 730000, China

Abstract— Ferromagnetic shape memory alloys (FSMAs) are potential candidates for new generation of magneto-mechanical actuators and sensors due to their high energy, large strain and high bandwidth properties. It is revealed that the giant reversible field-induced strain in FSMAs is from the rearrangement of martensitic variants. In order to describe the rearrangement mechanism related to the mobility of twin boundaries, a coupled mechanical-magneto-thermal model is proposed in this study. A general thermodynamic driving force to move the twin boundaries is derived based on the energy balance and the change of entropy at the twin boundary of FSMAs. An explicit evolution formulation for the mass fraction change between the different martensitic variants, which is the thermodynamic flux conjugate to the thermodynamic driving force, is obtained when the energy dissipation of system reaching a maximum. The model is used to a single crystal FSMA rod with an applied transverse magnetic field and an axial compression for characterizing the martensitic variants revolution and the mechanical-magneto-thermal behaviors. Some results are obtained and compared to the experimental measurements.

1. INTRODUCTION

Ferromagnetic shape memory alloys (FSMAs) are attaining more interests as a new class of magneto-mechanical smart materials exhibiting giant magnetic-field-induced strain which is called as magnetic shape memory effect (MSME). It has been shown that the mechanism of the MSME is due to the reorientation of twinned structures between different martensite variants by way of the movement of the martensite twin-boundary as an external magnetic field is applied to samples [1]. As one of the typical and proverbially investigated FSMAs, NiMnGa can produce giant strains of 6% and 10% in the martensite phase as a kind of ferromagnetic shape memory alloy via twin-boundary motion.

Recent work shows that the movement of the twin-boundary of the FSMAs is induced not only by a thermal field and/or a stress field, but also by a magnetic field. Intensive studies and attempts on the MSME at room temperature in FSMAs were performed theoretically [2, 3]. Some investigations on the temperature dependence of the maximum strain and the switching field defined as the magnetic field when the reorientation happened were also reported [4]. A model for describing the temperature dependence of the mechanism of the twin-boundary motion has been established [5]. However, the relationship between the temperature and the macroscopic reorientation behavior has not been clarified in their model.

In the present paper, a coupled mechanical-magneto-thermal model is developed to describe the macroscopic reorientation behavior of FSMAs under magnetic and stress fields. The macroscopic reorientation behavior in the low temperature region, the high temperature limit of which is the temperature of transformation to austenite will be predicted based on the model. The temperature dependence of the magnetization and the maximum strain are considered here.

2. COUPLED MECHANICAL-MAGNETO-THERMAL MODEL

To model the macroscopic reorientation behavior of FSMAs, the description of a general driving force to cause the twin-boundary motion is a very pivotal step. For the simplicity, we consider a FSMA sample totally in the Martensite phase below the phase transition temperature at which the FSMA finished the phase transition from Austenite to Martensite. Fig. 1 shows a simplified two martensite variants microstructure consisting of a field-preferred variant with mass fraction ξ and a stress-preferred variant with mass fraction $1 - \xi$ which is denoted by the subscript $\eta = 1, 2$, respectively.

In the two variants, the magnetization vectors can be respectively expressed by

$$\mathbf{M}_1 = M^{sat}(T)\mathbf{e}_x, \quad \mathbf{M}_2 = M^{sat}(T)\mathbf{e}_y \quad (1)$$

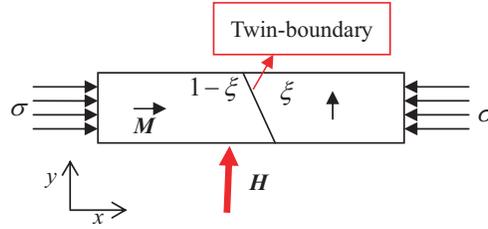


Figure 1: Simplified twin variant model.

where M^{sat} is the saturation magnetization which is usually a variable of temperature dependence [5].

From the mechanical point of view, phase transformation kinetics is associated with the deformation of phase interface or twin-boundary which leads to a volume change of materials of different martensite variants [6], we can write the energy balance law at the twin-boundary in the form

$$\rho [u] V + [\sigma V - Q] - \mu_0 [\mathbf{H} \cdot \mathbf{M}] V = 0 \quad (2)$$

where ρ , u , V , Q , σ , μ_0 , \mathbf{H} and \mathbf{M} respectively represent the mass density of the sample, the internal energy, the velocity of a material point, the heat flux, the external stress, the permeability of vacuum, the applied magnetic field vector and the magnetization vector. The term σV represents the mechanical energy power, and $\mu_0 \mathbf{H} \cdot \mathbf{M}$ represents the Zeeman energy from magnetic field. The sign $[A] = A^+ - A^-$ denotes a jump at the twin-boundary.

The change of entropy induced by the twin-boundary motion is expressed to be

$$-\gamma = \rho V [s] - [Q/T] \leq 0 \quad (3)$$

in which γ is the entropy production due to the twin-boundary motion, s is the entropy and T is the absolute temperature.

From Equations (1)–(3) and after tediously derivations, there gives the general driving force as follows

$$\pi = \frac{1}{2\rho} (\varepsilon_1 \sigma_1 - \varepsilon_2 \sigma_2) - \frac{1}{\rho} (\sigma_1 - \sigma_2) + \frac{\mu_0}{\rho} H M^{sat} + (a + b(T - T_0))(T - T_0) \quad (4)$$

Here T_0 is the reference temperature, σ_η ($\eta = 1, 2$) is the stress of the two variants, a , b are parameters related with chemical potential. The saturation magnetization dependence on the temperature by the form [5] $M^{sat} = M_0^{sat}(1 - (T/T_c)^2)$, where M_0^{sat} is the saturation magnetization at $T = 0$ K, T_c is the Curie temperature of the sample. The strains in two variants ε_η ($\eta = 1, 2$) composed by elastic strain ε_η^e ($\eta = 1, 2$) and thermal expansion strain ε_η^{th} ($\eta = 1, 2$), can be expressed

$$\varepsilon_1 = \varepsilon_1^e + \varepsilon_1^{th}, \quad \varepsilon_2 = \varepsilon_2^e + \varepsilon_2^{th} \quad (5)$$

The energy dissipation caused by the twin-boundary motion can be further denoted as

$$T\gamma = \rho V \pi \quad (6)$$

Here we will develop the evolution law for mass fraction ξ . The two assumptions are adopted for describing the rearrangement process of variants. Firstly, the dissipation induced by the twin-boundary motion reaches a maximum during the reorientation process based on the second law of thermodynamics [6], that is,

$$\pi \dot{\xi} \rightarrow \text{Max} \quad (7)$$

Secondly, the transformation starts only after a certain threshold value is reached. The driving force π dependence upon a threshold function, that is,

$$\Phi(\pi, \xi) = (\beta\pi)^2 - g(k) = 0 \quad (8)$$

in which

$$g(k) = \Phi_{cr} + \frac{c_1}{c_2} ((1 - c_3)(1 - \exp(-c_2 k)) + c_2 c_3 k) \quad (9)$$

It is similar to the procedure often used in material laws which describe plasticity. Here the phenomenological approach $g(k)$ is chosen and the parameters c_1, c_2, c_3 can be observed from experiments. The evolution of the internal variables β, k used to describe the reorientation behavior as follows [6]

$$\dot{\beta} = -\pi^{-1}\dot{\pi}\beta\xi, \quad \dot{k} = \beta\pi\dot{\xi} \quad (10)$$

The two variants are modeled with elastic moduli of E_1 and E_2 , and thermal expansion coefficients of α_1 and α_2 respectively. The effective elastic module E and the effective thermal expansion coefficient α of the sample are formulated by applying averaging method

$$E = (1 - \xi)E_1 + \xi E_2, \quad \alpha = (1 - \xi)\alpha_1 + \xi\alpha_2 \quad (11)$$

which in combination with the constitutive relations

$$\sigma = E\varepsilon^e, \quad \varepsilon^{\text{th}} = \alpha(T - T_0) \quad (12)$$

The total strain rate is considered to be composed of elastic strain rate, thermal expansion strain rate and the reorientation strain rate, i.e.,

$$\dot{\varepsilon} = \dot{\varepsilon}^e + \dot{\varepsilon}^{\text{th}} + \dot{\varepsilon}^r \quad (13)$$

Here, the elastic strain rate $\dot{\varepsilon}^e$, the thermal expansion strain rate $\dot{\varepsilon}^{\text{th}}$ of the sample, and reorientation strain rate which is linearly proportional to the rate of mass fraction $\dot{\xi}$, are expressed as

$$\dot{\varepsilon}^e = \frac{\dot{\sigma}E - \sigma\dot{\xi}(E_2 - E_1)}{E^2}, \quad \dot{\varepsilon}^{\text{th}} = (\alpha_2 - \alpha_1)(T - T_0)\dot{\xi} + \alpha\dot{T}, \quad \dot{\varepsilon}^r = \varepsilon^{\text{max}}(T)\dot{\xi} \quad (14)$$

where $\varepsilon^{\text{max}}(T)$ denotes the maximum reorientation strain which is temperature dependent.

By Equations of (4)–(10), (13) and (14), the mass fraction rate of the field-preferred variant further can be written as

$$\begin{aligned} \dot{\xi} = & \frac{1}{AE} (\sigma_1 - \sigma_2 + \varepsilon_1 E_1 - \varepsilon_2 E_2 + 2(E_2 - E_1)) \dot{\sigma} + \frac{2}{A} \mu_0 M^{\text{sat}} \dot{H} \\ & + \frac{1}{A} \left((\sigma_1 - \sigma_2) \alpha + 4\mu_0 H M_0^{\text{sat}} (T/T_c)^2 + 2\rho c_4 + 4\rho c_5 (T - T_0) \right) \dot{T} \end{aligned} \quad (15)$$

in which

$$\begin{aligned} A = & \frac{4\rho\beta(1 - \xi)}{c_1((1 - c_3)\exp(-c_2 k) + c_3)} - \left((\sigma_1 - \sigma_2)(\alpha_2 - \alpha_1)(T - T_0) \right. \\ & \left. - \frac{(E_2 - E_1)\sigma}{E^2} (\sigma_1 - \sigma_2 + \varepsilon_1 E_1 - \varepsilon_2 E_2 + 2(E_2 - E_1)) \right) \end{aligned} \quad (16)$$

3. RESULTS AND DISCUSSION

In this section, some simulations are performed based on the model developed in previous section to describe the macroscopic behavior of mechanical-magneto-thermal behaviors and rearrangement characteristics of martensite variants of a FSMA rod, as shown in Fig. 1. Fig. 2 shows the magnetic field-strain with applied field for different temperatures which are compared with the experimental measurements to show good agreement. It can be seen that the maximum strain and switching field decrease as an increase of the temperature which can be ascribed to the decrease of the elastic energy needed for the twin-boundary motion. A temperature threshold exists for the beginning of the reorientation which has been indicated in the experimental measurements [7]. Thermal dilatation strain is just observed as the temperature increasing without a magnetic field and/or an external stress. An abrupt strain is detected at the temperature threshold when a magnetic field is applied. These simulation results agree well with the experimental measurements.

The strains varying with applied magnetic field and the temperature at different external stresses are plotted in Fig. 4 and Fig. 5 respectively. As the external stress is increased at a constant temperature with value of 168 K or a magnetic field 0.655 T, the maximum rearrangement strain is decreased which is due to the increase of the initial elastic strain. Here the initial elastic strain is not considered and the constitutive response of the model is predicted, and more elastic energy needed

to be overcome for the rearrangement to start, resulting in a slowing down of the reorientation which is reflected by the slopes of the magnetic field-mass fraction curves and the mass fraction versus temperature. The corresponding evolutions of mass fraction ξ at different external stresses are shown in Fig. 6 and Fig. 7 respectively. As shown in Fig. 6 and Fig. 7, the slopes of the magnetic field-mass fraction curves at a constant temperature of 168 K or the mass fraction versus temperature curves at a magnetic field of 0.655 T decrease as the external stress increases, and the field-preferred variant grows geometrically at the expense of the stress-preferred variant and the mass fraction of the field-preferred ξ is changed from 0 to 1 during the rearrangement process of variants.

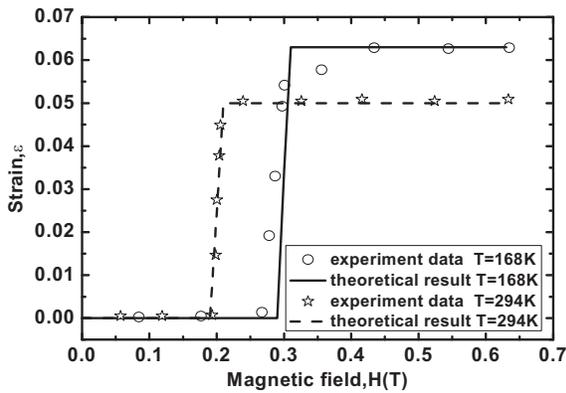


Figure 2: Strain vs. magnetic field at different temperatures.

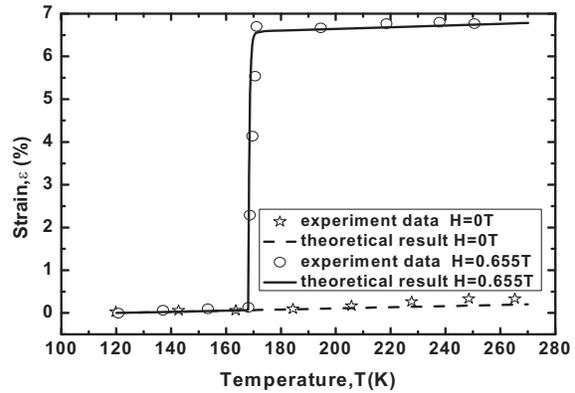


Figure 3: Strain vs. temperature at different magnetic fields.

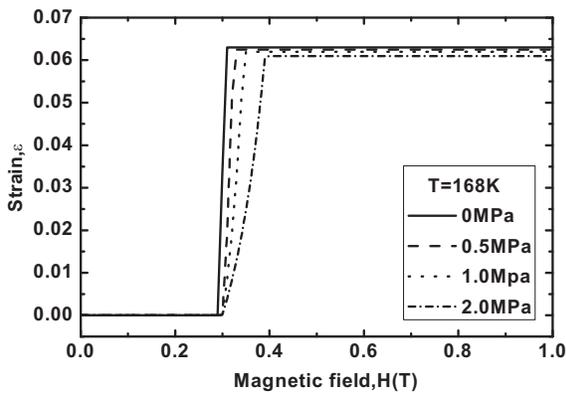


Figure 4: Strain vs. magnetic field at different stresses.

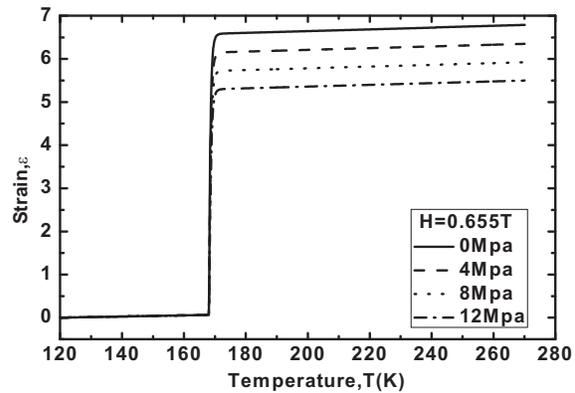


Figure 5: Strain vs. temperature at different stresses.

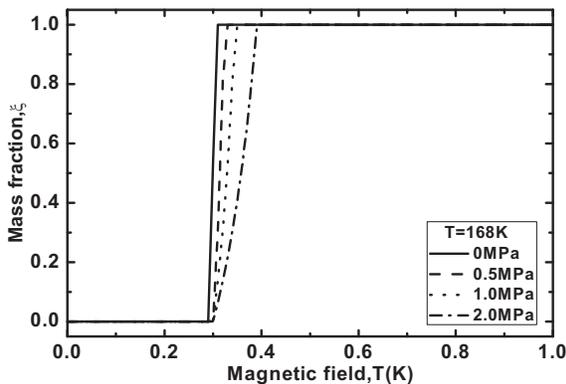


Figure 6: Mass fraction vs. magnetic field at different stresses.

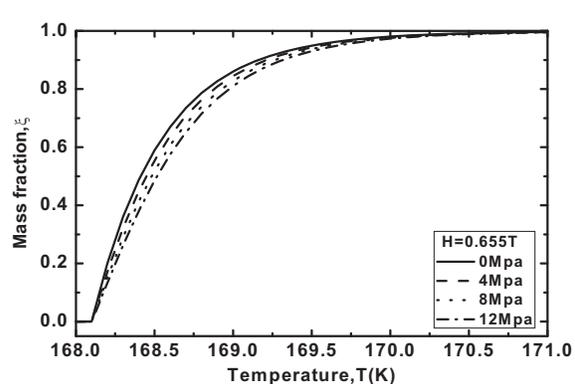


Figure 7: Mass fraction vs. temperature at different stresses.

4. CONCLUSION

A driving force for twin-boundary motion and the explicit expression of the mass fraction rate for martensite variants rearrangement are derived based on a phase transformation kinetics approach. The model can be used to describe the macroscopic rearrangement of the martensitic variants and the mechanical-magneto-thermal behaviors of the FSMAs. The theoretical and predicted results are in good agreement with the experimental measurements.

ACKNOWLEDGMENT

The authors acknowledge supports by the National Natural Science Foundation of China (10872081) and the Fok Ying-Tong Education Foundation for Young Teachers in the Higher Education Institutions of China (111005).

REFERENCES

1. Enkovaara, J., A. Ayuela, L. Nordstrom, and R. M. Nieminen, "Structural, thermal, and magnetic properties of Ni₂MnGa," *J. Appl. Phys.*, Vol. 91, Vol. 10, 7798–7800, 2002.
2. Kiefer, B. and D. C. Lagoudas, "Modeling of the magnetic field-induced martensitic variant reorientation and the associated magnetic shape memory effect in MSMA," *Smart Structures and Materials 2005: Active Materials: Behavior and Mechanics*, Vol. 5761, 454–465, 2005.
3. Sarawate, N. N. and M. J. Dapino, "A continuum thermodynamics model for the sensing effect in ferromagnetic shape memory Ni-Mn-Ga," *J. Appl. Phys.*, Vol. 101, Vol. 123522, 1–11, 2007.
4. Heczko, O. and L. Straka, "Temperature dependence and temperature limits of magnetic shape memory effect," *J. Appl. Phys.*, Vol. 94, Vol. 11, 7139–7143, 2003.
5. O'Handley, R. C., D. I. Paul, et al., "Model for temperature dependence of field-induced strain in ferromagnetic shape memory alloys," *Mater. Sci. Eng. A*, Vol. 438–440, 445–449, 2006.
6. Oberst-Brandenburg, C. and O. T. Bruhns, "A tensorial description of the transformation kinetics of the martensitic phase transformation," *Int. J. Plasticity*, Vol. 20, 2083–2109, 2004.
7. Jiang, C. B., J. M. Wang, and H. B. Xu, "Temperature dependence of the giant magnetostrain in a NiMnGa magnetic shape memory alloys," *Appl. Phys. Lett.*, Vol. 86, Vol. 252508, 1–3, 2005.

Analysis on Absorption and Thermal Stress of a Functionally Graded-absorbing Infinite Plate in Electromagnetic Fields

Hongyan Tian, Xingzhe Wang, and Youhe Zhou

Key Laboratory of Mechanics on Western Disasters and Environment
Ministry of Education of China, College of Civil Engineering and Mechanics
Lanzhou University, Lanzhou, Gansu 730000, China

Abstract— Based on the electromagnetic wave propagation theory and thermo-elasticity, this paper presents a solution to the inhomogeneous problem of functionally graded-absorbing material (FGM) of infinite plate in electromagnetic wave circumstance. The electromagnetic wave properties, including the absorption and reflection, as well the heat energy of the graded-absorbing material, are analyzed theoretically. The transient temperature and thermal stress fields in the graded-absorbing plate are calculated by mean of finite difference method. The results show that, with the proper material property design on the grade-absorbing plate there are good performance on the lower reflection and high absorption as well the gentle thermal stress distribution.

1. INTRODUCTION

Structural absorption materials become more important and have been attracted more attentions owing to the potential utilization of electromagnetic devices in industrial, commercial and military applications. There are various methods are employed to improve the absorption efficiency and the effective absorbing band of the absorber, such as adding optimized absorbent, compounding different types of absorbents and using multi-layer structure [1–3]. Some research on the absorption properties response to heat are also reported, like interaction between temperature and absorption [4, 5]. It turned out that the reflection loss of multi-layer structure is below -10 dB or -15 dB in 8–18 GHz. However, high performances and multi-function-absorbing materials and structures are increasingly demanded in various engineering applications where there always associate with the complex environments such as heat, humidity and corrosion. The gradient materials with attractive advantages of great design freedom, easy-to-show broadband, impedance gradient, high-temperature load characteristics are good absorbing material candidates. The experiment on the properties of electromagnetic wave absorption of functionally graded material was reported by Liu et al. to show that the graded plates have considerable absorbability for broadband frequencies and the reduction of thickness and weight [6]. Yoshihiro [7] conducted a theoretical analysis of the graded composition and plate thickness effects on the electromagnetic wave absorption. The recent research by Pendry and Christos [8, 9] proved that gradient shell can exhibit marvelous behavior even to make objects not visible. However, their focuses were on the electromagnetic wave properties in the graded medium and the temperature distribution and mechanical behavior have not been reported so far.

This paper presents a solution to the inhomogeneous problem of graded-absorbing material of infinite plate in electromagnetic wave circumstance. The electromagnetic wave properties, the transient temperature and thermal stress distribution of the graded absorbing material are analyzed and calculated. The results are expected to be helpful for designing graded-absorbing material parameters and mechanical parameters.

2. MODEL ANALYSIS

2.1. Electromagnetic Field Analysis of FGM Plate

We consider a functionally graded-absorbing infinite plate which is immersed in an electromagnetic wave circumstance as sketched in Fig. 1. The plate has nonhomogeneity of material properties only through the thickness (z -axis). Exact solutions cannot be obtained, as electromagnetic field in the plate is arbitrary inhomogeneity. Here FGM plate is considered to be a multilayered plate and the material properties of each layer are homogeneous. The incident electromagnetic wave (EM wave) with angular frequency ω propagates in the positive z direction. Superscripts I , R , F , B , T denote the incident, reflected, forward, backward and transmission waves, respectively. ε_i , μ_i , σ_i respectively indicates permittivity, permeability and conductivity of i th-layer, as shown in Fig. 2.

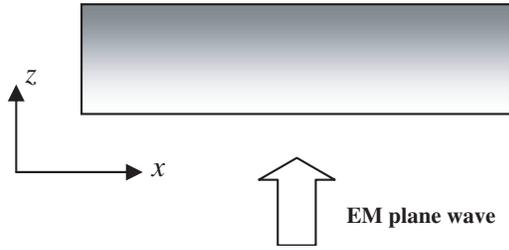


Figure 1: FGM plate with EM plane wave.

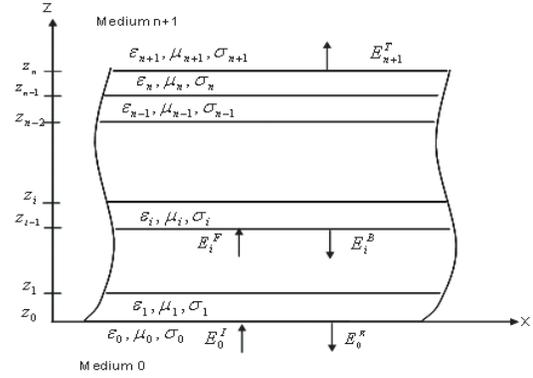


Figure 2: Analytical model of the FGM plate.

Based on the Maxwell's theory of electromagnetic wave, the electromagnetic field in each medium and each layer satisfies the following governing equation and boundary condition

$$\nabla \times \mathbf{E}_i(\mathbf{r}) + j\omega\mu_i\mathbf{H}_i(\mathbf{r}) = 0, \quad \nabla \times \mathbf{H}_i(\mathbf{r}) - (j\omega\epsilon_i + \sigma_i)\mathbf{E}_i(\mathbf{r}) = 0, \quad (i = 1, 2, \dots, n) \quad (1)$$

$$E_{yi}(z, t) = E_{yi+1}(z, t), \quad H_{xi}(z, t) = H_{xi+1}(z, t), \quad (z = z_i, i = 1, 2, \dots, n) \quad (2)$$

By solving the above time harmonic source free wave equations in one dimension, the general solution of the electric and magnetic fields in the i th-layer are [6]

$$E_{yi} = E_{yi}^F + E_{yi}^B = (A_i e^{-\gamma_i z} + B_i e^{\gamma_i z}) e^{j\omega t}, \quad -H_{xi} = H_{xi}^F + H_{xi}^B = \frac{\gamma_i}{j\omega\mu_i} (A_i e^{-\gamma_i z} - B_i e^{\gamma_i z}) e^{j\omega t}, \quad (i = 1, 2, \dots, n) \quad (3)$$

where $\gamma_i = \sqrt{(j\omega\epsilon_i + \sigma_i)(j\omega\mu_i)}$, and coefficients A_i, B_i denote the amplitude factor which can be obtained by boundary condition of Eq. (2). The values of A_i, B_i are further expressed by

$$\begin{bmatrix} A_i \\ B_i \end{bmatrix} = \frac{1}{2} \mathbf{M}_i \begin{bmatrix} A_{i+1} \\ B_{i+1} \end{bmatrix}, \quad (i = 1, 2, \dots, n); \quad B_{n+1} = 0 \quad (4)$$

in which $\mathbf{M}_i = \begin{bmatrix} (1 + u_{i+1}/u_i)e^{(v_i - v_{i+1})z_i} & (1 - u_{i+1}/u_i)e^{(v_i + v_{i+1})z_i} \\ (1 - u_{i+1}/u_i)e^{-(v_i + v_{i+1})z_i} & (1 + u_{i+1}/u_i)e^{-(v_i - v_{i+1})z_i} \end{bmatrix}$ is the characteristic matrix that determines the propagation behavior of the electromagnetic wave at each boundary, and $u_i = \gamma_i/(j\omega\mu_i)$, $v_i = \gamma_i$. Eq. (4) is the recurrence formula for the amplitude factor of the electromagnetic wave in each layer, the product of the \mathbf{M}_i is obtained as the following square matrix

$$\bar{\mathbf{M}} = \prod_{i=0}^n \mathbf{M}_i = \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} \quad (5)$$

The reflection ratio R_r at $z = z_0$, the transmission ratio T_r at the $z = z_n$, and the reflection coefficient R_{dB}, T_{dB} and the absorbing coefficient S , can be derived as

$$R_r = \left| \frac{m_{21} e^{2\gamma_0 z_0}}{m_{11}} \right|, \quad T_r = \left| \frac{2^{n+1} e^{-(\gamma_{n+1} - \gamma_0)z_n}}{m_{11}} \right|, \quad R_{dB} = -20 \log_{10} |R_r|, \\ T_{dB} = -20 \log_{10} |T_r|, \quad S = 1 - (R_r^2 + T_r^2) \quad (6)$$

2.2. Temperature Field and Thermal Stress Analysis of FGM Plate

For the absorbing materials, the EM energy is absorbed and converted to heat rather than reflect it back. The power absorbed by the graded-absorbing plate form EM wave in the i th layer can be expressed as

$$P_d(z_i, t_n) = \frac{\omega}{2} \epsilon_0 \epsilon_i'' |E_i(z_i, t_n)|^2 \quad (7)$$

where ε_i'' is the imaginary part of complex permittivity of $\varepsilon = \varepsilon' - j\varepsilon''$.

The energy balance equation of the FGM due to the electromagnetic wave assisted heat source gives

$$\rho(z)c(z)\frac{\partial T(z,t)}{\partial t} = \frac{\partial}{\partial z} \left[k(z)\frac{\partial T(z,t)}{\partial z} \right] + P_d(z,t) \quad (8)$$

where ρ is the density, c is the specific heat, and k is the thermal conductivity of the medium.

Now we consider the thermal stress in the FGM plate associated with the temperature field distribution of $T(z,t)$. The quasi-static analysis is restricted in present study. The generated inertia is small and can be neglected as the incident wave power is not large. The strain components $\hat{\varepsilon}_x$, $\hat{\varepsilon}_y$ and the stress components σ_x , σ_y are given by the relations

$$\hat{\varepsilon}_x = \hat{\varepsilon}_y = \hat{\varepsilon}_0 + \bar{z}/\rho_0 \quad (9)$$

$$\sigma_x(\bar{z}, t) = \sigma_y(\bar{z}, t) = \frac{\hat{E}(\bar{z})}{1 - \nu(\bar{z})} \times \left[\hat{\varepsilon}_0 + \frac{\bar{z}}{\rho_0} - \alpha(\bar{z})(T(\bar{z}, t) - T_0) \right] \quad (10)$$

where $\alpha(\bar{z})$, $\hat{E}(\bar{z})$, $\nu(\bar{z})$ are the coefficients of linear thermal expansion, the Young's modulus, and the Poisson's ratio, respectively. $\hat{\varepsilon}_0$ and $1/\rho_0$ denote the strain component and the curvature at the surface $z = 0$, they are unknown constants and determined from the mechanical conditions of the plate.

Assuming that the plate is traction-free, the following equilibrium relations for bending moment and in-plane force are given to evaluate the unknown constants

$$\int_0^1 \sigma_x d\bar{z} = 0 \quad \int_0^1 \bar{z}\sigma_x d\bar{z} = 0 \quad (11)$$

Finally, the thermal stress distribution in a transient state for the plate is given by

$$\sigma_x(\bar{z}, t) = \sigma_y(\bar{z}, t) = \frac{\hat{E}(\bar{z})}{1 - \nu(\bar{z})} \times \left\{ \frac{(B_2 D_0 - B_1 D_1) + (-B_1 D_0 + B_0 D_1)\bar{z}}{B_0 B_2 - B_1^2} - \alpha(\bar{z})(T(\bar{z}, t) - T_0) \right\} \quad (12)$$

in which $B_i = \int_0^1 \frac{\hat{E}(\bar{z})}{1 - \nu(\bar{z})} \bar{z}^i dz$, ($i = 0, 1, 2$), $D_i = \int_0^1 \frac{\hat{E}(\bar{z})\alpha(\bar{z})[T(\bar{z}, t) - T_0]}{1 - \nu(\bar{z})} \bar{z}^i dz$, ($i = 0, 1$).

As absorption power P_d is a function of electric field as seen in Eq. (7) and, hence, a functional representation of the electric field is necessary to solve the energy balance equation. The evaluation of the functional form of the electric field may be difficult for a multilayered sample, and we are unaware of such a solution till date. Alternatively, the energy balance and thermal stress are solved numerically.

3. EXAMPLE AND RESULTS

In this section, the finite difference method is employed in analyzing the transient temperature and thermal stress fields in the graded-absorbing plate. For calculating the temperature distribution, we use the convective boundary condition which considers heat transport and heat generation at the surface, where $k_0(\partial T/\partial z)|_{z=0} = h_t(T|_{z=0} - T_0)$, $-k_n(\partial T/\partial z)|_{z=z_n} = h_t(T|_{z=z_n} - T_0)$, with $T_0 = 300$ K, $h_t = 10$ W/(m² · K). The material parameters of the absorber vary as a function of the plate thickness, which is taken as $k = 0.188(1 + (z/z_n))$, $\rho = 1200(1 + (z/z_n))$ kg/m³, $c = 1046(1 + (z/z_n))$ J/(K · kg), $\alpha = 7.2e - 6(1 + (z/z_n))/K$, $\hat{E} = 1573(1 + (z/z_n))$ mpa, $\varepsilon_i = 3.5 + 12 \cdot (z/z_n) - j(0.1 + 4.5 \cdot (z/z_n))$.

Figure 3 shows the reflectivity R_{dB} versus frequency with different thickness for FGM plate by black color and for homogeneous plate (material properties uniform along thickness) by grey color. It is shown that the reflection loss at frequency range of 2 ~ 18 GHz is reduced and electromagnetic wave absorption performance is improved in the FGM plate compared to that of the homogeneous plate. The maximum reflectivity reaches 18.5 dB at 5.2 GHz, 10 GHz with 10 mm, 5 mm, respectively, and the reflectivity range above 8 dB is from 9 to 18 GHz with the 10 mm thickness or more. Additionally the peaks obviously shift toward high frequencies with reducing thickness of the graded-absorbing plate. The calculated absorptivity S is given in Fig. 4. There shows that the absorptivity in the FGM plate is improved by approximately 15% compared to homogenous one, and the FGM structure achieves an absorptivity of at least $S = 80\%$ within a frequency bandwidth

of 4 GHz when the thickness is more than 10 mm. However, for the very thin plate, there is not enough thickness to attenuate the incident wave. The EM wave characteristics predicted here are qualitatively coincident with the experimental and analytical results [6,7].

Figures 5 and 6 show the distribution of steady-state temperature and thermal stress at midpoint of the plate versus frequency with various thicknesses. The absorbed electromagnetic energy led to the rise of temperature. The temperature and thermal stress increase as the thickness and frequency increases. For the thin plate, the differences for temperature and stress distributions between FG and homogeneous case are small. There shows obvious difference as the increase of thickness of GM plate.

The analytical results of the transient temperature and thermal stress changes at the midpoint inside the absorber with different incidence power are shown in Fig. 7 and Fig. 8. It can be found that the transient temperature and thermal stress increase with increase in the incidence power, and the heat generated inside the absorber tends to be equal to the heat radiated from the surface of one under irradiation for 150 min.

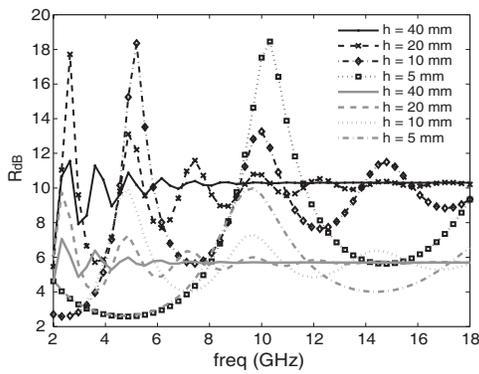


Figure 3: Reflectivity vs. frequency.

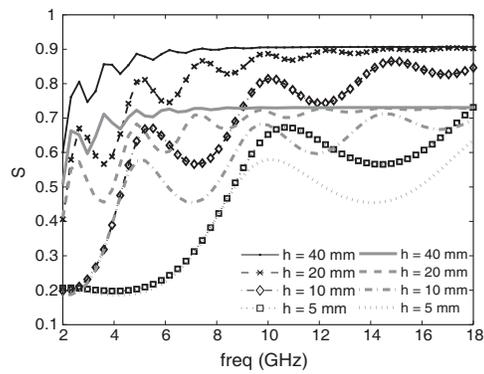


Figure 4: Absorptivity vs. frequency.

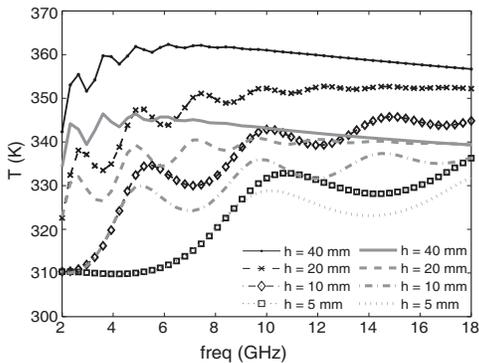


Figure 5: Steady-state temperature vs. frequency.

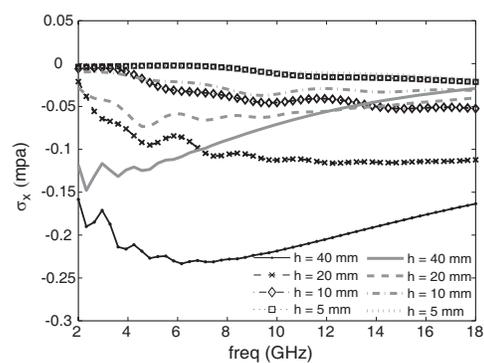


Figure 6: Steady-state thermal stress vs. frequency.

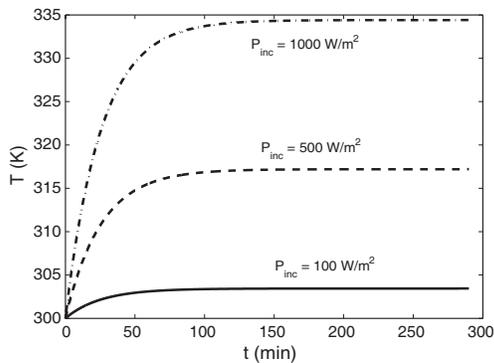


Figure 7: Transient temperature vs. incident power.

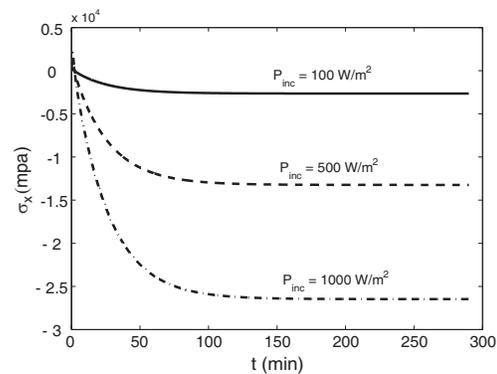


Figure 8: Transient thermal stress vs. incident power.

4. CONCLUSION

In this paper, the physics properties such as absorption, reflection properties and mechanical characteristics like temperature, thermal stress of a continuously graded-infinite absorbing plate in electromagnetic field are carried out. The results show that EM wave absorbing properties of the graded-absorbing structure are improved significantly, and its maximum reflection loss can reach 18.5 dB with thin thickness, at the same time, the temperature and thermal stress are lower and there is no obvious increase of thermal stress of FGM structure. The transient temperature and thermal stress arrive at steady-state eventually.

ACKNOWLEDGMENT

The authors acknowledge supports by the Fok Ying-Tong Education Foundation for Young Teachers in the Higher Education Institutions of China (111005) and the National Natural Science Foundation of China (10872081).

REFERENCES

1. Neo, C. P. and V. K. Varadan, "Optimization of carbon fiber composite for microwave absorber," *IEEE Transactions on Electromagnetic Compatibility*, Vol. 46, No. 1, 102–106, 2004.
2. Meshram, M. R., N. K. Agrawal, et al., "Characterization of M-type barium hexagonal ferrite-based wide band microwave absorber," *Journal of Magnetism and Magnetic Materials*, Vol. 271, No. 2–3, 207–214, 2004.
3. He, Y., R. Gong, et al., "Optimization of two-layer electromagnetic wave absorbers composed of magnetic and dielectric materials in gigahertz frequency band," *J. Appl. Phys.*, Vol. 98, 084903, 2005.
4. Suga, R., O. Hashimoto, et al., "Analytical study on change of temperature and absorption characteristics of a single-layer radiowave absorber under irradiation electric power," *IEEE Transactions on Electromagnetic Compatibility*, Vol. 47, 866–871, 2005.
5. Watanabe, S., A. Taniguchi, et al., "Computer simulation about temperature distribution of an EM-wave absorber using a coupled analysis method," *IEICE Trans. Electron.*, Vol. E91-C, 638–646, 2008.
6. Liu, H. T., M. S. Cao, C. J. Qiu, and J. Zhu, "Design of functionally graded materials towards ram and their microwave reflectivity," *Materials Science Forum*, Vol. 423–425, 427–437, 2003.
7. Sugano, Y. and S. Takahashi, "Material design of functionally graded plates with the function of electromagnetic noise suppression," *Journal of Solid Mechanics and Materials Engineering*, Vol. 2, No. 7, 912–923, 2008.
8. Pendry, J. B., D. Schurig, et al., "Controlling electromagnetic fields," *Science*, Vol. 312, No. 5781, 1780–1782, 2006.
9. Argyropoulos, C., E. Kallos, et al., "Manipulating the loss in electromagnetic cloaks for perfect wave absorption," *Optics Express*, Vol. 17, No. 10, 8467–8475, 2009.

A Model of Size Effect on Thermal Conductivity for Thin Metallic Films

Wei Luo and Xiaojing Zheng

Key Laboratory of Mechanics on Western Disaster and Environment
Ministry of Education, and Department of Mechanics and Engineering Science
College of Civil Engineering and Mechanics, Lanzhou University, Lanzhou, Gansu 730000, China

Abstract— This study examines the size effect on the thermal conductivity due to electrons scattering in metallic films on the basis of the Boltzmann equation in which the modification of the approximation of local thermal equilibrium is accounted for. Three dominant types of electron scattering mechanisms are considered: the isotropic background scattering, the grain boundaries scattering and the external surfaces scattering; furthermore, the grain boundaries scattering consists of two parts: grain boundaries perpendicular to the heat flux and grain boundaries parallel to the heat flux. The in-plane thermal conductivity of metallic thin films is obtained. The comparison between the theoretical calculations and the experimental results indicates that our model is valid and the grain boundary scattering plays an important role in the electron thermal transport.

1. INTRODUCTION

In order to effectively manage heat in integrated circuits and NEMS, designers require heat transport models with accurate thermal property data. Although many bulk materials have well-characterized thermal properties, different mechanisms govern conduction at small lengthscales and to various materials [1].

To metallic thin films, the surfaces and grain boundaries scattering are the dominant processes. The classical models are the FS and MS which are based on the Boltzmann transport equations and the relaxation time approximation [2–4]. But to super-thin films, those models have their deficiencies because of the quantum effects. A new simple model is founded by applying the transport time taking place of the relaxation time [6]. However, According to the related research [7, 8], the grain boundaries scattering effects to thermal conductivity can not simple be analogy to that of electrical conductivity. In the light of these arguments, we develop a model that study the surface and grain boundaries scatterings simultaneously based on the Boltzmann equation.

In this study, three types of the electron scattering mechanism, e.g., isotropic background (intrinsic), grain boundaries, and external surface boundaries scatterings, are considered. In order to thoroughly investigate the grain boundaries scattering, it is divided into two kinds; one is related to the grain boundaries perpendicular to the heat current; the other is related to the grain boundaries parallel to the heat flux (parallel to the film surface) which are neglected by previous scholars [4]. The scattering of grain boundaries parallel to the heat flux and the external surfaces scattering are considered together as the classical Fuch's [2] boundary conditions on the electron-distribution functions by using Hoffman's two-fluid model [11]. And the scattering of grain boundaries perpendicular to the heat current is treated as the potential barriers of suitable widths and heights by solving the Schrödinger's equation, and this effect influences the relaxation time with the isotropic background scattering. Additionally, the approximation of local thermal equilibrium in the quasi-classical electron thermal transport needs to be modified because of the decrease of film thickness.

2. MODEL

The general form of lineal Boltzmann equation with the relaxation time approximation is

$$\frac{\partial f}{\partial t} + \vec{v} \cdot \nabla f + \vec{F} \cdot \frac{\partial f}{\partial \vec{p}} = -\frac{f - f_0}{\tau}, \quad (1)$$

and $f = f_0 + f_1$, where f is the distribution function of the electrons velocity $\vec{v}(v_x, v_y, v_z)$ and the coordinate $\vec{r}(x, y, z)$. \vec{p} is the mobility of electron and τ is the relaxation time with isotropic background scattering. f_0 is the equilibrium distribution function of the kinetic energy ε and is given by the Fermi-Dirac statistics as $f_0 = 2(m/h)^3 / [\exp[(\varepsilon - \mu)/k_B T] + 1]$. μ is the Fermi energy level, which is the chemical potential of the electrons according to Gibbs thermodynamics and is

determined by the electron number density present. T is the temperature and k_B is the Boltzmann constant.

We consider an ideal metallic thin film of thickness d , with perfect external surfaces perpendicular to the z -axis, and take a temperature gradient $\nabla T = \partial T/\partial x$ and a thermal electric field E_x , which owing to the temperature gradient, to be the x direction, as shown in Fig. 1. In general, according to the approximation of local thermal equilibrium, it is assumed that $\nabla f \approx \nabla f_0$, because the thickness of films is further larger than the mean free path of the electrons. But, when the film thickness is of the same order of the mean free path, the approximation of local thermal equilibrium must be modified, So we take $\nabla f = \nabla (f_0 + f_1)$, where ∇f_1 is neglected by previous work [4, 6]. In considering $\frac{\partial f_1}{\partial T} \ll \frac{\partial f_0}{\partial T}$, and thinking that $\frac{\partial f_1}{\partial T}$ is not only the function of \vec{v} , but also is the function of z , and letting $\frac{\partial f_1}{\partial T} = \eta(\vec{v}, z) \frac{\partial f_0}{\partial T}$, where $|\eta(\vec{v}, z)| \ll 1$. Hence, the Boltzmann equation is reduced to

$$v_z \frac{\partial f_1}{\partial z} + \frac{f_1}{\tau^*} = \frac{eE_x}{m} \frac{\partial f_0}{\partial v_x} - v_x [1 + \eta(\vec{v}, z)] \frac{\partial f_0}{\partial T} \frac{dT}{dx}. \quad (2)$$

The general solution is

$$f_1 = \left(\frac{eE_x}{m} \frac{\partial f_0}{\partial v_x} - v_x \frac{\partial f_0}{\partial T} \frac{dT}{dx} \right) \tau^* \left[1 + \psi(\vec{v}, z) \exp\left(\frac{-z}{\tau^* v_z}\right) \right]. \quad (3)$$

With regard of grain boundaries perpendicular to the heat flux, the effective relaxing time τ^* has the form [9] $\frac{1}{\tau^*} = \frac{\lambda}{D} \frac{R(a, U) \cos \varphi \sin \theta}{T(a, U)} + \frac{1}{\tau}$, where $R(a, U)$ and $T(a, U)$ are the coefficients of reflection and transmission from the grain boundary, respectively; and they hold the following relation, $R(a, U) + T(a, U) = 1$. λ is the mean free path without considering grain boundaries and D is the average size of crystalline grains. We introduce a spherical system (v, φ, θ) spatial velocity with $v_z = v \cos \theta$, and a scattering parameter α ($\alpha = (\lambda/D)[R(a, U)/T(a, U)]$). The grain boundary is treated as a potential barrier characterized by the width a and the height U . By solving the Schrödinger's equation for the problem of particle transition through the potential barrier [10], we can obtain the following expressions for α (m is the mass of electrons and \hbar is the Planck constant),

$$\alpha = \frac{\lambda}{D} \frac{U^2}{4\varepsilon(U - \varepsilon)} \sinh^2 \left\{ \frac{a}{\hbar} [2m(U - \varepsilon)]^{1/2} \right\}, \quad \text{for } \varepsilon < U \quad (4a)$$

$$\alpha = \frac{\lambda}{D} \frac{a^2}{2\hbar^2} m\varepsilon, \quad \text{for } \varepsilon = U \quad (4b)$$

$$\alpha = \frac{\lambda}{D} \frac{U^2}{4\varepsilon(\varepsilon - U)} \sin^2 \left\{ \frac{a}{\hbar} [2m(\varepsilon - U)]^{1/2} \right\}, \quad \text{for } \varepsilon > U \quad (4c)$$

Considering the scattering of grain boundaries parallel to the heat flux and the external surfaces scattering, and treating them as the boundary condition by using the Hoffman's two-fluid

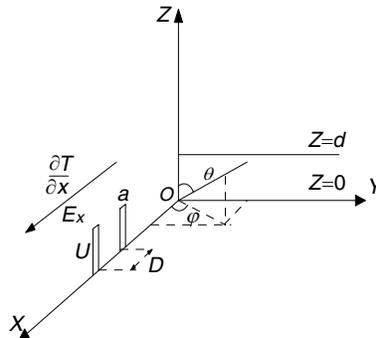


Figure 1: Schematic diagram of the system under consideration.

model [11], we obtain

$$\psi + (\vec{v}, z) = -\frac{1 - pe^{-\zeta z/\lambda}}{1 - pe^{-\zeta z/\lambda} e^{(z-d)/\tau^* v_z}}, \quad (v_z > 0) \quad (5a)$$

$$\psi^- (\vec{v}, z) = -\frac{(1 - pe^{-\zeta z/\lambda}) e^{d/\tau^* v_z}}{1 - pe^{-\zeta z/\lambda} e^{(d-z)/\tau^* v_z}}, \quad (v_z < 0) \quad (5b)$$

and substitute them to Eq. (5), then

$$f_1^+ (\vec{v}, z) = \left(\frac{eE_x}{m} \frac{\partial f_0}{\partial v_x} - v_x \frac{\partial f_0}{\partial T} \frac{dT}{dx} \right) \tau^* \left[1 - \frac{1 - pe^{-\zeta z/\lambda}}{1 - pe^{-\zeta z/\lambda} e^{(z-d)/\tau^* v_z}} e^{-z/\tau^* v_z} \right], \quad (v_z > 0) \quad (6a)$$

$$f_1^- (\vec{v}, z) = \left(\frac{eE_x}{m} \frac{\partial f_0}{\partial v_x} - v_x \frac{\partial f_0}{\partial T} \frac{dT}{dx} \right) \tau^* \left[1 - \frac{(1 - pe^{-\zeta z/\lambda})}{1 - pe^{-\zeta z/\lambda} e^{(d-z)/\tau^* v_z}} e^{(d-z)/\tau^* v_z} \right]. \quad (v_z < 0) \quad (6b)$$

where p is a fraction of the electrons are scattered specularly from the external surfaces and ζ is penetration parameter related to the electrons through grains. The electric current density $j_x(z)$, the average current density \bar{j}_x and the heat flux $q_x(z)$, the average heat flux \bar{q}_x are given by

$$j_x(z) = -e \iiint v_x f d\vec{v}, \quad \bar{j}_x = \frac{1}{d} \int_0^d j(z) dz, \quad (7a)$$

$$q_x(z) = \iiint v_x \varepsilon f d\vec{v}, \quad \bar{q}_x = \frac{1}{d} \int_0^d q_x(z) dz. \quad (7b)$$

We deduce the average current density and set it to zero for obtaining the expression of, E_x and then substitute it in Eq. (7b). Thereby, the film thermal conductivity is $\kappa_{film} = \bar{q}_x / (-dT/dx)$. Combining the definition of bulk thermal conductivity, the ratio of the film thermal conductivity to the bulk value is obtained

$$\frac{k_{film}}{k_{bulk}} = \frac{3}{4\pi} \int_0^{2\pi} d\varphi \int_0^\pi d\theta \int_0^d dz \frac{1}{d} \frac{1}{1 + \alpha \cos \varphi \sin \theta} \cos^2 \varphi \sin^3 \theta \left[1 - \frac{(1 - pe^{-\zeta z/\lambda})}{1 - pe^{-\zeta z/\lambda} e^{(z-d)/\tau^* v |\cos \theta|}} e^{-z/\tau^* v |\cos \theta|} \right]. \quad (8)$$

If don't think over the scattering of grain boundaries perpendicular to the heat flux, we can obtain a approximation,

$$\begin{aligned} \frac{k_{film}}{k_{bulk}} &= \gamma_0 = 1 - \frac{3}{4d} \int_0^\pi d\theta \int_0^d dz \sin^3 \theta e^{-z/\tau v |\cos \theta|} \quad (p = 0) \\ &= \gamma_p = 1 - \frac{3}{4d} \int_0^\pi d\theta \int_0^d dz \sin^3 \theta \frac{(1 - pe^{-\zeta z/\lambda})}{1 - pe^{-\zeta z/\lambda} e^{(z-d)/\tau v |\cos \theta|}} e^{-z/\tau v |\cos \theta|} \quad (p \neq 0) \end{aligned} \quad (9)$$

After defining $d/\lambda = \delta$, and using the so called exponential integral function Ei , γ_0 and γ_p can be written as

$$\gamma_0 = 1 - \frac{3}{8} \delta^{-1} - \frac{3}{4} \left(\delta - \frac{1}{12} \delta^3 \right) Ei(-\delta) - \exp(-\delta) \left[\frac{5}{8} + \frac{1}{16} \delta - \frac{1}{16} \delta^2 - \frac{3}{8} \delta^{-1} \right], \quad (p = 0) \quad (10)$$

$$\begin{aligned} \gamma_p &\simeq 1 - \frac{3}{16} (1 - pe^{-\zeta \delta}) e^{-\delta} + \frac{3}{16} \frac{1}{\delta} e^{-\delta} (1 - pe^{-\zeta \delta}) \ln (1 - pe^{-\zeta \delta}) - \frac{3}{16} \frac{1}{\delta} (1 - p) \ln (1 - pe^{-\delta}) \\ &\quad (1 - pe^{-\zeta \delta}) \left[-\frac{15}{16} \frac{1}{\delta} e^{-\delta} + \frac{3}{8} \delta \left(1 - \frac{1}{12} \delta^2 \right) Ei(-\delta) + \frac{1}{32} \delta (1 - \delta) e^{-\delta} \right] \quad (p \neq 0). \end{aligned} \quad (11)$$

When $\delta > 1$, Eq. (11) is reduced to $\gamma_p \simeq 1 - \left(\frac{3}{16} + \frac{15}{16} \frac{1}{\delta} \right) (1 - pe^{-\zeta \delta}) e^{-\delta}$. (*)

3. RESULTS AND DISCUSSIONS

The ratio of film to bulk thermal conductivity is a function of δ at different α according to Eq. (8) as shown in Fig. 2. It can be found that the value of k_f/k_b decreases apparently with the parameter α increasing. Because the parameter α is the function of the average grain size D , and we can conclude that the average size is smaller, and the ratio of film to bulk thermal is more litter.

Figure 3 presents that the comparison between the theoretical predictions from the expression (*) and the experimental results from previous work [5]. The theoretical calculations agree with the experimental results. It indicates that the simplification of the expression is reasonable. From more calculations, we find that the expression is else valid for super-thin films when $\delta < 1$ and $\delta > 0.1$. So the simple expression is suitable for application.

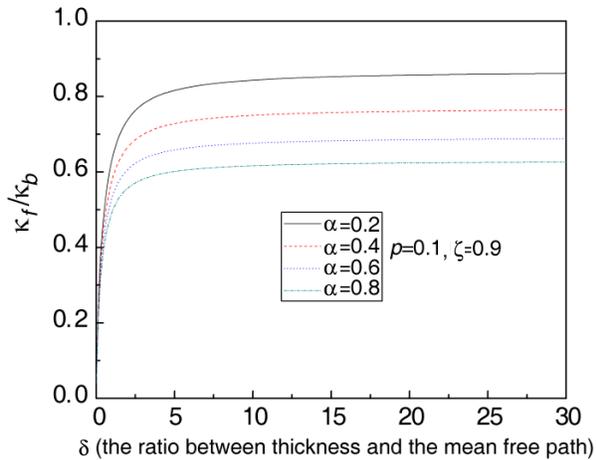


Figure 2: k_f/k_b as a function of δ at $p = 0.1$, $\zeta = 0.9$, but different α .

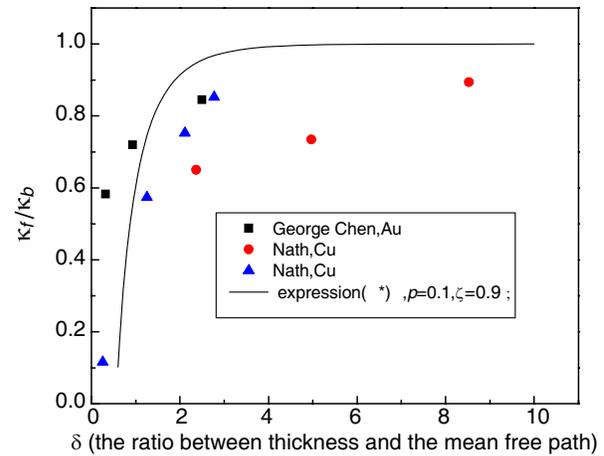


Figure 3: k_f/k_b as a function of δ . The line is the results that calculates according to the expression (*), and symbols are the experimental results [5].

4. CONCLUSIONS

In conclusion, we propose a theoretical model to describe the size effect in metallic thin films which are resulted from the external surfaces scattering and the grain boundaries scattering. The calculations indicate that: when the thickness of films is less than the mean free path of electrons, the grain boundaries scattering is as important as the external surfaces scattering; when the thickness is of the order of the mean free path, the external surfaces scattering is the main factor. The present results will be helpful to study the thermal properties of nano-materials and design new devices in the future.

ACKNOWLEDGMENT

Authors acknowledge support from the NSFC's program (No. 90405005), the "973 Program" (2007CB607506), the Ph.D. Fund (No. 20050730016), and the fund of No. WUT2005Z04.

REFERENCES

1. Tien, C. L. and G. Chen, "Challenges in microscale conductive and radiative heat transfer," *ASME J. Heat Trans.*, Vol. 116, No. 9, 799–807, 1994.
2. Sondheimer, E. H., "The mean free path of electrons in metals," *Adv. Phys.*, Vol. 50, No. 6, 499–537, 2001.
3. Mayadas, A. F. and M. Shatzkes, "Electrical-resistivity model for polycrystalline films: The case of arbitrary reflection at external surfaces," *Phys. Rev. B*, Vol. 1, No. 4, 1382–1389, 1970.
4. Kumar, S. and G. C. Vradis, "Thermal conductivity of thin metallic films," *ASME J. Heat Trans.*, Vol. 116, No. 2, 28–34, 1994.
5. Nath, P. and K. L. Chopra, "Thermal conductivity of copper films," *Thin Solid Films*, Vol. 20, No. 1, 53–62, 1974.
6. Jin, J. S., J. S. Lee, et al., "Electron effective mean free path and thermal conductivity predictions of metallic thin films," *Appl. Phys. Lett.*, Vol. 92, No. 17, 1719101–1719103, 2008.

7. Zhang, Q. G. and B. Y. Cao, “Influence of grain boundary scattering on the electrical and thermal conductivities of poly-crystalline gold nanofilms,” *Phys. Rev. B*, Vol. 74, No. 13, 1341091–1341095, 2006.
8. Zhang, X., H. Q. Xie, et al., “Thermal and electrical conductivity of a suspended platinum nanofilm,” Vol. 86, No. 17, 1719121–1719123, 2005.
9. Warkusz, F. J., “Electron scattering at the grain boundaries and total conductivity for metal films,” *Mater. Sci. Lett.*, Vol. 6, No. 10, 1123–1126, 1987.
10. Zeng, J. Y., *Quantum Mechanics Tutorial*, Science Press, Beijing, 2002.
11. Fang, P., K. Yi, et al., “Electrical transport in metallic films,” *J. Appl. Phys.*, Vol. 95, No. 5, 2527–2531, 2004.

Dynamic Analysis for Electrified Cantilever Conductive Thin Plates under Transverse Multi-pulse Magnetic Field

Huijuan Bai and Xiaojing Zheng

Department of Mechanics and Engineering Science, Lanzhou University
Lanzhou, Gansu 730000, China

Abstract— A mathematic model was established for electrified cantilever conductive thin plates under external transverse multi-pulse magnetic field. The distribution of eddy current induced by the transverse multi-pulse magnetic field would be influenced due to a subsistent inner uniform electric field in the conductive thin plates by adopting finite element numerical method, and also obtained the distribution of temperature field in thin plate from the law of heat conduction. Subsequently the influences of multi-pulse magnetic field and uniform electric field on the in-plane magnetic volume forces and the maximum deformation of thin plate were quantitatively simulated on the basis of electro-magneto-thermo-elastic theory. The simulation results indicate that the dynamic buckling phenomenon is caused by the in-plane magnetic volume compression force arising from the interaction between the eddy current induced by multi-pulse magnetic field and inner uniform electric field. The dynamic buckling phenomenon of cantilever thin plate is determined by the value of maximum magnetic field and impulse parameter, and also the inner uniform electric field.

1. INTRODUCTION

Abundant conductive shells and structures are broadly applied to the complicated electromagnetic circumstance in modern industry and electronic devices, so developing a relevant electro-magneto-thermo-elastic theory is important to make the guidance on the metallic thermo forming, optimum design of configuration for the electronic mechanical devices run under the electromagnetic field.

Much attention had been paid to the mechanical behaviors of a conductive thin plate under the magnetic field changing with time. Hua et al. [1, 2] took the limiter blade of a Tokamak fusion reactor as a cantilever beam in applied magnetic fields delayed exponentially with time. Zhou et al. [3] developed a coupled magneto-mechanical model, which can describe various mechanical behaviors for the ferromagnetic structures. Takagi et al. [4] conducted an experimental of thin plate deflection in the magnetic field and gave some computational analyses by using the T -method [5]. Recently, Wang et al. [6, 7] presented a theory model to analyze the magneto-thermo-elastic instability of ferromagnetic plates subjected to thermal and magnetic loading. Zhang et al. [8] and Zhu et al. [9] revealed some rules for the dynamic stability of a cantilever conductive plate under transverse impulsive magnetic field and strong magnetic field. Higuchi et al. [10] studied the magneto-thermo-elastic stressed of a conductive solid circular cylinder in a transient magnetic field. Laissaoui et al. [11] estimated the thermal impact on induction machines, for which a coupled electromagnetic thermal analysis was carried out.

Most researchers are focus on the dynamic behaviors of rectangular conductive beams or plate structures which are in transverse monopulse magnetic field, and fewer considered the existence of inner electric field. The present study considers the electrified cantilever conductive thin plate subjected to external transverse multi-pulse magnetic field. By using finite element numerical method, we obtain the distribution of eddy current induced by the transverse multi-pulse magnetic field and inner uniform electric field. Subsequently, the distribution of temperature, the in-plane magnetic volume forces and the maximum deformation of thin plate were quantitatively gained on the basis of electro-magneto-thermo-elastic theory. We also discuss the main influencing factors on the dynamic stability of the conductive plate.

2. BASIC EQUATION

Considering that a cantilever conductive thin plate is applied by the transverse multi-pulse magnetic field $B_0 = B_0(t)\mathbf{k}$ and the inner electric field $\mathbf{E} = E_0\mathbf{i}$ as shown in Fig. 1, where \mathbf{i} and \mathbf{k} are the unit vectors in the direction of x and z respectively. Ignore the displacement current, and Maxwell

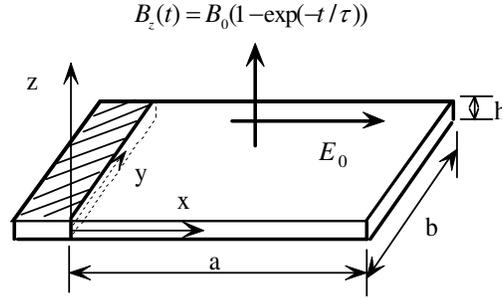


Figure 1: Schematic diagram of the electrified cantilever conductive thin plate in transverse multi-pulse magnetic field.

electromagnetic equations related to the distribution of eddy current can be described as

$$\nabla \times \mathbf{H} = \mathbf{J}_e, \quad (1a)$$

$$\nabla \times \mathbf{E}_e = - \left(\frac{\partial \mathbf{B}_e}{\partial t} + \frac{\partial B_0}{\partial t} \right). \quad (1b)$$

where \mathbf{H} , \mathbf{J}_e and \mathbf{E}_e are the vector of magnetic field, eddy current density and electric field intensity induced by time-varying magnetic field respectively. B_0 and \mathbf{B}_e are the applied magnetic field and the induced magnetic field duo to the eddy current.

The constitution relations for the magnetic field and the eddy current can be respectively expressed as

$$\mathbf{B} = B_0 + \mathbf{B}_e = \mu_0 \mathbf{H}, \quad \mathbf{J}_e = \sigma(\mathbf{E}_e + \mathbf{V} \times \mathbf{B}). \quad (2)$$

where \mathbf{V} is the velocity of vibration of the plate, σ , μ_0 are respectively the electrical conductivity of the conductive plate and the magnetic permeability of vacuum.

Applying divergence operator to both sides of Eq. (1a), we can get $\nabla \cdot \mathbf{J}_e = 0$, and the eddy current vector potential \mathbf{T} [5] can be defined as $\mathbf{J}_e = \nabla \times \mathbf{T}$. Introducing Coulombian gauge condition $\nabla \cdot \mathbf{T} = 0$, and using the Helmholtz's formula and Biot-Savart's law for the thin plate, the magnetic flux density arise from the eddy current in the plate can be written as

$$B_e = \mu_0 T - \frac{\mu_0}{4\pi} \int_s T'_n \frac{\partial}{\partial z} \left(\frac{1}{R} \right) ds'. \quad (3)$$

here $\mathbf{T} = T\mathbf{k}$, $\mathbf{B}_e = B_e\mathbf{k}$. T'_n is the component of T along the thin plate surface s , and R is a distance from an arbitrary point in the plate to a point occupied by the eddy current. Substituting Eq. (3) into Eq. (1b), and considering Eq. (2), we can obtain the basic equation of the vector potential of the eddy current T as

$$\nabla^2 T - \mu_0 \sigma \frac{\partial T}{\partial t} - \frac{\sigma \mu_0}{4\pi} \int_s \frac{\partial T'_n}{\partial t} \frac{\partial}{\partial z} \left(\frac{1}{R} \right) ds' = \sigma \frac{\partial B_0(t)}{\partial t}. \quad (4)$$

with the boundary condition $\partial T / \partial s = 0$ and the initial condition $T(x, y, 0) = 0$.

In the Eqs. (3) and (4), the eddy current vector potential \mathbf{T} for the thin plate is simplified as the only component T in the direction of z , which makes the magnetic flux density \mathbf{B}_e induced by the eddy current has the component in the direction of z .

Since an inner uniform electric field \mathbf{E} is existence in the thin plate, the current \mathbf{J} can be expressed as,

$$\mathbf{J} = \mathbf{J}_e + \sigma \mathbf{E} = \sigma(\mathbf{E}_e + \mathbf{E} + \mathbf{V} \times \mathbf{B}). \quad (5)$$

Assumed that the current is uniform along the thickness of the thin plate, the electromagnetic force density vector \mathbf{F} is given by Lorentz force formula,

$$\mathbf{F} = \int_{-h/2}^{h/2} (\mathbf{J} \times \mathbf{B}) dz = -h(B_0 + B_e) \frac{\partial T}{\partial x} \mathbf{i} - h(B_0 + B_e) \left(\frac{\partial T}{\partial y} + \sigma E_0 \right) \mathbf{j} \quad (6)$$

the Eq. (6) reveals that electromagnetic force exerted on the plate only has two in-plane component F_x and F_y with applied transverse time-varying magnetic field. If these in-plane components

were ignored, the thin plate would be not subjected to electromagnetic forces so that it would be impossible to analyze the mechanical behavior.

The plane temperature field induced by the heating effect of eddy current for the conductive plate can be expressed as

$$\lambda \left(\frac{\partial^2 \bar{T}}{\partial x^2} + \frac{\partial^2 \bar{T}}{\partial y^2} \right) = - \left(\frac{1}{\sigma} J^2 - c\rho \frac{\partial \bar{T}}{\partial t} \right) \quad (7)$$

where λ , c and ρ are respectively the thermal conductivity, specific heat ratio and mass density, \bar{T} is the temperature of one point in the plate. The convective boundary condition is $\lambda \frac{\partial \bar{T}}{\partial n} = -\beta(\bar{T}_{sc} - \bar{T}_e)$ and β is the convection coefficient.

Based on the analyses of the electromagnetic force mentioned above, the in-plane displacement of the plate under the in-plane electromagnetic force F_x and F_y and thermal stress is considered. The movement equations for the displaces u and v can be expressed as

$$\frac{Eh}{1-\mu^2} \left(\frac{\partial^2 u}{\partial x^2} + \frac{1-\mu}{2} \frac{\partial^2 u}{\partial y^2} + \frac{1+\mu}{2} \frac{\partial^2 v}{\partial x \partial y} \right) + F_x(x, y, t) - \frac{Eh\bar{\alpha}}{1-\mu} \frac{\partial \bar{\theta}}{\partial x} = \rho h \frac{\partial^2 u}{\partial t^2}, \quad (8)$$

$$\frac{Eh}{1-\mu^2} \left(\frac{\partial^2 v}{\partial y^2} + \frac{1-\mu}{2} \frac{\partial^2 v}{\partial x^2} + \frac{1+\mu}{2} \frac{\partial^2 u}{\partial x \partial y} \right) + F_y(x, y, t) - \frac{Eh\bar{\alpha}}{1-\mu} \frac{\partial \bar{\theta}}{\partial y} = \rho h \frac{\partial^2 v}{\partial t^2}. \quad (9)$$

The boundary conditions and the initial conditions are respectively written as

$$x = 0 : v = 0; \quad x = a, y = 0, b : N|_{\Gamma} = 0, \quad (10)$$

$$u(x, y, 0) = v(x, y, 0) = 0, \quad \partial u(x, y, 0)/\partial t = \partial v(x, y, 0)/\partial t = 0. \quad (11)$$

where E , μ , $\bar{\alpha}$ and $\bar{\theta}$ are the Young's modulus, the Poisson's ratio, bending strength and the temperature difference between the instantaneous temperature of one time and the next time. Γ represents the boundary of the plate except for the edge $x = 0$; $N|_{\Gamma}$ is the component of internal membrane force N at the boundary Γ . Since the internal membrane force N is related to the displaces u and v of the plate, its all components can be written

$$N_x = \frac{Eh}{1-\mu^2} \left(\frac{\partial u}{\partial x} + \mu \frac{\partial v}{\partial y} \right) - \frac{E\bar{\alpha}}{1-\mu} \int_{-h/2}^{h/2} \bar{\theta} dz, \quad N_y = \frac{Eh}{1-\mu^2} \left(\frac{\partial v}{\partial y} + \mu \frac{\partial u}{\partial x} \right) - \frac{E\bar{\alpha}}{1-\mu} \int_{-h/2}^{h/2} \bar{\theta} dz, \quad (12)$$

$$N_{xy} = \frac{Eh}{2(1+\mu)} \left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \right).$$

The vibration equation of the plate under the internal membrane forces N_x , N_y and N_{xy} can be written as

$$-D\nabla^2 \nabla^2 w - \frac{E\bar{\alpha}}{1-\mu} \int_{-h/2}^{h/2} \nabla^2 \bar{\theta} z dz + \left(N_x \frac{\partial^2 w}{\partial x^2} + 2N_{xy} \frac{\partial^2 w}{\partial x \partial y} + N_y \frac{\partial^2 w}{\partial y^2} \right) + \left(\frac{\partial N_x}{\partial x} + \frac{\partial N_{xy}}{\partial y} \right) \frac{\partial w}{\partial x} + \left(\frac{\partial N_y}{\partial y} + \frac{\partial N_{xy}}{\partial x} \right) \frac{\partial w}{\partial y} = \rho h \frac{\partial^2 w}{\partial t^2} \quad (13)$$

Considering the cantilever plate as shown in Fig. 1, the boundary conditions and the initial conditions are given

$$x = 0 : w(x, y, t) = 0, \quad \partial w(x, y, t)/\partial t = 0; \quad x = a, y = 0, b : M_n|_{\Gamma} = 0, \quad V_n|_{\Gamma} = 0, \quad (14)$$

$$t = 0 : \partial w(x, y, t)/\partial t = 0, \quad w(x, y, 0) = 0. \quad (15)$$

where $M_n|_{\Gamma}$ and $V_n|_{\Gamma}$ are a bending moment and an equivalent shearing force at the boundary Γ of the plate.

In the process of simulation calculation for the dynamic response of the electrified conductive thin plate under the multi-pulse magnetic field, we need to solve Eq. (4) first to obtain the eddy current vector potential T , and then get the total current density \mathbf{J} and the distribution of temperature. After having got the membrane forces N_x , N_y and N_{xy} arising from in-plane forces $F_x(x, y, t)$ and $F_y(x, y, t)$ by solving Eqs. (8)–(11), we will predict dynamic response of the plate according to the solution of Eqs. (13)–(15).

3. NUMERICAL RESULT

In order to simulate the electro-magneto-thermo-mechanical multi-field coupling behaviors of the cantilever conductive thin plate under the inner uniform electric field and external transverse multi-pulse magnetic field, as shown in Fig. 1, the theoretical model developed in preceding section is preformed. The parameters of geometry and material of the plate are taken as $a = 0.2$ m, $b = 0.1$ m, $h = 4$ mm, $E = 6.89 \times 10^{10}$ Pa, $\mu = 0.3$, $\rho = 2.713 \times 10^3$ kg/m³, $\sigma = 3.65 \times 10^7$ S/m, $\mu_0 = 4\pi \times 10^{-7}$ H/m, $\lambda = 263$ W/mK, $\bar{\beta} = 5$ W/m²K, $c\rho = 2.44 \times 10^6$ J/m³K. The applied time-varying magnetic field is $B_0 = B_0[1 - \exp(-t/\tau)]k$, and adopting the critical value $B_0 = 2$ Tesla, $\tau = 0.06$ s. The maximum number of impulse adopted here is $N_{\max} = 5000$. We propose a numerical technique of the finite element method using the *Crank-Nicolson* method and *Newmark* method in time for simulation, taking the element number of the plate as 20×10 and the time step as $\Delta t = 0.01$ s.

Choose one point of $(a/2, b/2)$ in the thin plate and its deflection response are plotted respectively for $E_0 = 0$ V/m and $E_0 = 0.3$ V/m in Fig. 2. It is found from the Fig. 2(a) that the vibration of the plate with a small amplitude tends to stability when no inner added electric field. However, with the uniform inner electric field in the Fig. 2(b), the amplitude increases greatly and the vibration will last long. Also, the center of the eddy current is moved and the distribution of the eddy current density in the conductive plate is changed owing to the added uniform electric field, as shown in Fig. 3.

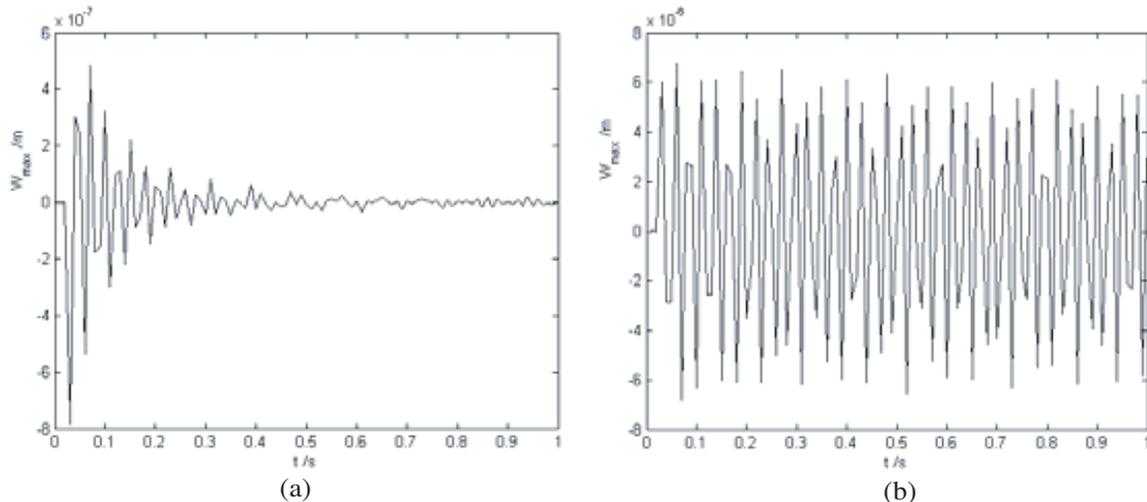


Figure 2: Response of the vibration of the conductive plate ((a) no electric field; (b) inner electric field $E_0 = 0.3$ V/m).

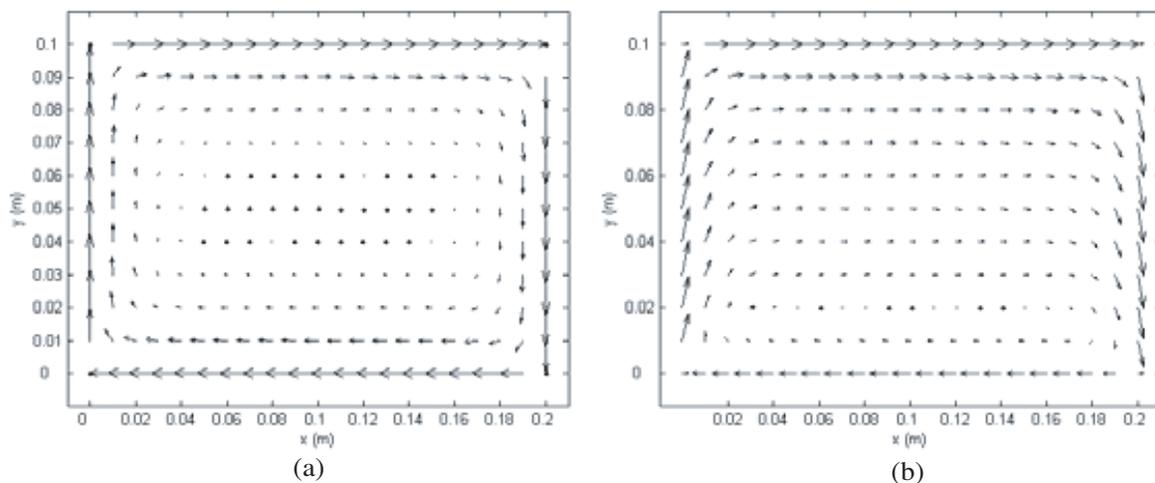


Figure 3: The distribution of the eddy current density ((a) no inner electric field; (b) inner electric field $E_0 = 0.1$ V/m).

4. CONCLUSIONS

A numerical analysis for the dynamic stability of the conductive thin plate subjected to the inner uniform electric field and transverse multi-pulse magnetic field is displayed. When the inner electric field is applied, the vibration of the plate becomes instable and the amplitude increases greatly. The changes of the distribution of eddy current density certainly alter the in-plane components of the electromagnetic force and the thermal stress, which can bring the influence on the distortion of the conductive thin plate.

ACKNOWLEDGMENT

This research was supported by the Fund of HSFC's program (No. 90405005), the Fund of "973 program" (2007CB607506), the Ph.D. Fund (No. 20050730016), and the Fund of No. WUT2005Z04.

REFERENCES

1. Hua, T. Q., M. J. Knott, and L. R. Turner, "Experimental modeling of eddy currents and deflection for Tokamak limiters," *Fusion Tech.*, Vol. 10, 1047–1052, 1986.
2. Turner, L. R. and T. Q. Hua, "Experimental study of coupling between eddy currents and deflections in cantilevered beams as models of Tokamak limiters," *Electromagnetomechanical Internationals in Deformable Solids and Structures*, 81–86, North Holland, Amsterdam, 1987.
3. Zhou, Y. H. and X. J. Zheng, "A general expression of magnetic force for soft ferromagnetic plates in complex magnetic field," *Int. J. Eng. Sci.*, Vol. 35, No. 15, 1405–1417, 1997.
4. Takagi, T., J. Tani, S. Matsuda, and S. Kawamura, "Analysis and experiment of dynamic deflection of a thin plate with a coupling effect," *IEEE Trans. Magn.*, Vol. 28, No. 2, 1259–1262, 1992.
5. Miya, K., J. E. Akin, and S. Hanai, "Finite element analysis of an eddy current induced in thin structure of magnetic fusion reactor," *Int. J. Numer. Mech. Eng.*, Vol. 17, 1613–1629, 1981.
6. Wang, X. Z., Y. H. Zhou, and X. J. Zheng, "A generalized variational model of magneto-thermoelasticity for nonlinearly magnetized ferroelastic bodies," *Int. J. Eng. Sci.*, Vol. 40, No. 17, 1957–1973, 2002.
7. Wang, X. Z., Y. H. Zhou, and X. J. Zheng, "Magneto-thermo-elastic instability of ferromagnetic plates in thermal and magnetic fields," *Int. J. Solids Struct.*, Vol. 40, 6125–6142, 2003.
8. Zheng, X. J., J. P. Zhang, and Y. H. Zhou, "Dynamic stability of a cantilever conductive plate in transverse impulsive magnetic field," *Int. J. Solids Struct.*, Vol. 42, 2417–2430, 2005.
9. Zhu, L. L., J. P. Zhang, and X. J. Zheng, "Multi-field coupling behavior of simply-supported conductive plate under condition of a transverse strong impulsive magnetic field," *Acta Mech. Solida Sin.*, Vol. 19, No. 3, 203–211, 2006.
10. Higuchi, M., M. Kawamura, and Y. Tanigawa, "Magneto-thermo-elastic stresses induced by a transient magnetic field in a conducting solid circular cylinder," *Int. J. Solids Struct.*, Vol. 44, 5316–5335, 2007.
11. Laissaoui, S., M. R. Mekideche, D. Sedira, and A. Ladjimi, "Dynamic modeling of induction motor taking into account thermal stresses," *COMPEL*, Vol. 26, No. 1, 36–47, 2007.

Crack Problem in a Thin Superconducting Disk

Feng Xue^{1,2} and You-He Zhou^{1,2}

¹Key Laboratory of Mechanics on Disaster and Environment in Western China, Ministry of Education
Lanzhou University, Lanzhou 730000, China

²Department of Mechanics and Engineering Science, College of Civil Engineering and Mechanics
Lanzhou University, Lanzhou 730000, China

Abstract— Bulk high-temperature superconductors have now been developed for applications in magnetic bearings, current leads, and flywheel energy storage systems. However, there is a major problem of such bulks, that is mechanical fracturing frequently occurs when a large magnetic field is applying. In this work, the general problem of a central crack in a thin superconducting disk is studied. The dependence of the stress intensity factor on the parameters, including the crack length and the applied field, is investigated. We calculate the body forces by numerical method, solve the magneto-elastic problem and present a simple model in which the effect of the crack on the critical current is taken into account. It is assumed that the crack forms a perfect barrier to the flow of current. The Bean model is considered for the critical state. Based on the complex potential and the boundary collocation methods, the stress intensity factor under the various magnetic field is obtained for a thin superconducting disk containing a central crack. The results show that the crack length and the applied field have significant effects on the fracture behavior of the superconductor. The stress intensity factor will arrive a peak when the magnetic field reduces, there is also a peak for the stress intensity factor as the relative length of the crack is reduced. However, the variation in the stress intensity factor with respect to the relative length of the crack is relatively small. Thus, the results show the times when mechanical fracturing is easy to happen.

1. INTRODUCTION

Top-seeded growth of bulk Y-Ba-Cu-O magnet has developed towards the production of large single-grain superconductors with very strong flux pinning during recent years [1, 2]. Superconductors of this type can trap very high magnetic fields in the remanent state, exceeding the field from permanent magnets by an order of magnitude and even more. However, the mechanical properties and thermal conductivity of a bulk $\text{YBa}_2\text{Cu}_3\text{O}_{7-y}$ magnet can strongly affect the magnetic behavior in a bulk Y-Ba-Cu-O magnet [3]. Thus, the magneto-elastic problems have received considerable attention in the past years [4]. A systematic study of these problems was first made by Ren et al. [5], later the modeling of the magneto-elastic behaviour was extended by including various geometrical cases [6–11].

Microcracking is a common and important phenomenon in melt-processed rare earth-Ba-Cu-O superconductors [12]. Cracks can be formed during oxygenation in the single grain bulk superconductor prepared by the top seeded melt-growth process [13]. Theoretical and experimental studies have been made recently [14–16]. However, the crack problem in a thin superconducting disk has not yet been studied.

In this paper, we use the complex potential method used in [16] to calculate the stress intensity factor of a thin superconducting disk with a central crack, which based on the critical state Bean model used in [17]. We present the results here and discussed how the stress intensity factor depends on the crack length and the applied field.

2. BASIC EQUATIONS

Consider a thin superconducting disk of radius R and thickness d , where $d \ll R$, see Fig. 1. The length of the crack is $2a$. We present a simple model on the assumption that there are no currents flow inside the crack region $0 < r < a$. While outside the crack region $a < r < R$, the current distribution is not disturbed by the crack. The flux distribution can be determined at the critical state. In the critical state model, the currents flow in different circular loops, and the Lorentz force f is generated on the flux lines.

For the convenience of mathematical analysis, we introduce a resultant force σ_t which replaces the body force. In this case, the resultant force is just applied at the boundary of the disk. By

integrating the body force along the radius r , the resultant force σ_t is given by

$$\sigma_t = \int_a^R f(r)rdr/R \tag{1}$$

where $f(r)$ is the body force and is defined as

$$f(r) = -\frac{1}{2\mu_0} \frac{d}{dr}(B^2) \tag{2}$$

In order to calculate the stress intensity factor, we use a complex potential method [16, 18]. The conformal mapping method is used to solve this problem. We use a function mapping a crack in the z -plane to a unit circle in the ζ -plane, which is [16]

$$z = \omega(\zeta) = (a/2)/(\zeta + \zeta^{-1}) \tag{3}$$

$$\zeta = z/a + [(z/a)^2 - 1]^{\frac{1}{2}} \tag{4}$$

where

$$z = x + iy \tag{5}$$

The crack and its exterior are mapped to the unit circle and its exterior by (3) (see Fig. 3).

The stresses and boundary condition in the elastic response yield

$$\sigma_x + \sigma_y = 4\text{Re}\Phi(\zeta) = 4\text{Re}[\phi'(\zeta)/\omega'(\zeta)] \tag{6}$$

$$\sigma_y - \sigma_x + 2i\tau_{xy} = 2[\overline{\omega(\zeta)}\Phi'(\zeta)/\omega'(\zeta) + \Psi(\zeta)] = 2\{\overline{\omega(\zeta)}[\phi'(\zeta)/\omega'(\zeta)]' + \psi'(\zeta)\}/\omega'(\zeta) \tag{7}$$

$$\phi(\zeta) + \omega(\zeta)\overline{\phi'(\zeta)}/\omega'(\zeta) + \overline{\psi(\zeta)} = i \int_s (X_n + iY_n)ds \tag{8}$$

where X_n and Y_n are the horizontal and vertical forces applied on an element of arc ds . With a same process as [16], the stress intensity factor K_I can be defined and calculated as

$$K_I = 2\pi^{1/2}\phi'(1)/[\omega''(1)]^{1/2} = 2\pi^{1/2}\phi'(1)/a^{1/2} \tag{9}$$

As only the tensile stress can lead to the fracture of the disk, we restrict the analysis to the decreasing field process. The applied field is decreased from its maximum value which is denoted by B_m .

The flux density profiles in a disk have been derived by Jonhansen in [17]. When the applied field is reduced, the flux density should be studied in three different regions, that is:

Region I (outer); $b \leq r < R$;

$$\frac{B(r)}{B_d} = \frac{B_a}{B_d} + \arccos h(R/r) + \frac{2}{\pi} \left[\int_{\arcsin(\hat{a}/r)}^{\arcsin(b/r)} \frac{1 - \theta \cot \theta}{\sqrt{1 - (r/R)^2 \sin^2 \theta}} d\theta - \int_{\arcsin(b/r)}^{\pi/2} \frac{1 - \theta \cot \theta}{\sqrt{1 - (r/R)^2 \sin^2 \theta}} d\theta \right] \tag{10}$$

$$J(r)/J_c = 1$$

where $b = \frac{R}{\cosh[(B_m - B_a)/2B_d]}$, $B_d = \mu_0 J_c d/2$, and $\hat{a} = \frac{R}{\cosh(B_m/B_d)}$.

Region II (middle); $\hat{a} \leq r < b$;

$$\frac{B(r)}{B_d} = \frac{B_m}{B_d} - \arccos h(R/r) + \frac{2}{\pi} \int_{\arcsin(\hat{a}/r)}^{\pi/2} \frac{1 - \theta \cot \theta}{\sqrt{1 - (r/R)^2 \sin^2 \theta}} d\theta \tag{11}$$

$$\frac{J(r)}{J_c} = -1 + \frac{4}{\pi} \arctan \left(\frac{r}{R} \sqrt{\frac{R^2 - b^2}{b^2 - r^2}} \right)$$

Region III (inner); $r < \hat{a}$;

$$B(r) = 0 \quad \frac{J(r)}{J_c} = \frac{2}{\pi} \left[2 \arctan \left(\frac{r}{R} \sqrt{\frac{R^2 - b^2}{b^2 - r^2}} \right) - \arctan \left(\frac{r}{R} \sqrt{\frac{R^2 - \hat{a}^2}{\hat{a}^2 - r^2}} \right) \right] \tag{12}$$

However, we consider that the crack length a is much longer than \hat{a} , so we replace \hat{a} by a in region II and region III is then displaced for that there is no current flow within this region.

3. RESULTS AND DISCUSSTION

In this section, we calculate the stress intensity factor of a superconducting disk with a central crack subjected to the electromagnetic forces. Because of the symmetry in geometry and loading, the mode II stress intensity factor K_{II} is equal to zero and only mode I stress intensity factor K_I will be considered. All the stress intensity factors here are normalized by

$$K_0 = \sigma_0 \sqrt{\pi a}$$

in which $\sigma_0 = B_d^2/2\mu_0$.

We can find that there are two factors that affect the value of the intensity factor K_I . One is that the applied field which determines the value of the total force σ_t , and the other one is that the length of the crack.

Figure 3 shows the distribution of the stress intensity factor of the disk when the applied field is reduced from B_m to B_a based on the Bean model.

In Fig. 3(a), the stress intensity factor firstly increase with the increase of the relative crack length a/R , then decreases after it reaches a peak value. This is because that, in this case, the effect of the crack length is larger than the other one when the crack is small. However, the effect will become insignificant as the crack length increases. In Fig. 3(b), however, the stress intensity factor reaches its peak value more quickly than the case in Fig. 3(a). So in most time the total force σ_t has dominant effect on the stress intensity factor. Another different from the previous case is that the magnitude of the stress intensity factor reduces relative to the case in Fig. 3(a). It is indicated that a larger stress intensity factor can be obtained due to a stronger applied field. Therefore, a stronger field should be avoid in energizing a trapped-field magnet.

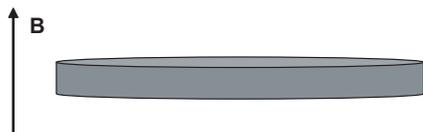


Figure 1: A thin superconducting disk with a central crack under a parallel magnetic field B_a .

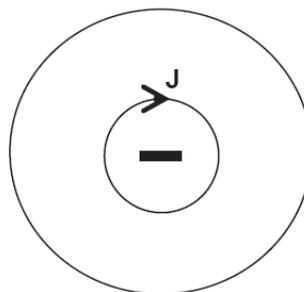


Figure 2: Flowing of shielding current in the circular loops.

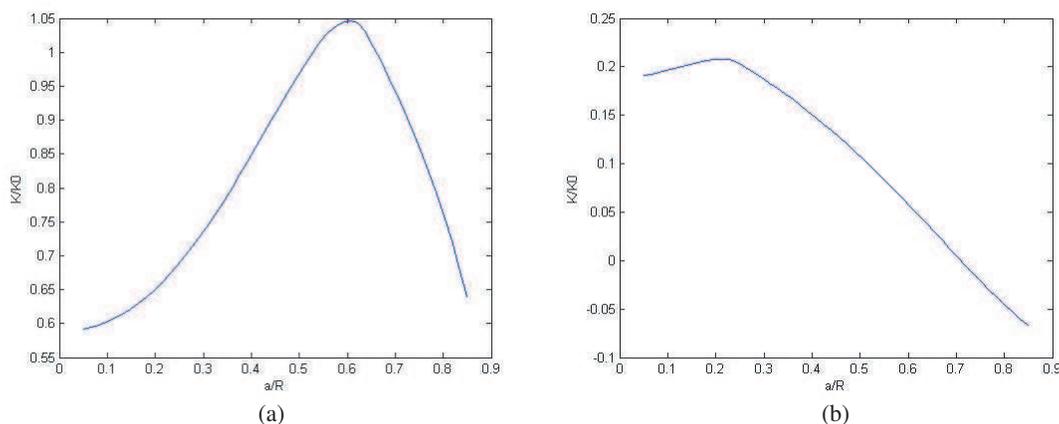


Figure 3: Variation in the stress intensity factor for the Bean model as the field decreases from B_m to B_a .

4. CONCLUSION

In conclusion, this article presents the fracture problem caused by flux pinning in a thin superconducting disk. Based on the complex theory and the assumption that there is no current flow in the crack region, results for the Bean model are obtained. It is worth pointing that the variation in the stress intensity factor with respect to the crack length is relatively small. It is indicated that the effect of the crack on the critical current distribution is not negligible, the perturbation of the critical current brought upon by the crack is significant and cannot be neglected.

ACKNOWLEDGMENT

This research was supported by a grant of the Fund of Natural Science Foundation of China, the Fund of Ministry of Education of the Program of Changjiang Scholars and Innovative Research Team in University (No. IRT0628), and the Science Foundation of the Ministry of Education of China for the Ph.D. program. The authors gratefully acknowledge these supports.

REFERENCES

1. Murakami, M., N. Sakai, T. Higuchi, and S. I. Yoo, *Supercond. Sci. Technol.*, Vol. 9, 1015, 1996.
2. Muralidhar, M., M. Jirsa, N. Sakai, and M. Murakami, *Supercond. Sci. Technol.*, Vol. 16, R1, 2003.
3. Tomita, M. and M. Murakami, *Nature*, Vol. 421, 517, London, 2003.
4. Johansen, T. H., *Supercond. Sci. Technol.*, Vol. 13, R121, 2000.
5. Ren, Y., R. Weinstein, J. Liu, R. P. Sawh, and C. Foster, *Physica C*, Vol. 251, 15, 1995.
6. Johansen, T. H., J. Lothe, and H. Bratsberg, *Phys. Rev. Lett.*, Vol. 80, 4757, 1998.
7. Johansen, T. H., *Phys. Rev. B*, Vol. 59, 11187, 1999.
8. Johansen, T. H., *Phys. Rev. B*, Vol. 60, 9690, 1999.
9. Johansen, T. H., *Supercond. Sci. Technol.*, Vol. 13, R121, 2000.
10. Johansen, T. H., C. Wang, Q.-Y. Chen, and W.-K. Chu, *J. Appl. Phys.*, Vol. 88, 2730, 2000.
11. Johansen, T. H., Q.-Y. Chen, and W.-K. Chu, *Physica C*, Vol. 349, 201, 2001.
12. Diko, P., *Supercond. Sci. Technol.*, Vol. 11, 68, 1998.
13. Zmorayova, K., P. Diko, and G. Krabbes, *Physica C*, 445–448, 436, 2006.
14. Johansen, T. H., *Phys. Rev. B*, Vol. 60, 9690, 1999.
15. Zhou, Y.-H. and H.-D. Yong, *Phys. Rev. B*, Vol. 76, 094523, 2007.
16. Yong, H.-D., Y.-H. Zhou, and J. Zeng, *J. Appl. Phys.*, Vol. 104, 113902, 2008.
17. Johansen, T. H. and D. V. Shantsev, *Supercond. Sci. Technol.*, Vol. 16, 1109, 2003.

Magnetoelastic Model of Magnetizable Media

Ke Jin, Yong Kou, and Xiaojing Zheng

Department of Mechanics and Engineering Science, Lanzhou University
Lanzhou, Gansu 730000, China

Abstract— The interaction between applied magnetic fields and magnetizable media has always been very complicated, since the distribution of magnetic field and magnetic forces in the media cannot be measured directly. Modeling the magnetic field and magnetic force in the deformable media which can agree to the measured results on the surface of media is very important. In this work, the authors review the various existing magnetic force models by comparing the premises on which the model is based, formulas and results.

1. INTRODUCTION

The interaction between applied magnetic fields and magnetizable media is very complicated. Firstly, magnetic domains in the magnetizable media will rotate along the direction of external applied magnetic field, and the media consequently becomes magnetized; then magnetic field will change inside and outside the media due to the magnetization, in which media acts as a inductive “source”; finally, magnetic force and torque are exerted on the media which cause deformation and movement, while the deformation and movement of magnetized media will further affect the distribution of the field of these “sources”. Modeling the magnetoelastic Interaction between magnetizable and deformable media and magnetic field which can agree to the measured results on the surface of media is very important. Although many works [1–11] exist in this area, researchers are still puzzled by a number of issues: why are there so many coexistent theoretical models for magnetic force of magnetizable media under a applied magnetic field? How to make a choice? How to calculate the interaction between applied magnetic fields and magnetizable media considering the effect of the magneto-elastic coupling? The rationality of magneto-elastic model and validity of calculation are related to whether the model can predict the stability of ferromagnetic structure accurately under the magnetic field in the nuclear industry, high-tech area, etc. In this paper, the authors review the various existing magnetic force models by comparing the premises on which the model is based, formulas and results, and discuss their scope of application.

2. MAGNETIC DIPOLE MODEL

According to the Ampere’s molecular current hypothesis, each molecular in the magnetic media can be considered as a closed circuit. When there’s no applied magnetic field, the moments become randomly disordered, canceling each other and eliminating the magnetic field. As external magnetic field is applied, magnetic moments aligned to the direction of magnetic field, which makes media becomes magnetized. We study a molecular closed circuit near the point P (Fig. 1).

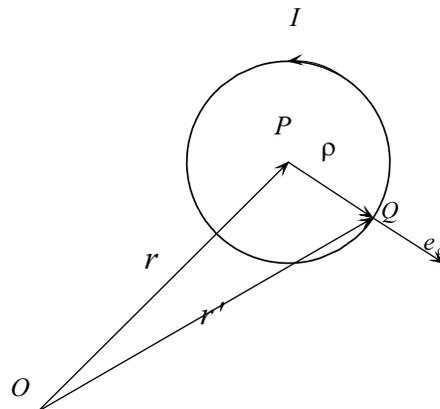


Figure 1.

The total force and torque on the circuit are then given by Lorentz law:

$$\mathbf{F} = \oint \mathbf{I} d\mathbf{l} \mathbf{e}_\tau \times \mathbf{B}(\mathbf{r}') \quad \mathbf{C} = \oint \rho \mathbf{e}_\rho \times [\mathbf{I} d\mathbf{l} \mathbf{e}_\tau \times \mathbf{B}(\mathbf{r}')] \quad (1)$$

Here, $\mathbf{B}(\mathbf{r}')$ is the total magnetic induction at point Q , which can be expressed with the magnetic induction at point P by expanding $\mathbf{B}(\mathbf{r}')$ in a Taylor series:

$$\mathbf{B}(\mathbf{r}') = \mathbf{B}(\mathbf{r}) + \rho \mathbf{e}_\rho \cdot \nabla \mathbf{B}(\mathbf{r}) + \dots \quad (2)$$

Hence:

$$\mathbf{F} = \oint \mathbf{I} d\mathbf{l} \mathbf{e}_\tau \times (\mathbf{B}(\mathbf{r}) + \rho \mathbf{e}_\rho \cdot \nabla \mathbf{B}(\mathbf{r})) = I_\rho \oint \mathbf{e}_\tau (\mathbf{e}_\rho \cdot \nabla) \times \mathbf{B}(\mathbf{r}) d\mathbf{l} = \frac{\mathbf{I}}{2} \left[\oint (\rho \mathbf{e}_\rho \times \mathbf{e}_\tau d\mathbf{l}) \times \nabla \right] \times \mathbf{B}(\mathbf{r}) \quad (3)$$

We defined the equivalent magnetic moment of the circuit:

$$m = \frac{\mathbf{I}}{2} \oint \rho \mathbf{e}_\rho \times \mathbf{e}_\tau d\mathbf{l} \quad (4)$$

Thus the total force on the circuit can be written as:

$$\mathbf{F} = (\mathbf{m} \times \nabla) \times \mathbf{B}(\mathbf{r}) = [\nabla \mathbf{B}(\mathbf{r})] \cdot \mathbf{m} - \mathbf{m} [\nabla \cdot \mathbf{B}(\mathbf{r})] = [\nabla \mathbf{B}(\mathbf{r})] \cdot \mathbf{m} \quad (5)$$

In the Eq. (5), the condition $\nabla \cdot \mathbf{B}(\mathbf{r}) = 0$ is used. With the aid of definition of magnetization vector \mathbf{M} , the magnetic body force in unit volume can be given:

$$\mathbf{f}(\mathbf{r}) = [\nabla \mathbf{B}(\mathbf{r})] \cdot \mathbf{M}(\mathbf{r}) \quad (6)$$

A similar calculation leads to torque in unit volume formula as:

$$\mathbf{c}(\mathbf{r}) = \mathbf{M}(\mathbf{r}) \times \mathbf{B}(\mathbf{r}) \quad (7)$$

Considering linear constitutive relationship between magnetic induction and magnetization $\mathbf{M}(\mathbf{r}) = \chi_m \mathbf{B}(\mathbf{r}) / \mu_0 \mu_r$, we have $\mathbf{c}(\mathbf{r}) = 0$. The research object of magnetic dipole model is infinite media, since magnetic body force formula (6) has nothing to do with the magnetic force at boundary surface.

3. MAXWELL'S MODEL BASED ON ELECTROMAGNETIC STRESS TENSOR

According to Maxwell [1], "the force on a piece of magnetic substance carrying a current may for convenience of calculation" be divided into two parts: the force on the element due to the current and the force in consequence of the magnetism. The first part is the same as the force on a non-magnetic media, written as $\mathbf{J} \times \mathbf{B}$ (Here, \mathbf{J} is conduction current, while magnetization current is not included); the second part calculate the magnetic force due to magnetization, denoted as $\mu_0 (\nabla \mathbf{H}) \cdot \mathbf{M}$, where μ_0 is vacuum permeability. Compared to the dipole model, it can be noticed that the force in consequence of the magnetism in this manner is analogous to taking $\mu_0 \mathbf{H}$ instead of the magnetic induction \mathbf{B} at point inside the body in the dipole model, which means that calculating the force with external magnetic induction replacing the total magnetic induction due to magnetization.

Using $\nabla \times \mathbf{H} = \mathbf{J}$, the magnetic force exerted on an element of a body can be expressed as:

$$\mathbf{f} = \mu_0 (\nabla \mathbf{H}) \cdot \mathbf{M} + \mathbf{J} \times \mathbf{B} = \nabla \cdot \left(\mathbf{B} \mathbf{H} - \frac{1}{2} \mu_0 H^2 \boldsymbol{\delta} \right) \quad (8)$$

in which $\boldsymbol{\delta}$ is unit tensor. Maxwell had proved that the quantity in the bracket is tensor, subsequently, which is the famous "Maxwell stress tensor". Hence, the total force on a body can be calculated by integration over a surface surrounding it,

$$\mathbf{F} = \int_{\Omega} \nabla \cdot \left(\mathbf{B} \mathbf{H} - \frac{1}{2} \mu_0 H^2 \boldsymbol{\delta} \right) d\tau = \int_S \mathbf{n} \cdot \left(\mathbf{B} \mathbf{H} - \frac{1}{2} \mu_0 H^2 \boldsymbol{\delta} \right) dS \quad (9)$$

When S surrounds the whole body, the total magnetic force can be obtained. The treatment above requires no modification for elastic solid; but all the coordinates in this treatment must be the instantaneous coordinates of the particles after the deformation [2]. We can also see from Eq. (9) that magnetic traction at the boundary of the magnetized media is not taken into account since formula (9) just transform the force inside the body to the force over the surface in the treatment. Moreover, magnetic induction contributed by magnetization of media is not taken into consideration, therefore, the magnetic body force in the magnetized media without carrying current omits the term $\mu_0(\nabla M^2)/2$ compared to dipole model. Some works [2, 10, 11] let the surface S expand to outside the body, hence, \mathbf{B} may be replaced by $\mu_0\mathbf{H}$, since the surface integration can be carried out outside the body. Thus, we have:

$$\mathbf{F} = \frac{1}{\mu_0} \int_S \mathbf{n} \cdot \left(\mathbf{B}\mathbf{B} - \frac{1}{2}B^2\boldsymbol{\delta} \right) dS \quad (10)$$

However, the magnetic induction at point outside but near the magnetized media is not easy to calculate since the external magnetic field near the media will change due to the magnetization of media.

4. BROWN'S MODEL

In Brown's thesis [2], magnetic force exerted on a magnetized and conduction-current-carrying body subjected to an externally applied magnetic field was discussed by consideration of external magnetic induction \mathbf{B}_0 instead of the real magnetic induction \mathbf{B} at points inside body. Thus, Eq. (6) can be expressed as:

$$\mathbf{f}(\mathbf{r}) = [\nabla\mathbf{B}(\mathbf{r})] \cdot \mathbf{M}(\mathbf{r}) = [\mathbf{M}(\mathbf{r}) \cdot \nabla]\mathbf{B}_0 - (\nabla \times \mathbf{B}_0) \times \mathbf{M}(\mathbf{r}) = [\mathbf{M}(\mathbf{r}) \cdot \nabla]\mathbf{B}_0 \quad (11)$$

where $\nabla \times \mathbf{B}_0 = 0$ is considered in Eq. (11). The total force on the body can expressed by the integration:

$$\mathbf{F} = \int \mathbf{M} \cdot \nabla\mathbf{B}_0 d\tau \quad (12)$$

Alternative expressions for the total force on the body can be derived by use of vector identities and $\nabla \cdot \mathbf{B} = 0$ always, $\nabla \times \mathbf{B}_0 = 0$ in external magnetic field:

$$\mathbf{F} = \int \mathbf{n} \cdot \mathbf{M}\mathbf{B}_0 ds - \int (\nabla \cdot \mathbf{M})\mathbf{B}_0 d\tau \quad (13)$$

$$\mathbf{F} = - \int (\mathbf{n} \times \mathbf{M}) \times \mathbf{B}_0 ds + \int (\nabla \times \mathbf{M}) \times \mathbf{B}_0 d\tau \quad (14)$$

Eq. (13) and Eq. (14) are called the *pole model* and *Ampere-current model*, respectively. The above three representation of the magnetic body total force are equivalent, while the distribution of magnetic forces are widely different. Moreover, magnetic induction \mathbf{B} in these three model are external magnetic induction \mathbf{B}_0 , in which magnetic induction contributed by the matter in the magnetized body is not taken into consideration.

Then, Brown modified the model by evaluating the long-range part of the magnetic force exerted on the matter in an arbitrary volume of a magnetized body by all sources outside the volume, including the rest of body. In Fig. 2, τ_1 is the volume of magnetized body, and τ_2 is the arbitrary volume of τ_1 , S_1 and S_2 are closed surfaces surrounding τ_1 and τ_2 , respectively.

Defined: $\mathbf{H}_0 = \mathbf{H} - \mathbf{H}_{12}$, $\mathbf{B}_0 = \mathbf{B} - \mathbf{B}_{12}$. In which \mathbf{H}_0 and \mathbf{B}_0 are external magnetic field intensity and induction, \mathbf{H} and \mathbf{B} are total intensity and induction at point inside τ_2 , \mathbf{H}_{12} and \mathbf{B}_{12} are the parts of them contributed by the matter in τ_1 , calculated at points inside τ_2 . Thus

$$\mathbf{F} = \mu_0 \int \mathbf{M} \cdot \nabla\mathbf{H}_0 d\tau = \mu_0 \int \mathbf{M} \cdot \nabla\mathbf{H} d\tau + \mathbf{F}_1 \quad (15)$$

where

$$\mathbf{F}_1 = - \lim_{S_1 S_2 \rightarrow S} \left\{ \mu_0 \int_{S_2} \mathbf{n}_2 \cdot \mathbf{m}_2 \mathbf{H}_{12} dS_2 + \mu_0 \int_{\tau_2} (-\nabla_2 \cdot \mathbf{M}_2) \mathbf{H}_{12} d\tau_2 \right\} \quad (16)$$

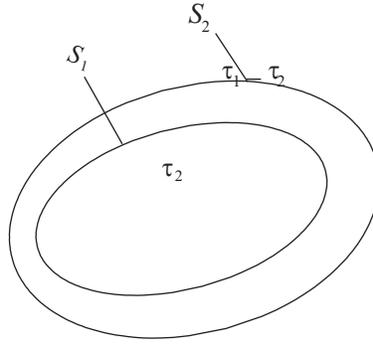


Figure 2: Method of calculating the force on part of a body [2].

With the aid of discontinuity of \mathbf{H} across the surface S , we can obtain:

$$\mathbf{F}_1 = \frac{1}{2}\mu_0 \int \mathbf{n}M_n^2 dS \tag{17}$$

Therefore

$$\mathbf{F} = \mu_0 \int \mathbf{M} \cdot \nabla \mathbf{H} d\tau + \frac{1}{2}\mu_0 \int \mathbf{n}M_n^2 dS \tag{18}$$

Formula (18) can be rewritten by setting $\mathbf{B} = \mu_0(\mathbf{H} + \mathbf{M})$ and considering $\mathbf{M} \times (\nabla \times \mathbf{B}) = \nabla \mathbf{B} \cdot \mathbf{M} - (\mathbf{M} \cdot \nabla)\mathbf{B}$:

$$\mathbf{F} = \int \nabla \mathbf{B} \cdot \mathbf{M} d\tau - \frac{1}{2}\mu_0 \int \mathbf{n}M_t^2 dS \tag{19}$$

Formula (19) can be separated into the representation of magnetic body force and magnetic traction:

$$\mathbf{f}^{em} = \nabla \mathbf{B} \cdot \mathbf{M}, \quad \mathbf{F}^{em} = -\frac{1}{2}\mu_0 M_t^2 \mathbf{n} \tag{20}$$

Compared with formulas (6), the representation of magnetic body force is consistent with the magnetic body force based on dipole model, however, the magnetic traction is added. The total magnetic force calculated by formulae (18) and (19) are equivalent, while the distribution of magnetic forces are different. As can be seen from the above derivation, the procedure for obtaining the model by Brown it is the following: First calculate the magnetic force exerted on the magnetized body by source of field entirely outside it, without considering the magnetic induction due to magnetizing, thus, formulae (12) (13) and (14) are obtained; then modify, by calculating the long-rang part of the magnetic force contributed by the matter in the body, thus, formulae (18) and (19) are got. Though the representation of magnetic body force is as same as dipole model, the magnetic traction is just calculated by carrying the volume integral of the magnetic body force enclose a region which is separated by a discontinuity surface. Therefore, the magnetic traction calculated in this manner is the representation of magnetic body force on the boundary surface, not the physically significant magnetic traction.

5. ZHOU-ZHENG'S MODEL BASED ON GENERALIZED VARIATIONSL PRINCIPLE

Zhou-Zheng [7, 9] generalize the variational principle of magneto-elasticity into the system of magnetoelastic interaction of magnetizable media in magnetic fields. This generalization is achieved by establishing a functional of total energy of the system through considering the summation of internal energy, and magnetic energy.

$$\Pi[\mathbf{u}, \phi] = \frac{1}{2} \int_{\Omega^+(\mathbf{u})} \mu_0 \mu_r (\nabla \phi^+)^2 dV + \frac{1}{2} \int_{\Omega^-(\mathbf{u})} \mu_0 (\nabla \phi^-)^2 dV + \frac{1}{2} \int_{s_0} \mathbf{n} \cdot \mathbf{B}_0 \phi^- dS + \frac{1}{2} \int_{\Omega^+} \boldsymbol{\varepsilon} : \mathbf{t} dV \tag{21}$$

Considering the variations of ϕ and u independent from each other, according to the arithmetic of variational approach, we have

$$\delta \Pi = \delta_\phi \Pi + \delta_u \Pi \tag{22}$$

On the assumption that the magnetization of media is linear and the material is incompressible, one can reduce Eq. (22) into the form

$$\begin{aligned} \delta\Pi = & - \int_{\Omega^+(\mathbf{u})} \mu_0\mu_r(\nabla^2\phi^+)\delta\phi^+dV - \int_{\Omega^-(\mathbf{u})} \mu_0(\nabla^2\phi^-)\delta\phi^-dV + \oint_S \mu_0 \left\{ \mu_r \frac{\partial\phi^+}{\partial n} - \frac{\partial\phi^-}{\partial n} \right\} \delta\phi dS \\ & + \int_{s_0} \left\{ \mu_0 \frac{\partial\phi^-}{\partial n} + \mathbf{n} \cdot \mathbf{B}_0 \right\} \delta\phi^- dS \end{aligned} \quad (23)$$

For Eq. (23), let $\delta\Pi = 0$, Due to arbitrariness and independence of $\delta\phi$, we can get

$$\mathbf{f}^{em} = \frac{\mu_0\mu_r\chi_m}{2} \nabla(\mathbf{H}^+)^2, \quad \mathbf{F}^{em} = -\frac{\mu_0\chi_m(\mu_r + 1)}{2} (H_r^+)^2 \mathbf{n}^+ \quad (24)$$

Magnetic body force in formula (24) is consonant with magnetic body force in formula (6) based on the dipole model, but the magnetic surface force is added, which is consistent with the formula of magnetic stress on both sides of the interface of magnetic media given by Faraday, which therefore physical significance will be attributed.

ACKNOWLEDGMENT

Supported by the National Science Foundation of China (No. 90405005) and the Ph.D. Fund of the Ministry of Education of China (No. 20050730016).

REFERENCES

1. Maxwell, J. C., *A Treatise on Electricity and Magnetism*, Oxford University Press, 1881.
2. Brown, W. F., *Magnetoelastic Interactions*, Springer, 1966.
3. Eringen, A. C. and G. Maugin, *Electrodynamics of Continua*, Springer, New York, 1989.
4. Eringen, A. C., *Int. J. Engr. Sci.*, Vol. 27, 363, 1989.
5. Moon, E. C. and Y. H. Pao, *ASME J. App. and Mech.*, Vol. 35, 53, 1968.
6. Moon, E. C., *Magneto-solid Mechanics*, Wiley, New York, 1984.
7. Zhou, Y. H., X. J. Zheng, and K. Miya, *Fusion Eng. and Design*, Vol. 30, 325, 1995.
8. Pao, Y. H. and C. S. Yeh, *Int. Z. Engr. Sci.*, Vol. 11, 415, 1973.
9. Zhou, Y. H. and X. J. Zheng, *Int. J. Engng. Sci.*, Vol. 35, No. 15, 1405–1417, 1997.
10. Landau, L. D., E. M. Lifshitz, and L. P. Pitaevskii, *Electrodynamic of Continuous Media*, Elsevier, 1960.
11. Jackson, J. D., *Classical Electrodynamics*, John Wiley & Sons, New York, 1976.

Theoretical Analysis on Quantum Well at Undoped GaN/In_xGa_{1-x}N/GaN Heterostructure Interface

Shah Mohammad Bahauddin, Farha Diba Sumana, Md. Rubaiyat Hossain,
Md. Ahsan Uddin, and Zahid Hasan Mahmood

Department of Applied Physics, Electronics & Communication Engineering
University of Dhaka, Bangladesh

Abstract— Energy band diagram of undoped GaN/In_xGa_{1-x}N/GaN hetero-structure is simulated which consists of a GaN cap layer with an InGaN layer sandwiched on GaN buffer layer. The formation of triangular quantum well at the hetero-interface with the existence of 2DEG and 2DHG has been confirmed and the dependence of sheet carrier density on the alloy composition and the thickness of InGaN layer for different temperature have been simulated.

1. INTRODUCTION

III Nitride wide band gap semiconductors have become most functional materials in recent years to emit light in ultra-violet/visible wavelength region as well as to provide high power and high frequency applications in devices like HEMTS, HFETS and MODFETS. The significant lattice mismatch between InN and GaN or AlN can result in a large piezoelectric charge, which is advantageous for various device applications. Again, InGaN alloys on GaN based heterostructure possesses tunable band gap along thermal stability and high resistivity which promises that In-rich heterostructure can be used for high frequency shorter to longer wavelength signal detection. In III Nitride heterostructure, the difference in the spontaneous and piezoelectric polarization between two different layers would result in a fixed sheet of polarization charge at the interface. This charge tends to attract high concentration of electrons or holes depending on the net polarization at the interface. For a net positive polarization at the interface, this confinement results in 2DEG and for a net negative polarization this result in 2DHG [1, 2].

In our present study, we have analyzed the energy band diagram of undoped GaN/In_xGa_{1-x}N/GaN hetero-structure using 1D Poisson/Schrodinger solver [3]. The heterostructure consists of 50 nm GaN cap layer with an InGaN layer on 100 nm GaN buffer layer (Fig. 1). The formation of triangular quantum well at the hetero-interface with the existence of 2DEG and 2DHG has been confirmed. In Fig. 2, the Fermi level is taken as the reference energy level having 0.0 eV energy. The penetration of Fermi energy by both conduction and valence band indicates the creation of quantum well. It is evident from the figure that the well is formed at InGaN layer. The dependence of sheet carrier density on the alloy composition for different temperature has also been investigated.

2. SIMULATION AND RESULTS

The dependence of sheet carrier density on the alloy composition for temperature range of 150 K to 350 K has been simulated. From Fig. 3, it can be stated that the density of 2DHG and 2DEG increases with the increase of alloy composition and almost immune to the variation of temperature. The results presented in Fig. 3 also dictate that 2DHG is slightly dependent on temperature. Though the density of 2DHG is more dependent [1] to temperature than 2DEG, they remain unaffected at higher alloy composition.

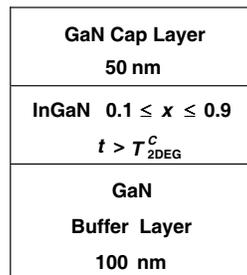


Figure 1: Undoped GaN/In_xGa_{1-x}N/GaN hetero-structure.

For a given alloy composition the thickness of InGaN layer was varied and corresponding sheet carrier density was observed (Fig. 4). The thickness was taken from 1.0 nm to 50 nm. Again for a fixed InGaN layer thickness, the compositional parameter x was varied and resultant sheet carrier density was also observed (Fig. 5). From the figures, it can be shown that the critical thickness, T^C is decreased with the increase of alloy composition. Thus the relation between the sheet carrier concentrations with the thickness of InGaN layer has been studied and consequently critical thickness for the formation of 2DHG and 2DEG has been found. Table 1 presents simulated critical thickness, T^C for different alloy composition at 300 K. Moreover, it can be obtained from Fig. 6 that the sheet carrier concentration reaches a saturation value after about 50 nm. We believe that the critical thickness is dependent on the degree of band bending and the surface charge and the piezoelectric polarization charge of the interface. This critical and saturation thickness can be very useful to predict the performance of ultra-thin structures.

An important electronic transport property, the low field mobility values for 2DEG has been also investigated for different mole fraction of Indium with corresponding sheet concentrations. Here

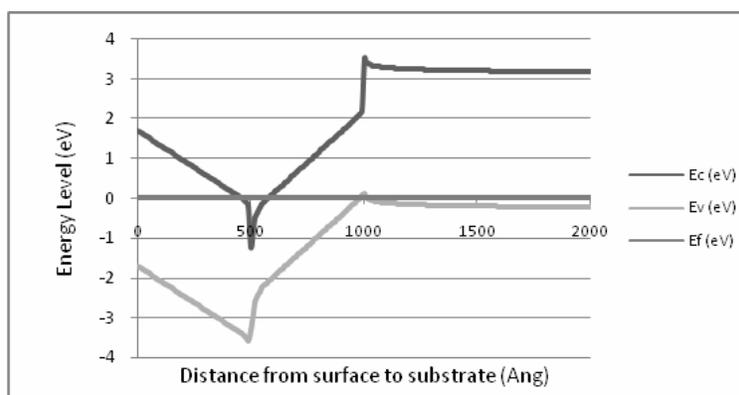


Figure 2: Band diagram of undoped $\text{In}_{0.5}\text{Ga}_{0.5}\text{N}$, Layer thickness of 50 nm at $T = 300$ K.

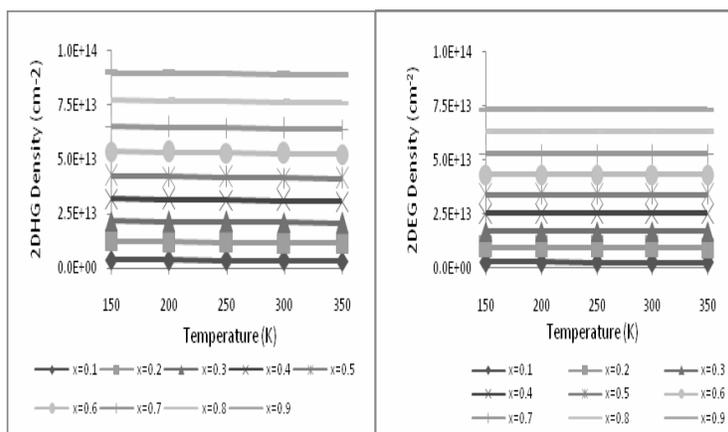


Figure 3: Dependence of 2DHG and 2DEG concentration on temperature for variable alloy composition.

Table 1: Critical thickness for the formation of 2DEG and 2DHG in different alloy composition estimated from Figs. 5 and 6.

Alloy Composition, x	Critical Thickness for 2DEG $T_{2\text{DEG}}^C$ (nm)	Critical Thickness for 2DHG $T_{2\text{DHG}}^C$ (nm)
0.1	> 20	> 10
0.2	≈ 10	≈ 5
0.3	> 5	> 2.5
0.5	≈ 2.5	≈ 2
0.7	< 2	1
0.9	< 1	> 0

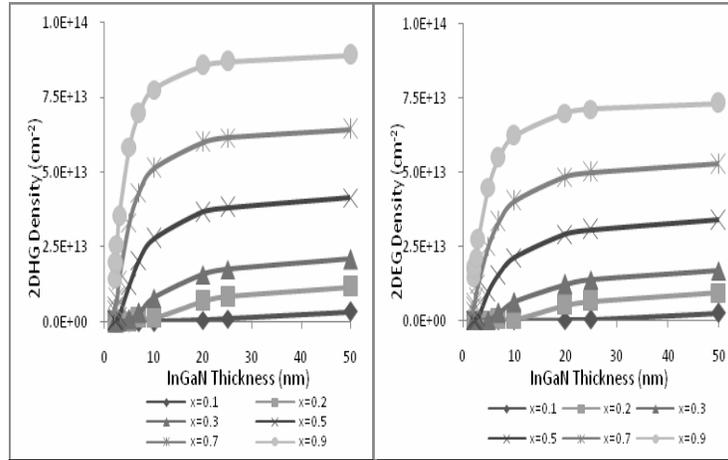


Figure 4: Dependence of 2DHG and 2DEG concentration on InGaN thickness for different alloy composition.

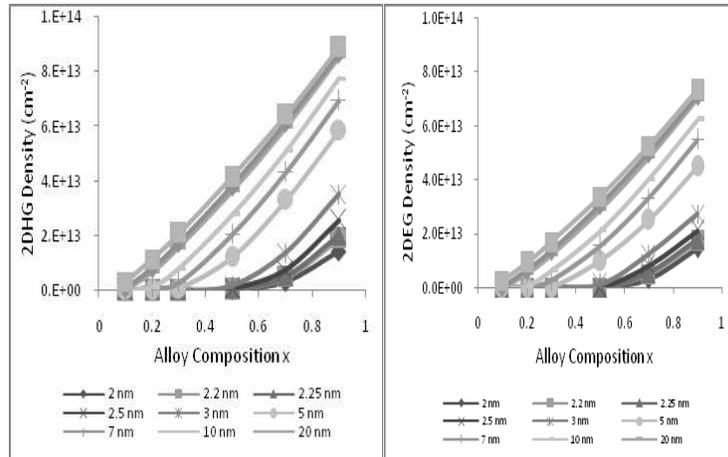


Figure 5: 2DHG and 2DEG density with Alloy composition for different InGaN layer thickness.

scattering by two types of phonons are relevant for our study which are acoustic phonon by deformation potential coupling and longitudinal optical phonon. But for higher temperature (≈ 300 K) the mobility is limited by polar optical phonon scattering [2]. Thus we have neglected acoustic phonon scattering as our interest is restricted for operations at room temperature. Scattering by polar optical phonons is highly inelastic. An analytic expression for the momentum relaxation rate in 2DEG can be evaluated which is able to match experimental data [2, 4]. Thus the 2DEG mobility in undoped InGaN/GaN quantum well has been calculated by using the following equation,

$$\mu = \frac{2\mathcal{E}^* q_0 \hbar^2 F(y)}{em^* \omega_0 N_B G(k_0)}$$

where, q_0 is the polar optical phonon wave vector and $F(y)$ is a dimensionless function of sheet carrier density and temperature. Here, N_B is the Bose-Einstein distribution function and $G(k_0)$ is the form factor, which is dependent on the geometry of the device and can be assumed unity for perfect 2DEG in this case.

In our present work, the temperature dependence of the mobility for LO phonon has been observed (Fig. 6). The figure shows that mobility increases as the Indium mole fraction rises. An average mobility of $3000 \text{ cm}^2/\text{V} \cdot \text{s}$ has been obtained at room temperature though mobility is gradually decreased at higher temperature. In contrast to AlGaIn/GaN, where $\mu_{300\text{K}} \approx 2000 \text{ cm}^2/\text{V} \cdot \text{s}$ has been reported [5], InGaIn/GaN samples have shown considerably large mobility in room temperature.

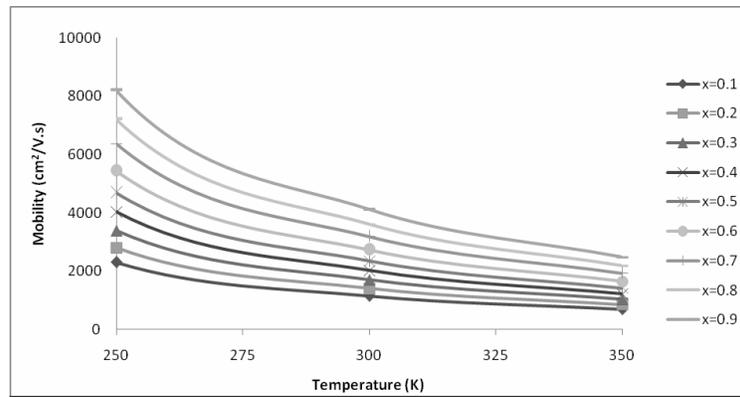


Figure 6: LO Phonon limited Mobility calculation of 2DEG as a function of temperature for different alloy composition.

REFERENCES

1. Upal, T. N., M. A. Uddin, M. Hossain, F. Jahan, and Z. H. Mahmood, "Study of charge density at $\text{In}_x\text{Ga}_{1-x}\text{N}/\text{GaN}$ heterostructure interface," *2nd IEDST'09*, 1–4, June 2009.
2. Wood, C. and D. Jena, *Polarization Effects in Semiconductors: From Ab Initio Theory to Device Applications*, Debdeep Jena, Springer, 2008.
3. Tan, I. H., G. L. Snider, L. D. Chang, and E. L. Hu, "A self consistent solution of Schrodinger-Poisson equations using a nonuniform mesh," *J. Appl. Phys.*, Vol. 68, No. 8, 4071–4076, 1990.
4. Gelmont, B. L., M. Shur, and M. Stroscio, "Polar optical-phonon scattering in three- and two-dimensional electron gases," *J. Appl. Phys.*, Vol. 77, 657, 1995.
5. Gaska, R., J. W. Yang, A. Osinsky, Q. Chen, M. A. Khan, A. O. Orlov, G. L. Snider, and M. Shur, "Electron transport in AlGa_xN-GaN heterostructures grown on 6H-SiC substrates," *Appl. Phys. Lett.*, Vol. 72, 707, 1998.

Optimization for Rotating-scanning Ring Arrays of Synthetic Aperture Radiometer

Weiying Sun, Hao Liu, Zhang Cheng, Shengwei Zhang, and Ji Wu
Center for Space Science and Applied Research, Chinese Academy of Sciences
Beijing 100190, China

Abstract— During the past decade, interferometric techniques, inherited from radio astronomy, has been successfully applied to Earth remote sensing. Achieving adequate sensitivity is a critical issue of a synthetic aperture radiometer, which depends largely on the type of an antenna array. To reach the sensitivity of each sampling of visibility function as identical as possible is one of requirements of the optimum antenna array. That is, total integral time of the antenna array distributes as average as possible to each sampling of visibility function.

In this paper, optimization of rotating-scanning ring arrays is examined on the view of u - v coverage and sensitivity of each sampling point. The relations between the sensitivity and length of baselines and aperture of element antenna are studied. The objective functions and solution space of optimization for rotating scan ring arrays are established, and based on simulated annealing algorithm, the simulation results of rotating scan ring array are presented. The simulation results proved that visibility coverage on u - v plane and equality of sensitivity of each sampling are quite improved.

The methods and results may be used to the similar antenna array design of radiometer, such as millimeter wave radiometer in Geostationary Orbit in the future.

1. INTRODUCTION

With rapid development of single chip millimeter-wave integrated circuits and multi-channel digital correlators, synthetic aperture interferometric radiometer (SAIR) becomes mainstream instrument in passive microwave remote sensing domain since 1990s.

Imaging principle of SAIR is based on Fourier transform relationship between amendatory brightness temperature and visibility function. Antenna array of SAIR can obtain visibility functions of observation objects. Inversion algorithm can realize image reconstruction from visibility functions to amendatory brightness temperatures. Therefore, sparse-aperture design for the antenna array of SAIR is one of important parts of general design for SAIR. It is also an important factor that influencing applications of SAIR.

Many efforts have been made on sparse-aperture design for the antenna array of SAIR in recent years. However, as passive microwave remote sensing domain continuously enlarging, new application backgrounds and purposes appear. And new requirements of general design and properties for SAIR have been imported. The current sparse-aperture design for antenna array cannot satisfy new application requirements of SAIR.

Two-dimensional ring arrays can adopt isotropic radiation and multi-channel receiving techniques. The multi-beam of them covers the whole detective spatial domain. And flexible gain in different directions realizes easily. Main objectives of the current sparse-aperture design for the antenna array are low redundancy and side-lobe [1–4]. In recent years, a new scanning type, called rotating scanning, is introduced to SAIR. It can achieve similar samplings with much less antenna elements.

In view of requirements of general design and properties for SAIR with rotating-scanning ring arrays, in this paper, we first study the relationship between the distribution of antenna array elements and temperature sensitivity. Next, optimization objectives are proposed and traditional simulated annealing algorithm is mended. We optimize the distribution of 26-element ring array for SAIR. The ring array consists of 22-element antenna with large aperture and 4-element antenna with small aperture. The large antenna aperture is twice as large as the small antenna aperture. Finally, considering installation error, we add a set of random errors to the optimization result and analyze the influence of installation error on the system properties of SAIR.

2. THE EFFECT OF ANTENNA ARRAY DISTRIBUTION ON SENSITIVITY OF SAIR

Temperature sensitivity is one of key indicators of properties of SAIR. There are two methods to improve temperature sensitivity. One is to prolong integral time, the other is to enlarge antenna aperture.

As sensitivity is identical, integral time of each visibility function has relation to element antenna aperture, which is given by [5, 6]

$$\tau = \left[\frac{2\sqrt{3}}{\eta_a \pi} \cdot \left(\frac{\Delta d}{d_a} \right)^2 \frac{T_A + T_R}{NeDT} \cdot \alpha_w \right]^2 \frac{\alpha_{ds}}{B} \cdot N_v. \quad (1)$$

where τ is integral time of visibility function, η_a is aperture efficiency of element antenna, Δd is length of the shortest baseline, d_a is diameter of element antenna aperture, d_{a1} is diameter of small element antenna aperture and equal to Δd , d_{a2} is diameter of large element antenna aperture and equal to $2\Delta d$, T_A is noise temperature of element antenna, T_R is noise temperature of receiver, $NeDT$ is sensitivity of SAIR, α_w is impact factor of windowed function, α_{ds} is impact factor of finite bits, B is bandwidth, and N_v is number of sampling in spatial frequency domain.

The integral time of visibility function that obtained by small aperture element antenna is $\tau_{\Delta d}$. We have

$$\tau_{\Delta d} = \left[\frac{2\sqrt{3}}{\eta_a \pi} \cdot \left(\frac{\Delta d}{\Delta d} \right)^2 \frac{T_A + T_R}{NeDT} \cdot \alpha_w \right]^2 \frac{\alpha_{ds}}{B} \cdot N_v = A. \quad (2)$$

And then we rearrange integral time given in (1) in terms of $d_{a2} = 2\Delta d$ so that the integral time of visibility function that obtained by large aperture element antenna can be carried out. That is,

$$\tau_{2\Delta d} = \left(\left(\frac{\Delta d^2}{d_a^2} \right) \right)^2 \cdot A = \left(\frac{\Delta d}{4\Delta d^2} \right)^2 \cdot A = \frac{\tau_{\Delta d}}{16}. \quad (3)$$

Moreover, the integral time of visibility function that obtained by compounding that between large aperture and small aperture element antenna can be calculated. That is,

$$t_{\text{Com}} = \left(\left(\frac{\Delta d^2}{d_a^2} \right) \right)^2 \cdot A = \left(\left(\frac{\Delta d^2}{(\Delta d^2 + 4\Delta d^2)/2} \right) \right)^2 \cdot A = \frac{4}{25} \tau_{\Delta d}. \quad (4)$$

If the sequence of baseline length is $\{\Delta d, 2\Delta d, 3\Delta d, 4\Delta d, \dots, n\Delta d, \dots\}$; And the number of sampling points of the shortest baseline Δd as rotating scanning, is N ; And the interval between sampling points in a certain circle that obtained by different baselines is even and equal to those between the sampling points that obtained by the shortest baseline as antenna array rotating. Therefore, the number of sampling points of the baseline, which length is equal to $2\Delta d$, should be $2N$; And the integral time of every sampling points decrease to $\tau_{\Delta d}/2$. On the analogy of this, the relation between integral time τ and baseline length $y = n\Delta d$ is $\tau \propto 1/n$.

3. THE OPTIMIZATION OBJECTIVES AND OPTIMIZATION STRATEGY

In order to simply the optimization problem, the diameter of the ring is normalization. The optimization objectives are the following:

- The sensitivity of each sampling point is as identical as possible.
- The center of gravity of the array approaches the geometrical center of the ring.
- Ascending sequence of baseline length of the array is an arithmetic series. And the ratio between the longest baseline and the shortest baseline is equal to 86.

Some optimizing algorithms have been proposed to optimize antenna array of SAIR, such as Leech algorithm, Ishiguro algorithm, Blanton algorithm, genetic algorithm and simulated annealing algorithm, etc. [7–10]. Simulated annealing algorithm has advantages of simple describing, flexibly exercising, high operation efficiency and few restrictions by initial condition. The procedure of traditional simulated annealing algorithm first creates solution series of combinatorial optimization problems by using of Metropolis algorithm; then the transfer from current solution i to new solution j is decided by transfer probability P_t of Metropolis algorithm, which is expressed as [11]

$$P_t(i \Rightarrow j) = \begin{cases} 1, & \text{while } f(j) \leq f(i) \\ \exp\left(\frac{f(i)-f(j)}{t}\right), & \text{else} \end{cases}, \quad (5)$$

A large initial value of t should be selected; and then the value gradually decrease during transfer procedure; finally, simulated annealing algorithm is stopped as a certain stop principle is satisfied after several repeated calculate.

In this paper, in view of the characteristic of metropolis algorithm, that is, jumping out of local optimal solution, we add memory function to traditional simulated annealing algorithm. Memory function can record those objective function values that obtained in the whole searching procedure and guarantee the global optimal solution can be saved. Moreover, we add a local search algorithm to simulated annealing algorithm. The local search algorithm can iterate within neighborhood solution and make the objective functions gradually optimizing until it cannot be more optimized. The initial positions of element antennas on the ring are random; the initial temperature t_0 is 900; cooling rate is 0.9995; and terminal temperature t_{end} is 1.

4. OPTIMIZATION RESULTS

We optimize the distribution of 26-element rotating-scanning ring array. The ring array consists of 22-element antenna with large aperture and 4-element antenna with small aperture. The large antenna aperture is twice as large as the small antenna aperture. The optimization results are shown in Fig. 1.

As can be seen from Figs. 1(a), (b), (c) and (d), the center of gravity of the optimization array, which is denoted by '*', approaches the geometrical center of the ring; the baseline sequence is close to arithmetic; the ratio between the longest baseline and the shortest baseline is equal to 86; the equality of sensitivity of each sampling point is quite improved and the SLL of the array factor is comparatively low.

Moreover, we add stochastic phasic error to each element antenna. The stochastic phasic errors are less than 0.1° . The optimization antenna array and its properties that affected by stochastic phasic errors are shown in Fig. 2.

As can be seen from Figs. 2(a), (b), (c) and (d), under influence of stochastic phasic errors,

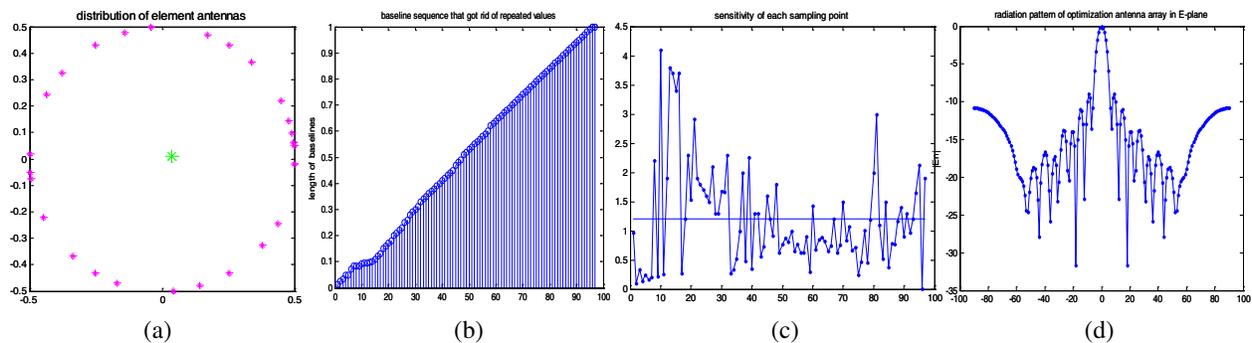


Figure 1: Properties of the optimization ring array. (a) Distribution of element antennas; (b) Baseline sequence that got rid of repeated values; (c) Sensitivity of each sampling point; (d) Radiation pattern in E -plane.

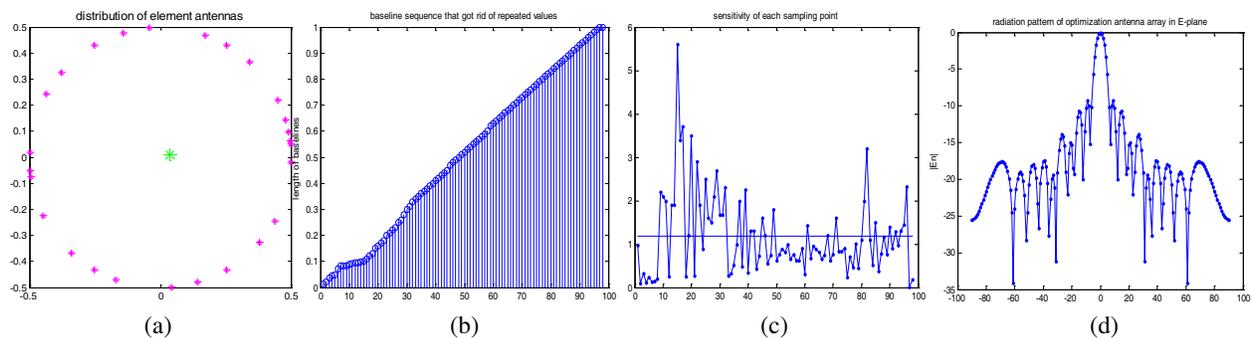


Figure 2: Properties of the optimization ring array that affected by stochastic phasic errors. (a) Distribution of element antennas; (b) Baseline sequence that got rid of repeated values; (c) Sensitivity of each sampling point; (d) Radiation pattern in E -plane.

variation of the center of gravity, the baseline sequence and radiation pattern of the optimization array are very little and can be ignored. The equality of sensitivity of each sampling point is a little worse.

5. CONCLUSIONS

In this paper, we first study the relationship between distribution of antenna array and its temperature sensitivity. Secondly, the optimization objectives are proposed. Thirdly, we optimize the distribution of 26-element rotating scan ring array with improved simulated annealing algorithm. Finally, we add stochastic phasic errors to the optimization array and calculate its properties as effecting by those errors.

The simulation results proved that visibility coverage on u - v plane is perfect and equality of sensitivity of each sampling is quite improved. Therefore, the quality of retrieval image can be greatly bettered. The SLL of the array factor is comparatively low. Moreover, the center of gravity of the optimization array approaches the geometrical center of the ring, which diminishing the difficulty of engineering application. The influence of stochastic phasic errors on the center of gravity and the baseline sequence and radiation pattern of the optimization array are very little and can be ignored. However, the equality of sensitivity of each sampling point is a little worse under influence of stochastic phasic errors.

The methods and results may be used to the similar antenna array design of SAIR, such as millimeter wave radiometer in geostationary orbit in the future.

ACKNOWLEDGMENT

The authors acknowledge funding for this research from the National High Technology Research and Development Program of China.

REFERENCES

1. Boeringer, D. W. and D. H. Werner, "Particle swarm optimization versus genetic algorithm for phased array synthesis," *IEEE Trans. Antennas Propagat.*, Vol. 52, No. 3, 771–779, 2004.
2. Sun, W. and J. Wu, "Optimization of antenna array for interferometric synthetic aperture radiometer," *IEEE 2005 International Symposium on Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications Proceedings*, 293–296, 2005.
3. Sun, W., B. He, and J. Wu, "Optimization of fourier plane coverage of antenna arrays for SPORT," *PIERS Proceedings*, 533–537, Hangzhou, China, August 22–26, 2005.
4. He, Y. T., Y. S. Jiang, and H. T. Chen, "Studies of optimization and imaging properties of two-dimensional circle array for mm-wave synthetic aperture system," *Journal of Remote Sensing*, Vol. 11, No. 1, 33–38, 2007.
5. Le Vine, D. M., "The sensitivity of synthetic aperture radiometers for remote sensing of applications from space," *Radio Science*, Vol. 25, No. 4, 441–453, 1990.
6. Camps, A., I. Corbella, J. Bara, and F. Torres, "Radiometric sensitivity computation in aperture synthesis interferometric radiometry," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 36, No. 2, 680–685, 1998.
7. Leech, J., "On the representation of $1, 2, \dots, n$ by differences," *J. London Math. Soc.*, Vol. 31, 160–169, 1956.
8. Ishiguro, M., "Minimum redundancy linear arrays for a large number of antennas," *Radio Science*, Vol. 15, 1163–1170, 1980.
9. Blanton, K. A. and J. H. McClellan, "New search algorithm for minimum redundancy linear arrays," *Acoustics, Speech, and Signal Processing*, Vol. 2, 1361–1364, 1991.
10. Guo, J. M., B. Liu, and E. K. Mao, "Synthesis of circular array patterns of VHF band," *Chinese Journal of Radio Science*, Vol. 23, No. 4, 792–796, 2008.
11. He, B. and J. Wu, "Two dimensional least redundant array for interferometric synthetic aperture radiometer," *IEEE International Geoscience and Remote Sensing Symposium*, 4371–4373, 2003.

A Physically Based Parameterized Method to Estimate Cloud Liquid Water over Land Using AMSR-E

Yongqian Wang^{1,2}, Jiancheng Shi¹, and Bangsen Tian^{1,3}

¹Institute of Remote Sensing Applications, Chinese Academy of Sciences, Beijing 100101, China

²Graduate University of Chinese Academy of Sciences, Beijing 100049, China

³Center for Earth Observation and Digital Earth of Chinese Academy of Sciences, Beijing 100101, China

Abstract— This paper presents a new scheme that retrieves cloud liquid water (CLW) over land using AMSR-E without the help of any ancillary data. Surface emission model (AIEM) and one dimensional atmosphere transfer model (1DRTM) were combined to generate a database. Through analyzing the simulated database, it's found that the ratio of the polarization difference obtained from 36.5 and 89 GHz (PDR_CLW) is sensitive to CLW. The algorithm was validated using AMSR-E observations over SGP. The validation results are very encouraging.

1. INTRODUCTION

Cloud liquid water (CLW) is difficult to measure because the emissivity over land is often very high and non-uniform. Currently, almost all kinds of the methods to retrieve CLW have a drawback that ancillary data is indispensable [1]. In this paper, a physically based parameterized technique for retrieving CLW in non-precipitating clouds over land will be described. This technique is not dependent on any other ancillary data so that this method provides a new opportunity to establish a long term global dataset about CLW.

2. THEORETICAL BASIS FOR THE ALGORITHM

In this study, we analyze the characteristics of microwave frequencies by generating a simulated surface emission database for the sensor parameters of AMSR-E with its six frequencies for both horizontal and vertical polarization at the incidence angle of 55°, using the Advanced Integral Equation Model (AIEM) [2]. This database includes a wide range of volumetric soil moisture (2% to 44% at 2% interval). Surface roughness parameters included rms height from 0.25 cm to 3 cm at a 0.25 cm interval and correlation length from 2.5 cm to 30 cm at a 2.5 cm interval. The commonly used Gaussian correlation function was used in the simulation since it is a good approximation for the microwave measurement frequencies of this study. Through analyzing the database, the relationship of emissivities polarization difference between two adjacent frequencies as can be described as:

$$\Delta\varepsilon(23.8 \text{ GHz}) = 0.959 \cdot \Delta\varepsilon(18.7 \text{ GHz}) \quad (1)$$

$$\Delta\varepsilon(89 \text{ GHz}) = 0.778 \cdot \Delta\varepsilon(36.5 \text{ GHz}) \quad (2)$$

where $\Delta\varepsilon$ is the emissivity polarization difference of one particular frequency. Equations (1) and (2) indicate that the relationship of $\Delta\varepsilon$ between two adjacent AMSR-E frequencies is proportional as a constant, independent of surface properties, which supply the prerequisite to minimize the surface information in the CLW retrieval process.

Considering the proportional relationship between surface emissivity polarization differences of two adjacent frequencies, the ratio of the brightness temperature polarization differences obtained from two adjacent AMSR-E frequencies can be represented as:

$$\frac{T_{Bv}(f_1) - T_{Bh}(f_1)}{T_{Bv}(f_2) - T_{Bh}(f_2)} = b(f_1, f_2) \cdot \frac{\Gamma(f_1) \cdot T_s - \Gamma(f_1) \cdot T_a(f_1) \cdot (1 - \Gamma(f_1))}{\Gamma(f_2) \cdot T_s - \Gamma(f_2) \cdot T_a(f_2) \cdot (1 - \Gamma(f_2))} \quad (3)$$

Using the polarization difference ratio (**PDR**) in (3) results in the cancellation of the effects of ground surface emission signals, so the ratio of brightness temperature polarization differences in (3) now is influenced by atmosphere parameters such as transmittance $\Gamma(f)$, atmosphere effective temperature $T_a(f)$ and surface temperature T_s .

In order to study the sensitivity of **PDR** in (3) to different atmosphere parameters, one dimensional radiative transfer model (**1DRTM**) was used to simulate AMSR-E brightness temperatures

adapted from a plane-parallel model described in [3]. The **AIEM** simulated surface emissivity database and atmosphere radiosonde profiles from Southern Great Plains (SGP) of the United States in 2007 were set as input in **1DRTM**.

The algorithm developed in [4] was found that errors were greatest for the lowest level clouds. Greenwald [5] gave the results that for the type of cloud which situates from 2 to 3 km, the response at 85.5 GHz to cloud liquid water is very small. But if the same cloud is moved to a greater height (4–5 km), the brightness temperature for both polarizations at 85.5 GHz will have an obvious change according the CLW. In this study, we tried to use $\Delta T_B(36.5)/\Delta T_B(89)$ (PDR_CLW) to retrieve CLW, so PDR_CLW will also be sensitive to the height of cloud. An index which is sensitive to the height of cloud was developed (using HI to stand for this index). The index can be written as:

$$HI = \frac{T_s - T_v(89 \text{ GHz})}{T_s + T_v(89 \text{ GHz})} \quad (4)$$

This index is constructed based on the sensitivity of 89 GHz to the height of cloud [5]. The pattern of HI showed in (4) can minimize the surface information and maximize the cloud height information contained in 89 GHz brightness temperature. The sensitivity of HI to cloud height was analyzed using simulated database. The results are plotted in Fig. 1(a). When cloud height moves from 2~3 km to 3~4 km, the relative variation of HI is about 25% and it reaches 40% when the height is from 3~4 km to 4~5 km. From the figure, it also can be seen that PDR_CLW which is used to retrieve CLW is also sensitive to cloud height. However, it is around 10%, much smaller compared to HI. In Fig. 1(b), the sensitivity of HI and PDR_CLW to CLW was plotted. When CLW was changing 0.01 cm, the relative variation of PDR_CLW is about 10%, while for HI, it is less than 2%. Consequently, Compared to PDR_CLW, HI is sensitive to cloud height, but not CLW.

After developing the cloud height index, the uncertainty in the CLW retrieval algorithm caused by water vapor also should be considered. Greenwald et al. [5] pointed out that the polarization difference of 89 GHz brightness temperature will decrease due to the increase of water vapor in the atmosphere, as a result, the sensitivity of brightness temperatures at 89 GHz to CLW will decrease. Because 23.8 GHz is close to the water vapor line, the ratio of polarization difference of 18.7 and 23.8 GHz (PDR_WV) can be used to detect water vapor. Because we want to use PDR_WV to detect the water vapor influences in the CLW retrieval, PDR_WV should not sensitive to CLW. In Fig. 1(b), it also has plotted the sensitivity of PDR_WV to CLW. The relative variations are less than 1%, almost no change according to the increase of CLW. Deeter in [6] used PDR_WV to retrieve WVP and the variability of WVP on the order of 0.6 cm can be detected. It's reasonable to use PDR_WV to remove the uncertainty caused by water vapor in the CLW retrieval.

Another factor which influenced the value of PDR_CLW is surface temperature. In this paper, the surface temperature based on TIR measurements cannot be used in the CLW retrieval algorithm. The best solution is that surface temperature can be obtained by using AMSR-E measurements themselves. AMSR-E observations at 36.5 GHz vertical polarization are used to estimate surface temperature. The approach used in this study is a modification of that described in [7]. The relationship between surface temperature and Ka-band measurements of AMSR-E is established using ground based observations.

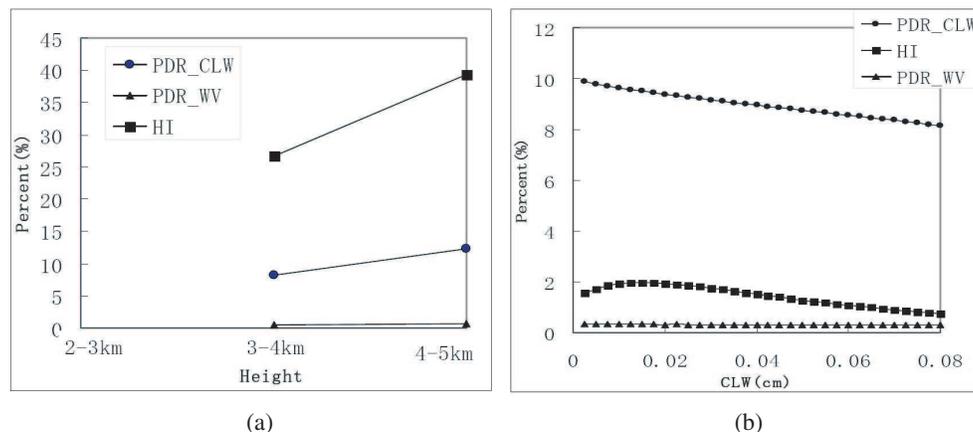


Figure 1: The sensitivity of HI, PDR_CLW and PDR_WV to cloud height and CLW.

3. REGRESSION ANALYSIS

Five hundred atmosphere radiosonde profiles in SGP of the United States were chosen to simulate the brightness temperatures at AMSR-E configuration and build the database. For each profile, CLW is set to the value from 0.0 to 0.08 cm at the interval of 0.0025 cm and at three heights for 2–3, 3–4, and 4–5 km. At the same time, the surface emissivity database based on AIEM was also used in the simulation. We parameterized this relationship and gave a form as follows:

$$\text{PDR_CLW} = a \cdot \exp(b \cdot \text{CLW}) \quad (5)$$

Table 1: The slope and intercept for the linear relationship in Fig. 11.

	Slope	Intercept
2–3 km	−0.0506	17.85
3–4 km	−0.0506	18.026
4–5 km	−0.0506	18.463

In (5), a and b are fitting coefficients. Firstly, through analyzing the data, it's found that coefficient a is only dependent on water vapor and the relationship is linear. The relationship between b and water vapor is weak, however, coefficient b is highly dependent on surface temperature and cloud height. The coefficient b is linear dependent on surface temperature. The linear coefficients are listed in Table 1. It can be seen from Table 1 that among these three sets of linear coefficients, slopes are same and the differences only exist in intercepts. The relationship between HI and intercepts is also linear.

Until now, (5) can be expanded to consider all other affected factors, and the final form can be written as:

$$\text{PDR_CLW} = (P_1 \cdot \text{PDR_WV} + P_2) \cdot \exp((P_3 \cdot T_s + P_4 \cdot \text{HI} + P_5) \cdot \text{CLW}) \quad (6)$$

In (6), P_i values are fitting coefficients to be determined through regressing the simulated database. For retrievals of CLW, (6) can be changed to the form as follows:

$$\text{CLW} = \frac{\ln \frac{\text{PDR_CLW}}{P_1 \cdot \text{PDR_WV} + P_2}}{P_3 \cdot T_s + P_4 \cdot \text{HI} + P_5} \quad (7)$$

Equation (7) is the final form to retrieve CLW. It relates CLW to different variables which can be directly obtained from AMSR-E measurements such as PDR_CLW and PDR_WV, or can be obtained indirectly from satellite observations including HI and surface temperatures.

Values of the coefficients in (6) were determined for using simulated database described in Section 2.3. Fig. 2 has shown the comparison between the regression results for PDR_CLW and the original ones which were computed by using simulated database. By the way, the solid lines in the figure are the plots of the function of $y = x$. Obviously, the figure is fitting well and the relative error is about 4.5%.

4. VALIDATION

The first validation way is that the results are validated in clear sky. The MWR CLW is used to discriminate the clear sky. We define all the cases in which MWR CLW is less than 0.05 mm as clear sky. AMSR-E brightness temperatures at the clear sky were used to retrieve CLW. The retrieved results indicate the magnitude of the retrieval error, because in the clear sky conditions, obviously, the CLW would be very small. The root mean square (RMS) errors in clear sky are typically about 0.04 mm. Comparisons of AMSR-E based and MWR-based CLW retrievals is the second validation method. The scatter-plot of the comparisons is presented in Fig. 3(a). The statistical RMS is 0.11 mm. CLW retrieved by MODIS which is also in the Aqua platform is used to validate the AMSR-E retrieved results. The scatter plot is shown in Fig. 3(b) and comparisons are correlated well with the RMS of 0.14 mm, but also indicate large discrepancies for some points.

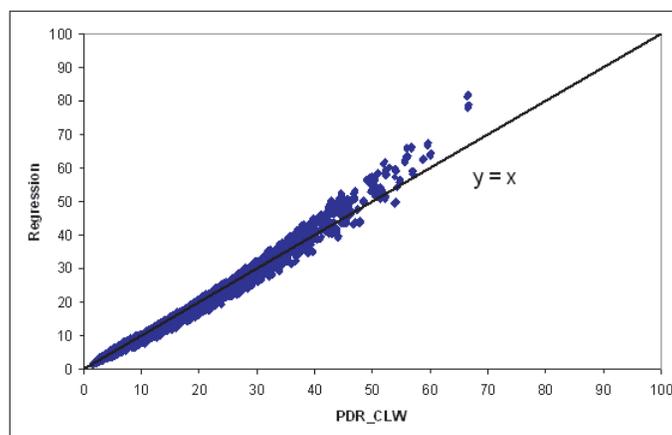


Figure 2: Comparison between the regression results for PDR_CLW and the original ones.

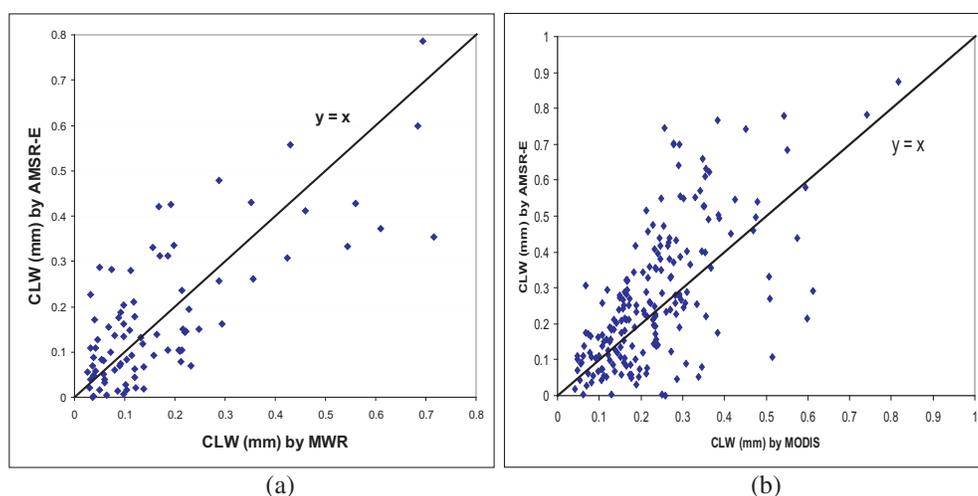


Figure 3: Comparison of CLW retrieved by MWR, MODIS and AMSR-E.

5. CONCLUSIONS

This algorithm attempts to retrieve CLW without the dependence on any other ancillary data and the retrieval result is very encouraging. Besides being physically based, the algorithm has the characteristics of mathematical simplicity. Therefore, it has the potential to process large satellite radiometer data sets to generate CLW product on global scale.

REFERENCES

1. Deeter, M. N. and J. Vivekanandan, "New dual-frequency microwave technique for retrieving liquid water path over land," *J. Geophys. Res.*, Vol. 111, D15209, doi:10.1029/2005JD006784, 2006.
2. Chen, K. S., T. D. Wu, L. Tsang, Q. Li, J. Shi, and A. K. Fung, "Emission of rough surfaces calculated by the integral equation method with comparison to three-dimensional moment method simulations," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 41, 90–101, 2003.
3. Olson, W. S., P. Bauer, N. F. Viltard, D. E. Johnson, W. Tao, R. Meneghini, and L. Liao, "A melting-layer model for passive/active microwave remote sensing applications. Part I: Model formulation and comparison with observations," *Journal of Applied Meteorology*, Vol. 40, 1145–1163, 2001.
4. Jones, A. and H. Vonder, "Passive microwave remote sensing of cloud liquid water over land regions," *Journal of Geophysical Research*, Vol. 95, 16, 673–16, 683, 1990.
5. Greenwald, T. J., C. L. Combs, A. S. Jones, D. L. Randel, and T. H. Vonder Haar, "Further developments in estimating cloud liquid water over land using microwave and infrared satellite measurements," *Journal of Applied Meteorology*, Vol. 36, 389–405, 1997.

6. Deeter, M. N., “A new satellite retrieval method for precipitable water vapor over land and ocean,” *Geophysical Research Letters*, Vol. 34, L02815, doi:10.1029/2006GL028019, 2007.
7. Jackson, T. J., M. H. Cosh, and R. Bindlish, “A five-year validation of AMSR-E soil moisture products,” *IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2008*, 2008.

Transient Responses Analysis of Ultra-wideband Filters Illuminated by High-power Electromagnetic Pulses (EMP)

Zheng Jiang¹, Jian Wang², and Wen-Yan Yin^{1,2}

¹Center for Optical and EM Research (COER), Zhejiang University, Hangzhou 310058, China

²Center for Microwave and RF Technologies (CMRFT), Shanghai Jiao Tong University
Shanghai 200240, China

Abstract— A modified FDTD method is used to analyze the transient response of two ultra-wideband (UWB) filters with a switch diode enclosed, which are illuminated by a high-power electromagnetic pulse (EMP), respectively. Our attention will be paid to the characterization of EMP interference effects on their input and output signals described by the S -parameters. The incident EMP can take any direction and polarization state. We introduce failure parameter to denote the impact effectiveness of the EMP on the filter. Some statistical analysis can enhance our confidence in the suppression of certain intentional high-power EMP interference on ultra-wideband filters which are widely used in modern ultra-wideband communication systems.

1. INTRODUCTION

Impacts of intentional electromagnetic interference (IEMI) on electric systems are attracting much attention in EM community more recently. For example, terrorist threats are increasing over the world [1, 2]. On the other hand, significant research interest on UWB techniques has been aroused in both academic and industrial community since the U.S. Federal Communications Commission (FCC) released the unlicensed use of UWB devices in February 2002 [3]. In particular, the bandpass filters (BPFs) with large fractional bandwidths (FBWs) have been driven by their wide applications of indoor and outdoor UWB communication systems. Unfortunately, an ultra-wideband communication system could be easily interfered by an ultra-wideband EMP (UWEMP). Thus, transient analysis of UWB BPFs illuminated by external electromagnetic waves will provide fundamental knowledge about the protection of electric devices against an EMP.

We know that two types of UWB bandpass filters have been proposed in [3] and [4] recently. In [3], the miniaturized reconfigurable UWB filter with PIN diodes consists of two bended open stubs of quarter-wavelength and a connecting stub of half-wavelength, and in [4] a compact ultra-wideband bandpass filter with improved upper stopband can be found.

In this paper, we use the FDTD method to investigate transient responses of UWB filters illuminated by an EMP. In Section 2, the geometry of filters is provided. In Section 3, the FDTD equations, the definition of S -parameters and breakdown failure error based on S -parameters are presented. In Section 4, statistical analysis is performed to demonstrate the effects of external interference on the UWB filters. Some conclusions are drawn in Section 5.

2. PROBLEM DEFINITION

A 3-D structure of a reconfigurable UWB filter is shown in Fig. 1(a) [4, 5], which was fabricated on single-layer Rogers substrate with the dielectric constant of $\epsilon_r = 3.38$ and the thickness of $h = 0.8$ mm. The bandstop filter is shown in Fig. 1(b), which consists of two bended open stubs l_2 , and bended connecting line l_1 , in the form of square loop. The bended short stubs have a length of $l_2 = \lambda_g/4$ and bended connecting line $l_1 = \lambda_g/2$, where λ_g is the guided wavelength at the center frequency (6.35 GHz) of the bandstop. Between the folded stubs is the gap $g = 0.3$ mm, and the reconfigurable filter changes to bandpass case when the connecting stubs are grounded by the pin diodes D1 and D2, as shown in Fig. 1(c). The filter has fixed dimensions referred to Fig. 1 of $w_0 = 5$ mm, $w_1 = 1.7$ mm, $w_2 = 1.2$ mm. Fig. 2 also shows the schematic of the 2-D UWB bandpass filter, which has a square slot-line resonator (SLR) with slot-line stub etched in the ground plane and the modified T-shaped feed lines. The filter dimensions are $w_1 = 1.2$ mm, $w_2 = 0.3$ mm, $w_3 = 0.4$ mm, $L_1 = 3.6$ mm, $L_2 = 8.95$ mm, $S = 0.9$ mm, $D = 8.35$ mm [4, 5].

The incident EMP wave vector is described by

$$\vec{k}_{inc} = \vec{e}_x \sin \theta \cos \varphi + \vec{e}_y \sin \theta \sin \varphi + \vec{e}_z \cos \theta \quad (1)$$

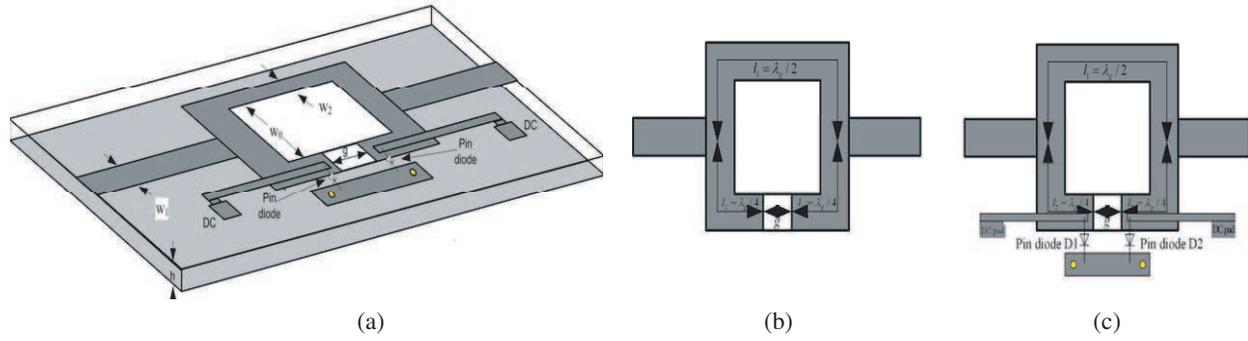


Figure 1. The reconfigurable UWB filters [4,5]: (a) 3-D structure, (b) UWB-bandstop and (c) UWB-bandpass filters.

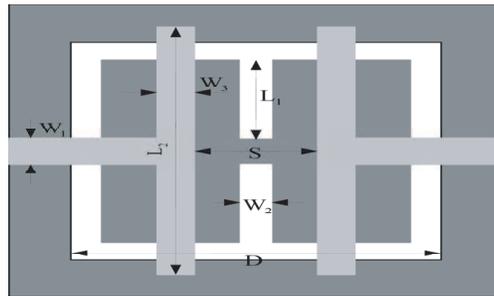


Figure 2. The UWB bandpass filter with an improved upper stop band [4,5].

where \vec{e}_x , \vec{e}_y and \vec{e}_z are three unit vectors in the Cartesian coordinate system; with $0^\circ < \theta < 180^\circ$ and $0^\circ < \varphi < 360^\circ$. The polarizations of the external EMP are determined by the polarization angle ψ , and the incident field component $E^{(inc)}$ is described by a triple-exponential function as follows:

$$E^{(inc)}(t) = \sum_{i=1}^3 \frac{a_i}{b_i \sqrt{\pi/2}} e^{-2\left(\frac{t-t_i}{\tau_i}\right)^2} \quad (2)$$

where the coefficients a_i , b_i , t_i and τ_i are obtained according to the pulse waveform such as shown in [2].

3. FDTD FORMULATION

It is well known that the time-dependent Maxwell's curl equations can be written as

$$\nabla \times \vec{H} = \varepsilon \frac{\partial \vec{E}}{\partial t} + \sigma \vec{E} \quad (3a)$$

$$\nabla \times \vec{E} = -\mu \frac{\partial \vec{H}}{\partial t} \quad (3b)$$

where \vec{E} and \vec{H} are the electric and magnetic vector fields, respectively, and ε is the permittivity, σ is the conductivity, and μ is the permeability of the medium. (3a) and (3b) can be solved by the FDTD method:

$$\vec{H}^{n+1/2} = \vec{H}^{n-1/2} - (\Delta t/\mu) \nabla \times \vec{E}^n \quad (4a)$$

$$\vec{E}^{n+1} = \frac{\varepsilon - \sigma \Delta t/2}{\varepsilon + \sigma \Delta t/2} \vec{E}^n + \frac{1}{\varepsilon + \sigma \Delta t/2} \nabla \times \vec{H}^{n+1/2} \quad (4b)$$

The frequency-dependent scattering matrix coefficients can be calculated from the transient results obtained the FDTD method, i.e.,

$$\vec{V}^r = \vec{S} \vec{V}^i \quad (5)$$

where \vec{V}^r and \vec{V}^i are the reflected and incident voltage vectors, respectively, and \vec{S} is the scattering matrix. The two-step method will be used in this study to capture the S -parameters of the filter. At first, the FDTD simulation calculates the sum of incident and reflected waveforms at the input port. Then, the input port microstrip line will be extended from the source to the far absorbing wall to obtain the incident waveform, and the incident waveform will be recorded. The reflected waveform can be obtained by subtracting the whole signal from the incident plus at the input port. The output ports only need register the transmitted waveforms at the first step. The scattering parameters $S_{i,j}$ can then be obtained by the Fourier transform of the transient waveforms as

$$S_{i,j} = FT [v_j(t)] / FT [v_i(t)] \quad (6)$$

Note that the reference planes are chosen with large distance from the circuit discontinuities so as to eliminate the evanescent waves. These distances are included in the definition of the circuit so that no phase correction is performed for the scattering coefficients.

To describe the grade of failure effects caused by the external high-power EMP, the relative failure ratio $FS_{i,j}$ is defined by

$$FS_{i,j} = \left| S_{i,j}^{(\text{EMP})} - S_{i,j}^{(\text{without EMP})} \right| / \left| S_{i,j}^{(\text{without EMP})} \right| \quad (7)$$

where $S_{i,j}^{(\text{EMP})}$ is the $S_{i,j}$ of the filter illuminated by an EMP, $S_{i,j}^{(\text{without EMP})}$ is the $S_{i,j}$ with no external EMP considered. The failure means that the component is unable to perform its function, but without any physical damage [6].

4. NUMERICAL RESULTS AND DISCUSSION

In order to check the accuracy of FDTD code, we calculate the S -parameters of the UWB filter with no external EMP source introduced at first. Figs. 3(a) and (b) show the simulated S -parameters of the reconfigurable UWB bandstop/bandpass filter, in comparison with the simulated (HFSS) results, respectively. It is obvious that excellent agreements are obtained between them. Fig. 3(c) shows the simulated S -parameters of the UWB bandpass filter, compared with the measurement results given in [4], and excellent agreements are also obtained between them.

In our calculation, the incident EMP direction is set to be $\varphi = 90^\circ$ and $\theta = 90^\circ$, with a polarization angle $\psi = 90^\circ$. In the FDTD simulation, it is assumed that the filters are shielded by a metallic enclosure when illuminated by an external high-power EMP (HPEMP), so the inner input signal magnitude will be much smaller than that of the external EMP (about 24 dB decrease in magnitude). It is seen that both UWB filters can be disturbed by the external EMP seriously, and in particular for the S_{21} and S_{11} of these UWB-bandpass filters.

Finally, statistical analysis is carried out so as to predict susceptibility of the filter S -parameter, and in particular for S_{11} . We introduce different EMP illumination, such as double-exponential pulse, a fast- and UWB- pulse as described in Table 1 [6].

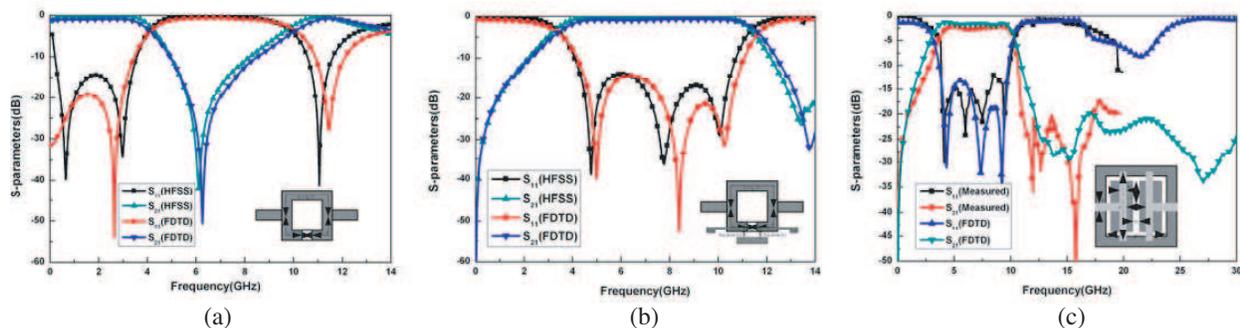


Figure 3. Comparisons between the S -parameters between the FDTD and HFSS. (a) The case of reconfigurable filter-UWB bandstop filter, (b) the reconfigurable filter-UWB bandpass filter, and (c) the simulated and measured S -parameters of the UWB bandpass filter.

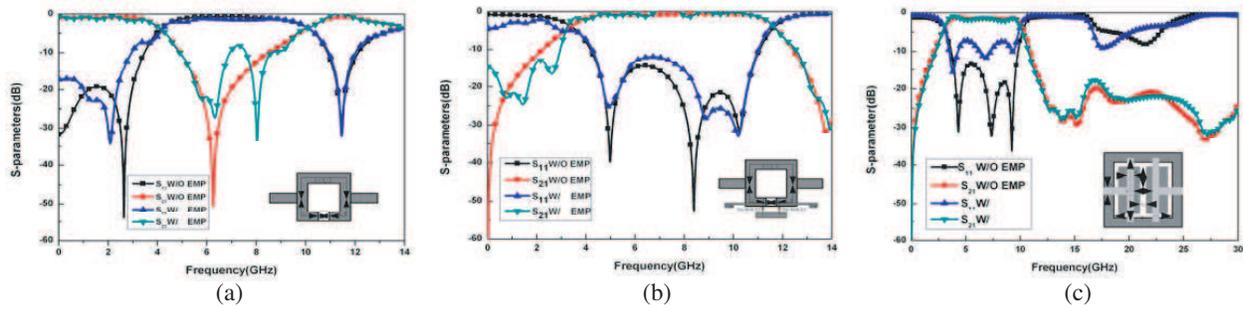


Figure 4. The S -parameters of the filters enclosed by a metallic enclosure, and W/O: without EMP, and W/: with an EMP. The case of (a) reconfigurable UWB-bandstop filter, (b) reconfigurable UWB-bandpass filter, and (c) UWB bandpass with improved upper stopband.

Table 1.

Pulse	α/s^{-1}	β/s^{-1}
EMP (fast)	6.0×10^{10}	4.0×10^9
EMP (UWB)	3.0×10^{10}	2.5×10^8

The external pulse is described by

$$E_{\alpha}(t) = E_0 k(\alpha, \beta) \left(e^{-\beta t} - e^{-\alpha t} \right) \tag{8a}$$

$$k(\alpha, \beta) = \left(e^{-\beta(\ln(\alpha) - \ln(\beta)) / (\alpha - \beta)} - e^{-\alpha(\ln(\alpha) - \ln(\beta)) / (\alpha - \beta)} \right)^{-1} \tag{8b}$$

The average values of $FS_{i,j}$ are calculated from 3 to 11 GHz for different magnitudes of the incident pulse, as shown in Fig. 5. It is found that the fast-pulse can affect the UWB bandpass filter more seriously, compared with other UWB pulse.

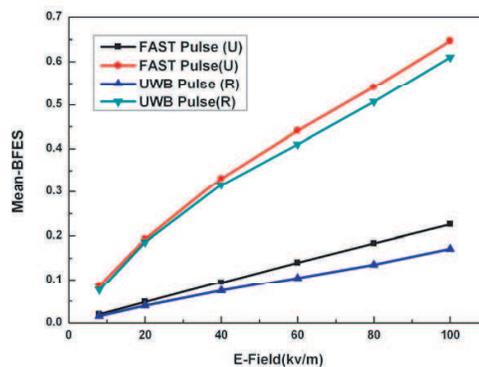


Figure 5. The average values of $FS_{i,j}$ of two types of bandpass filters illuminated by a fast- and UWB-pulse, respectively. U: bandpass filter with improved upper bandstop, R: reconfigurable bandpass filter.

5. CONCLUSION

The FDTD method is implemented to characterize EM interference in UWB filters caused by an external EMP, and the S -parameters are captured for different geometrical and physical parameters of filters. Our results show that the susceptibility of two filters is very sensitive to the fast and UWB pulse waveforms. Although metallic enclosure is often adopted to protect inner circuit systems against external EMI, further protection method is still needed.

REFERENCES

1. Radasky, W. A., C. E. Baum, and M. W. Wik, "Introduction to the special issue on high-power electromagnetics (HPEM) and intentional electromagnetic interference (IEMI)," *IEEE Trans. Electromagn. Compat.*, Vol. 46, No. 3, 314–321, 2004.
2. Giri, D. V. and F. M. Tesche, "Classification of intentional electromagnetic environments (IEME)," *IEEE Trans. Electromagn. Compat.*, Vol. 46, No. 3, 322–328, 2004.
3. "Revision of part 15 of the commission's rules regarding ultra-wideband transmission systems," Tech. Rep., ET-Docket 98-153, FCC02-48, 2002.
4. Karim, M. F., Y.-X. Guo, Z. N. Chen, and L. C. Ong, "Miniaturized reconfigurable and switchable filter from UWB to 2.4 GHz WLAN using PIN diodes," *IEEE MTT-S Int. Dig.*, 509–512, 2009.
5. Mondal, P., M. K. Mandal, and A. Chakrabarty, "Compact ultra-wideband bandpass filter with improved upper stopband," *IEEE Microw. Wireless Compon. Lett.*, Vol. 17, No. 9, 643–645, 2007.
6. Camp, M., H. Garth, H. Garbe, and H. Haase, "Predicting the breakdown behavior of micro-controllers under EMP/UWB impact using a statistical analysis," *IEEE Trans. Electromagn. Compat.*, Vol. 46, No. 3, 368–379, 2004.

Research on New Technology on Protection of Electronic Systems from High Power Electromagnetic Pulse

Zhonghao Lu, Chunxiao Jian, Shuanglin Wan, and Peiguo Liu

School of Electronic Science and Engineering, NUDT, Changsha 410073, China

Abstract— The definition of high power microwave (HPM) weapon and its damage to people and other weapons is discussed first. Recent progress and perspectives of the research on the new technology on protection of electronic systems from high power electromagnetic pulse (EMP) damage are expatiated in the form of new structure and material. The hypothesis is proved by computer simulation at last.

1. INTRODUCTION

In the high technology war of the modern times, the HPM weapons, the orientational weapons and the infrasonic weapons come into our eyes rather frequently that the menace in the battlefield becomes complex and multiplex and the damage it caused is greatly increased that the electromagnetic environment in the battlefield tends to become more and more complex, which will bring great effect on the weapon's efficacy and security. Therefore, in order to face the challenge and ensure the people's safety, the protection capability of the weapons must be strengthened to protect itself from electromagnetic strike [1, 2].

2. HPM WEAPON AND ITS MENACE

EMP and HPM weapons have the ability to disturb the ionosphere and the magnetosphere in the atmosphere, which will not only paralyze the communication network and the power transmission equipment system, but also damage the national defense command and control systems, information and communication systems etc. Hence these weapons are of great efficiency and truculence in the modern battlefield. Considering its great importance, the chief power countries in the world are putting a large number of scientists and a great much of money into the research of the HPM weapons, which drives this new type of weapon cut a brilliant figure in the modern war [3].

The microwave energy can thrill through all the aperture with the size of wavelength and poor conductor like glass and fiber etc and enter the inside of the target, thus HPM weapons can cause damage to people hiding in the hermetical blindage and armored vehicle. When the microwave energy takes act on the body, low intensity energy will cause people's psychology dysfunction, which will leads to awful symptom like dysphonic, headache, decrease of memory etc which will cause the soldiers' disability and lose battle effectiveness. As the intensity of energy increases, skin burn and damage of internal apparatus will be caused. Expressly, when the power density of microwave energy comes to 800 MW/cm^2 , it will leads to death in just one second [2, 4, 5]. High power EMP brings deadly damage to equipment of electronic, information, electric power photoelectricity and microwave, it can also cause the semiconductor insulation or the integrate circuit of the electronic equipment burned, which will invalidate or damage the equipment permanently. When the EMP produced by nucleus bomb is used strategically, it will be of great harm to the electronic equipment of the enemies. Similarly, when the HPM weapons are used in the tactics, it will be possible to strike the enemies accurately, so that hostile party's command and communication system will be paralyzed, which plays an important role in the modern high technology battlefield [4, 5].

3. EMP PROTECTION AND ITS DEVELOPMENT

With the rapid development and extensive application of the high technology, a variety of advanced electronic technology weapon systems present themselves to the battlefield of earth, sea, outer space, sky and the electromagnetism. It's known that the key components of the high technology weapons are the advanced computer, the integrated semiconductor circuit and the software systems controlled by feebleness electric current. Generally, the more integrated circuits the equipment uses, the flimsier it will be, and it will be more sensitive to the variance in the electromagnetic field of the environment. Therefore, it's of great importance to control the electromagnetic environment of the system, whose critical factor is EMP protection.

The power countries like the USA and Russia laid much weight on the EMP protection of weapon equipment. The experiment of effect that caused by EMP to microelectronic circuit and

the technology of EMP protection were achieved in the USA in the last years of the 20th century, new material and structure were investigated to adjust and control the equipment under the complex electromagnetic environment. As a result, the electromagnetic interference between systems was decreased and the system's ability to withstand the EMP strike was strengthened. Simultaneously, the military standard of the system was established [3, 4]. In China, the ability of the electronic and information system to defense the HPM weapon under complex electromagnetic environment is so weak that it's urgently necessary to pay more emphasis on the research. However, the research on HPM and the protection from it in China still has a long way to go to catch up with the developed countries. We still have a large number of work like new ideas, new mechanism, new material and new structures to do, as well as the new military standard is to be established. It's a new question for discussion on how to carry out effective protection for our electromagnetic space.

4. NEW TECHNOLOGY OF ELECTROMAGNETIC PROTECTION

Intelligent integration weapons are developed in order to meet the demands of war under the high technology. At the same time, the complex electromagnetic environment requires the electronic equipments must have a good ability to anti-electromagnetic interference, by which to ensure their safety, reliability and stability. Therefore, the development of new technologies protecting the equipments from HPM is increasingly important and urgent, application research on a variety of new materials and structures is imperative [1, 2, 6].

4.1. New Structure

4.1.1. The New Sub-wavelength Artificial Electromagnetic Structure (SWAEMS)

The new SWAEMS breaks the traditional design philosophy which is used in the current physical protection. Under the condition that no disturbance is made to the wave's amplitude and phase, this new structure can distort the trace of the electromagnetic waves smoothly, and at the same time do not produce any electromagnetic interference to the internal region that is surrounded by it. Hence the isolation between the protected region and the out electromagnetic environment is realized, which will protect the weapons from attacks and destruction of HPM weapons.

Based on the left-handed materials of quasi-periodic SWAEMS and guided by the theory of Maxwell's equations in the form of invariance under the coordinate transformation, electromagnetic parameters' distribution of the expected structure in the space was given by Smith in the United States and Pendry in the United Kingdom. Smith's Study Group given the narrowband artificial electromagnetic structures in 2006, published in Science, as shown in Figure 1 [7].

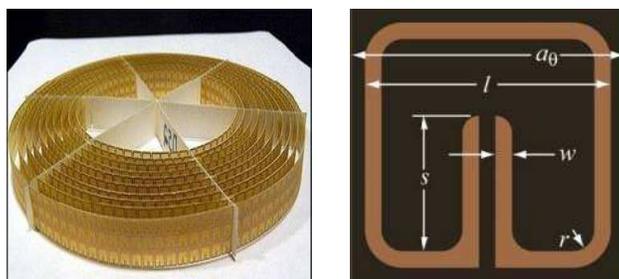


Figure 1: Entities of SWAEMS and its structural unit.

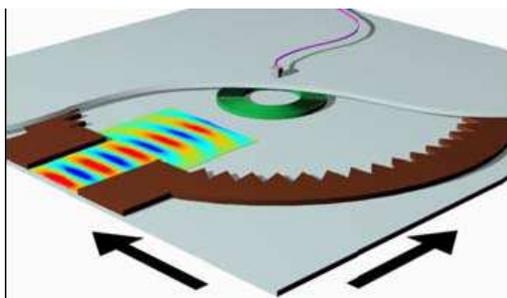


Figure 2: The measurements of SWAEMS.

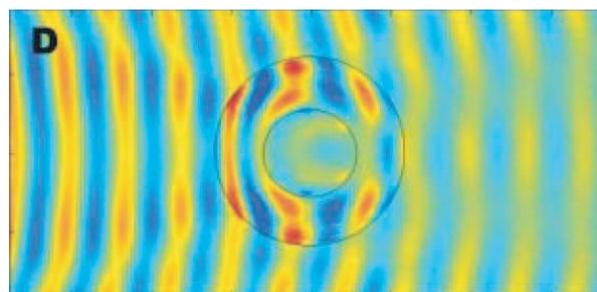


Figure 3: Field distribution around the SWAEMS.

Figure 2 shows a cross-section map of the structure irradiated by an open-ended waveguide plane wave. In the middle of the device is the objective to be protected. Test probes are installed on the top of the device to measure the amplitude and phase of the electric field that can be got by move the device along the black arrow step by step. The distribution of the field is given in Figure 3, from which we can see that when the wave pass through this region, it produce no reflection but maintained its original state, the protected region is bypassed and a beautiful banana ball curve was formed in the region [7].

Actually, the artificial electromagnetic structure is a closed object. It's difficult to establish a physical contact between the internal and outside of it. Simultaneously, the property parameter of the structure is too complex. In addition, the study of SWAEMS were mostly concentrated on two-dimensional structure, the study of three-dimensional structure, which is of more significance, still remain theoretical. Therefore, the achievement of three-dimensional magnetic structures and the electromagnetic protection mechanism should be paid more attention.

4.1.2. Frequency Selective Surface Structure (FSS)

With a two-dimensional periodic structure, FSS is a kind of surface whose reflection and transmission characteristics could be expressed as a function of frequency, for which it has a characteristic of selective to passed electromagnetic wave. It consists of metallic patch unit arranged periodically or aperture unit arranged periodically on the metal screen. When the structure is working on the resonant frequency of the unit, characteristics of whole transmission (aperture type) or total reflection (SMD type) will come into being. That's to say, FSS is a filter to the incident electromagnetic wave.

As Shown in Figure 4, FSS unit has a variety of shapes, such as the “+”-shaped unit, quadrilateral unit, round ring unit, “Y”-shaped unit and hexagonal unit. It's of great importance to select the shape of FSS. A well-designed FSS not only has a relatively stable resonant frequency for the large incident angle and a good transmission character for all kinds of polarization wave, but also is conducive to obtaining a wide bandwidth.

The FSS periodic array is so thin that the support of dielectric materials is needed to obtain sufficient strength. In addition, the resonant frequency will be more stable when the dielectric material is loaded. The characteristic of wide band and drastically cut-off frequency can be obtained by connecting Multi-layer FSS. Figure 5 shows a FSS with two layers “Y” shaped unit embedded in five layers of dielectric materials, and its *S*-parameters is given simultaneously.

FSS is mainly used as a radome, by which the performance can be enhanced or splendid design can be obtained. In the technology of target obscure in radar, the radome made by FSS not only can guarantee that the antenna's operating characteristics in-band are not affected, but also can inhibit electromagnetic waves out of band to achieve the purpose of hiding the antennas. It can also be used as a frequency division surface, separating electromagnetic waves containing different frequency signals, that is, one of the frequencies is selected while the other is abandoned.

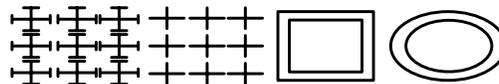


Figure 4: Typical unit geometry of FSS.

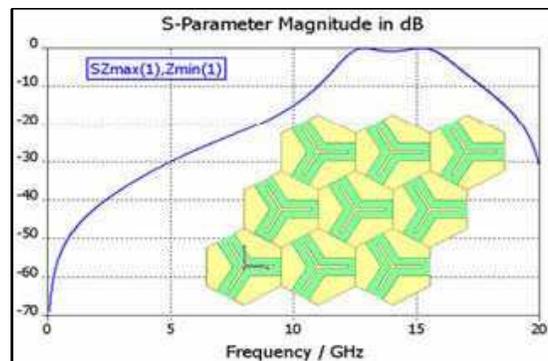


Figure 5: “Y”-shaped FSS with band-pass characteristics.

4.2. New Materials

4.2.1. Radar Absorbing Coating (RAC) with Wideband

The RAC has always been popular with all powered countries for its inexpensive, and widely been used in the radiation protection of various of electronic products, equipment and systems. Absorber plays an important part in the contest of research on technology of protection from electromagnetic wave between the powered countries. According to the record, 80% of the shielding methods is using absorber, which is the main body that determine the absorbing properties of materials and directly restricts the development level and engineering application of absorbing materials. At present, absorb that is well researched includes ultra fine ferrite absorber, metal and oxide ultra-fine powders absorber, black conductive carbon absorber and carbonyl iron absorber.

Ultra fine ferrite absorber's absorption capacity is so good that even in the conditions that low-frequency, small thickness it still maintaining very good absorption properties, its relative permeability is much larger and relative dielectric constant is much smaller. The Nano-composite absorber consisted of Ferrite microwave absorbers and conductive polymer will be a promising new type of shielding material.

The real and imaginary parts of ultra fine nickel powder absorber can be adjusted in a certain range through different ratio of particle size. The magnetic transition temperature of this material is so high that the performance of wave-absorbing is stable in a large range of temperature; there is almost no the possibility of deterioration. This metal powder is commonly used as packing material added to the binder. The performance of shielding and absorption are greatly affected by the distribution of the metal powder in the blinder, the more symmetrical the powder distributes, the easier the conductive channels forms, and the easier to advance the conductivity of the coating surface and to ameliorate the performance of shielding.

Conductive carbon black absorber is one of the more commonly used materials in microwave absorbent. Compared with the metal and organic absorbing materials, it has less weight, more predominant performance of anti-oxidation and more excellent chemical stability.

4.2.2. Novel Voltage-controlled Conductive Material (VCCM)

The existing protective materials can not meet the demands that the weapons should send and receive normal signals effectively and at the same time avoid the EMP weapons' attack in the complex electromagnetic environment, thus, a new material with the capability of selection to signals, that is, the signals that will do damage to the receiver will be attenuated and the safe ones will be allowed to pass, is greatly needed. This demand could be satisfied by a new material called VCCM.

VCCM is similar to medium in low-power electromagnetic irradiation, while it is similar to metal in the high-power electromagnetic irradiation. In other words, this material is not conductive under normal circumstances, but may be conductive in the strong electromagnetic field. This material has a turn-on voltage threshold. Its property will be suddenly changed when the external electromagnetic fields or voltage is greater than the threshold, and the reaction to the electromagnetic wave is changed from transmission to shielding.

VCCM has been researched by the Characteristics of Target Environment and EMC Center in National University of Defense Technology and some results have been achieved. The performance of shielding and transmission of current design are shown in Figure 6, from which we can see that in the frequency band of 1 GHz–15 GHz, when the VCCM is conductive, the attenuation is greater

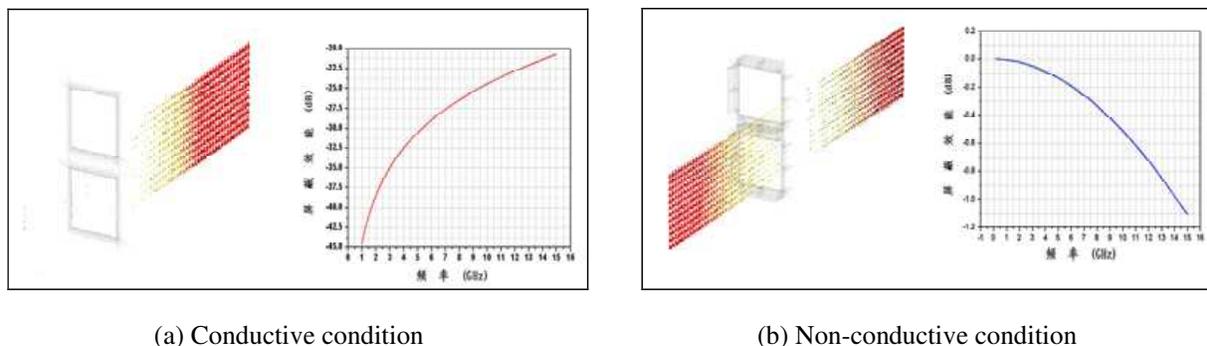


Figure 6: Shielding efficient of VCCM.

than 20 dB that the EMP is effectively shielded; while it is not conductive, the attenuation is less than 1.5 dB, the Useful electromagnetic signals can thrill through almost without attenuation.

The property of the VCCM has no relationship with the frequency. That's to say, it has a wide range of spectrum, for which it's an excellent material to be used to protect the equipments from high power EMP. Nowadays, the study of novel VCCM has achieved certain results that can be easily translate into actual products and be widely used in the protection of crucial receive equipments, communication systems and other large venues like command center.

5. CONCLUSION

HPM weapon is one of the new menace in the battlefield of the future, and EMP protection is one of the requisite instrument that protect the government, the military center, the national finance, the insurance, the telecommunication and the manufacture of precision electronic equipment from damage. Hence, it's necessary for us to put more bankroll and energy on the development of EMP weapon and its protection, for which the EMP protection should be planed as a whole and carried out, the protection of personnel and weapon equipments from EMP pulse, should be executed. By clearing up the stochastic electromagnetic interference and depressing the noise, the protection of equipments from super high EMP, especially HPM weapon's strike can come to be reality; the ability of weapons' protection, subsistence and application under complex electromagnetic environment could be enhanced. Consequently, the damage of EMP to people and equipments can be decreased to the fullest extent.

REFERENCES

1. Zhou, B., B. Chen, and L. Shi, *EMP and Its Engineering Protection*, The Publisher of Industry of National Defense, 2003.
2. Tan, Z., S. Liu, and G. Wei, *EMP and the Technology of Protection*, 103–105, The Publisher of College of Ordnance Projection, 1999.
3. Zhang, Z., J. Zhang, and C. Yin, *HPM weapons and Its Application*, The Publisher of the PLA, 2003
4. Celozzi, S. and M. Feliziani, "EMP coupling to twisted wire cables," *IEEE International Symposium on Electromagnetic Compatibility Symposium Record*, 1990.
5. Taylor, C. D. and C. W. Harrison, Jr., "On the coupling of microwave radiation to wire structures," *IEEE Transactions on Electromagnetic Compatibility*, 1992.
6. Lai, Z., *Protection of Electromagnetic Interference and EMC*, The Publisher of Atomic Energy, 1993.
7. Schurig, D., J. J. Mock, B. J. Justice, S. A. Cummer, J. B. Pendry, A. F. Starr, and D. R. Smith, "Metamaterial electromagnetic cloak at microwave frequencies," *Science*, Vol. 1133628, No. 10, 1126, 2006.

A Novel Hybrid Method for Solving the Response of Non-uniform Transmission Line Network

Yujian Qin, Peiguo Liu, and Jianguo He

School of Electronic Science and Engineering, NUDT, China

Abstract— For solving the responses of transmission line networks which include non-uniform parts, a novel hybrid method based on Baum-Liu-Tesche (BLT) equation and Finite Different Time Domain method (FDTD) is proposed. FDTD is applied for calculating the S -parameters of the non-uniform parts, and these parts are treated as junctions in BLT equation. Finally the response of the whole network will be solved by BLT equation. Numerical results illustrate the efficacy of the proposed method.

1. INTRODUCTION

In the analysis of transmission line networks on complex systems (automobile, aircraft and warships, etc.), the most appropriate method is BLT equation, which was invented by Baum, Liu and Tesche in 1978 [1]. However, BLT equation is inapplicable in non-uniform circumstances. In practice, non-uniform transmission lines are unavoidable. Therefore, methods in time domain are appropriate rather than methods in frequency domain, such as Finite Difference Time Domain (FDTD) [2]. But FDTD is more time-consuming than BLT equation. For example, the length of cables in “Typhoon” fighter is 30 km; and that in Boeing 747 is 274 km, if the excitation is a nuclear electromagnetic pulse (NEMP), the spatial mesh will be achieve the level of 10^6 , the computational cost will be high. Not only the system-level cable analysis, the PCB-level analysis has the same problems, non-uniform transmission lines are everywhere, such as via holes, corners, etc.

If a method could calculate the responses of large-scale transmission line networks rapidly, and could handle the non-uniform lines, simultaneously; the practical analysis requirement can be satisfied primely. Based on that, we combine the advantages of BLT equation and FDTD method, propose a BLT-FDTD hybrid method. While maintaining the precision, this method could maximize the computational efficiency.

2. BLT EQUATION

BLT equation is based on Electromagnetic Topology (EMT) [3]. Tubes and Junctions are used to represent transmission lines and terminals. The relationship of all the tubes and junctions in the network is [4]

$$\bar{\mathbf{V}} = [\bar{\mathbf{I}} + \bar{\mathbf{S}}] \cdot [\bar{\mathbf{\Gamma}} - \bar{\mathbf{S}}]^{-1} \cdot \bar{\mathbf{V}}_S \quad (1)$$

where, $\bar{\mathbf{V}}$ is the total voltage vector. $\bar{\mathbf{V}}_S$ is the source vector. $\bar{\mathbf{I}}$ is the unit super-matrix. $\bar{\mathbf{\Gamma}}$ is the propagation super-matrix, denotes the propagation functions of all the tubes, if there are N tubes, the rank of $\bar{\mathbf{\Gamma}}$ is $2N \times 2N$; and it's the block diagonal matrix, each sub-matrix denotes the corresponding tube's propagation function. $\bar{\mathbf{S}}$ is the scattering super-matrix, as shown in Eq. (2) (where $\bar{\mathbf{Z}}_c$ denotes the characteristic impedance super-matrix for all the tubes, $\bar{\mathbf{Y}}$ denotes the admittance super-matrix for all the junctions); it is sparse, but not necessarily block diagonal, this depends on the connection status of junctions.

$$\bar{\mathbf{S}} = [\bar{\mathbf{I}} - \bar{\mathbf{Z}}_c \times \bar{\mathbf{Y}}] [\bar{\mathbf{I}} + \bar{\mathbf{Z}}_c \times \bar{\mathbf{Y}}]^{-1} \quad (2)$$

Equation (1) is the expression of the classical BLT equation.

3. FDTD METHOD

Based on Kirchhoff's law, and utilizing the first-order central difference formula, the Telegrapher's Equations for multiconductor transmission lines can be expressed as

$$\begin{cases} \mathbf{U}_n^m = \left(\frac{\Delta x}{\Delta t} \mathbf{C}' + \frac{\Delta x}{2} \mathbf{G}'\right)^{-1} \left[\left(\frac{\Delta x}{\Delta t} \mathbf{C}' - \frac{\Delta x}{2} \mathbf{G}'\right) \mathbf{U}_n^{m-1} + \mathbf{I}_{n-1}^{m-1} - \mathbf{I}_n^{m-1}\right] & 1 < n < N + 1 \\ \mathbf{I}_n^m = \left(\frac{\Delta x}{\Delta t} \mathbf{L}' + \frac{\Delta x}{2} \mathbf{R}'\right)^{-1} \left[\left(\frac{\Delta x}{\Delta t} \mathbf{L}' - \frac{\Delta x}{2} \mathbf{R}'\right) \mathbf{I}_n^{m-1} + \mathbf{U}_n^m - \mathbf{U}_{n+1}^m\right] & 1 \leq n < N + 1 \\ \mathbf{U}_1^m = \left(\frac{\Delta x}{\Delta t} \mathbf{Z}_0 \mathbf{C}' + \mathbf{I}\right)^{-1} \left[\left(\frac{\Delta x}{\Delta t} \mathbf{Z}_0 \mathbf{C}' - \mathbf{I}\right) \mathbf{U}_1^{m-1} + 2\mathbf{V}_s^{m-1} - 2\mathbf{Z}_0 \mathbf{I}_1^{m-1}\right] \\ \mathbf{U}_{N+1}^m = \left(\frac{\Delta x}{\Delta t} \mathbf{Z}_L \mathbf{C}' + \mathbf{I}\right)^{-1} \left[\left(\frac{\Delta x}{\Delta t} \mathbf{Z}_L \mathbf{C}' - \mathbf{I}\right) \mathbf{U}_{N+1}^{m-1} + 2\mathbf{Z}_L \mathbf{I}_N^{m-1}\right] \end{cases} \quad (3)$$

where, \mathbf{U}_n^m denotes the voltage at n spatial mesh cell and m temporal mesh cell; \mathbf{I}_n^m denotes the current at n spatial mesh cell and m temporal mesh cell; \mathbf{I} denotes the unit matrix; \mathbf{Z}_0 and \mathbf{Z}_L denotes the impedance matrix of the two terminals, respectively; Δx and Δt denotes the spatial step and temporal step, respectively.

The spatial step Δx and the temporal step Δt must check the stability condition: $\Delta t \leq \frac{\Delta x}{v_p}$. Where v_p is the highest modal velocity.

4. BLT-FDTD HYBRID METHOD

It is complicated and difficult, using methods in frequency domain, to obtain the propagation characteristics of non-uniform transmission lines. It is very convenient to solve out by methods in time domain. Therefore, if we separate the non-uniform parts, and solve it by a time domain method, then integrate them into a frequency domain method; the responses of transmission line networks which include non-uniform parts could be resolved out.

The process of BLT-FDTD hybrid method is: (1) Calculating the S parameters of non-uniform parts by FDTD; (2) Treat them as the junctions in BLT equation, and solve the whole network.

4.1. S Parameters of Non-uniform Parts

As shown in Figure 1, there is a single non-uniform line over a PEC ground plane, and it is separated form a transmission line network. The length is L , and this is a two port microwave network.

The line is discretized into M parts, the length of each part is Δx . The characteristic impedance of part m , Z_{cm} , can be obtained easily. Connecting two fictional impedances, Z_{L1} , Z_{L2} , to the two ports, respectively, and

$$\begin{cases} Z_{L1} = Z_{c1} \\ Z_{L2} = Z_{cM} \end{cases} \quad (4)$$

where, Z_{c1} is the characteristic impedance of the discrete part connected to port 1; Z_{cM} is the characteristic impedance of the discrete part connected to port 2.

Adding a perfect voltage source at port 1, the loads voltage responses v_{L1} , v_{L2} and current responses i_{L1} , i_{L2} can be solved out by FDTD.

Using FFT, we could obtain \tilde{V}_{L1} , \tilde{V}_{L2} and \tilde{I}_{L1} , \tilde{I}_{L2} , which are the Fourier transform of v_{L1} , v_{L2} and i_{L1} , i_{L2} . Defining a_1 , b_1 is the normalized ingoing wave and outgoing wave of port 1, respectively; a_2 , b_2 is the normalized ingoing wave and outgoing wave of port 2, respectively; and according to

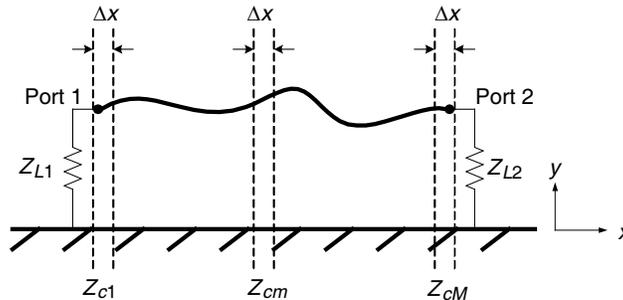


Figure 1: Non-uniform transmission line.

the definition of S parameters, separating the ingoing wave and the outgoing wave, we obtain

$$\begin{aligned} a_1 &= \frac{1}{2} \left(\frac{\tilde{V}_{L1}}{\sqrt{Z_{c1}}} + \tilde{I}_{L1} \sqrt{Z_{c1}} \right) \\ b_1 &= \frac{1}{2} \left(\frac{\tilde{V}_{L1}}{\sqrt{Z_{c1}}} - \tilde{I}_{L1} \sqrt{Z_{c1}} \right) \\ b_2 &= \frac{1}{2} \left(\frac{\tilde{V}_{L2}}{\sqrt{Z_{cM}}} + \tilde{I}_{L2} \sqrt{Z_{cM}} \right) \end{aligned} \quad (5)$$

And then s_{11} (reflective coefficient of port 1) and s_{21} (propagation coefficient from port 1 to port 2) could be obtained

$$\begin{aligned} s_{11} &= \frac{b_1}{a_1} \Big|_{a_2=0} \\ s_{21} &= \frac{b_2}{a_1} \Big|_{a_2=0} \end{aligned} \quad (6)$$

In a similar way, we can obtain

$$\begin{aligned} a_2 &= \frac{1}{2} \left(\frac{\tilde{V}_{L2}}{\sqrt{Z_{cM}}} - \tilde{I}_{L2} \sqrt{Z_{cM}} \right) \\ b_1 &= \frac{1}{2} \left(\frac{\tilde{V}_{L1}}{\sqrt{Z_{c1}}} - \tilde{I}_{L1} \sqrt{Z_{c1}} \right) \\ b_2 &= \frac{1}{2} \left(\frac{\tilde{V}_{L2}}{\sqrt{Z_{cM}}} + \tilde{I}_{L2} \sqrt{Z_{cM}} \right) \end{aligned} \quad (7)$$

And then s_{22} (reflective coefficient of port 2) and s_{12} (propagation coefficient from port 2 to port 1) are

$$\begin{aligned} s_{22} &= \frac{b_2}{a_2} \Big|_{a_1=0} \\ s_{12} &= \frac{b_1}{a_2} \Big|_{a_1=0} \end{aligned} \quad (8)$$

Combining Eqs. (5)–(8), we can obtain the S -parameters of the non-uniform transmission line:

$$\mathbf{S} = \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix} \quad (9)$$

4.2. Response of the Whole Network

In BLT equation, the influence of a junction is embodied by the scattering parameter S . Therefore, putting Eq. (9) into Eq. (1), then solving BLT equation by analytical means or by numerical means; the response of the whole network will be solved out.

5. NUMERICAL RESULTS

5.1. Uniform Transmission Line

First, we use a simple example to validate the proposed method. Suppose a single line over the PEC ground plane, the radius of the line is 1.5 mm, the height over the plane is 0.1 m; the length of the line is 30 m. A voltage source with 50Ω internal resistance lies in one side, and a resistance of 100Ω lies in the other side. The excitation form is double exponential pulse [5], which is defined by Eq. (10).

$$V(t) = 1.05 \left(e^{-4 \times 10^6 t} - e^{-4.76 \times 10^8 t} \right) \quad (10)$$

From the source end, this line is divided into three parts, with the length of 6 m, 15 m and 9 m; the middle part is solved by FDTD, the others is solved by BLT equation.

Figure 2 shows the transient load response calculated by BLT-FDTD hybrid method and full FDTD, the latter is treated as reference. Figure 3 shows the frequency load response calculated by

BLT-FDTD hybrid method and full BLT, the latter is treated as reference. Obviously, no matter the transient or the frequency response, the accuracy of the proposed method is extremely high. This shows the proposed method is fit for the uniform conditions.

5.2. A Transmission Line with a Non-uniform Part

Suppose a single line over the PEC ground plane, a voltage source with $50\ \Omega$ internal resistance lies in one side, and a resistance of $100\ \Omega$ lies in the other side. the radius of the line is $1\ \text{mm}$, the height over the plane of the source end is $0.1\ \text{m}$, the height over the plane of the load end is $1\ \text{m}$; the structure is shown in Figure 4. The excitation form is Gaussian pulse, which is defined by Eq. (11).

$$V(t) = e^{-\frac{4\pi(t-t_0)^2}{T^2}} \tag{11}$$

From the source end, this line is divided into three parts, with the length of $6\ \text{m}$, $15\ \text{m}$ and $9\ \text{m}$; the middle part is solved by FDTD, the others is solved by BLT equation. Figure 4 shows the transient load response calculated by BLT-FDTD hybrid method and full FDTD, the latter is treated as reference. Because the mismatch of the load, energy oscillates on the line, the peaks on Figure 4 show the phenomena. Among the peaks, there are some small oscillations, which are caused by the non-uniformity. The two results are in good agreement. This indicates the proposed method is fit for the non-uniform conditions.

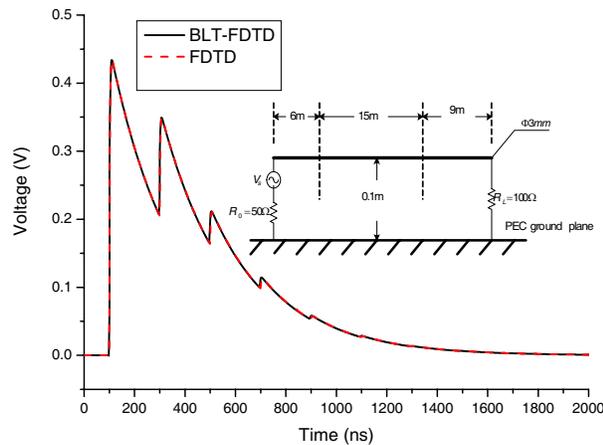


Figure 2: Transient response of the $100\ \Omega$ load of the uniform line.

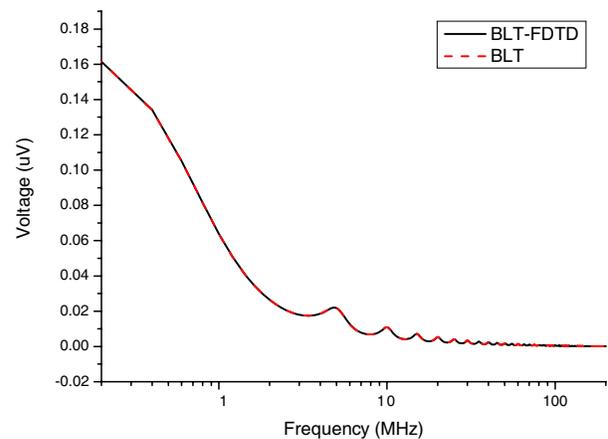


Figure 3: Frequency response of the $100\ \Omega$ load of the uniform line.

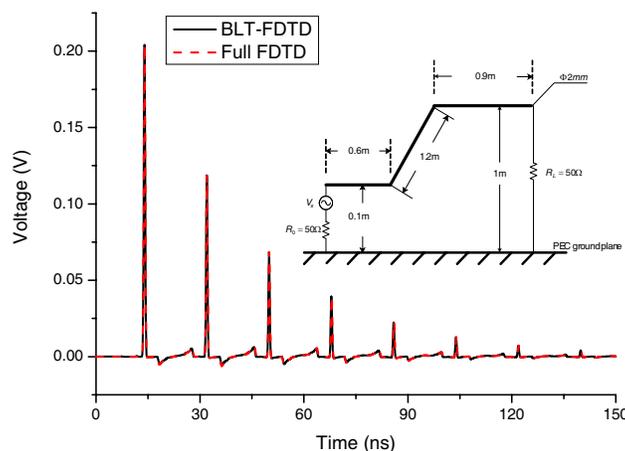


Figure 4: Transient response of the $100\ \Omega$ load of the non-uniform line.

6. CONCLUSION

To solve the transmission line network with non-uniform parts, a hybrid method from frequency domain and time domain, BLT-FDTD method, is proposed. FDTD is applied to calculate the S parameters of the non-uniform parts; then these parts are treated as junctions in BLT equation; and solving BLT equation, the responses of the network could be done. The proposed method overcomes the inapplicability of the methods in frequency domain and avoids the high computational cost of the methods in time domain. It has the capability for solving complex transmission line networks.

REFERENCES

1. Baum, C. E., T. K. Liu, and F. M. Tesche, "On the analysis of general multi-conductor transmission — Line networks," *Interaction Notes*, Vol. 350, 1978.
2. Mimouni, A., et al., "Lighting-induced overvoltages on overhead lines: Modelling and experimental validation," *J. Electrical. Engineering*, Vol. 58, No. 3, 146–151, 2007.
3. Tesche, F. M., "Topological concepts for internal EMP interaction," *IEEE Trans. EMC*, Vol. 20, No. 1, 60–64, 1978.
4. Qin, Y.-J., D.-M. Zhou, and J.-G. He, "Crosstalk analysis of arbitrary layout transmission line using BLT equation," *Journal of National University of Defense Technology*, Vol. 31, No. 2, 55–58, 2009.
5. Tesche, F. M., et al., *EMC Analysis Methods and Computational Models*, John Wiley and Sons, New York, 1997.

Solving Method for Electromagnetic Pulse Propagation Based on Combination of EMT and TDIE

Gaosheng Li, Yujian Qin, Peiguo Liu, and Jianguo He

College of Electronic Science and Engineering, National University of Defense Technology
Changsha 410073, China

Abstract— A solving method for propagation and coupling of Electromagnetic Pulse (EMP) is provided. The idea is to combine the theory of Electromagnetic Topology (EMT) and the method of Time Domain Integral Equation (TDIE). The former offers methodology and the latter carries out the computations, which can improve the efficiency of analysis and the size of area that can be computed. By this way, the difficulty of direct computation of complex electronics system can be conquered, and virtues of time domain computing can be used adequately. Principal ideas of EMT for analyzing and solving problems are introduced, after which the four main steps of using EMT to find solutions of propagation and coupling of EMP are summarized. Methods to combine EMT and TDIE are put forward, together with the advantages of TDIE for analyzing EMP. Finally, to improve computing efficiency of TDIE, relative speeding algorithms are expatiated.

1. INTRODUCTION

Both military and civil electromagnetic radiators have increased sharply since several years ago, resulting in the complexity of electromagnetic environment. Electronic systems are in face of intentional or inadvertently interferences. The electromagnetic pulses with low power may influence the normal operation of the system, while pulses with high power may damage the electronic system. It's incident field can bring voltage or current on the susceptible components when it enters the system through apertures, holes, power cords or signals cables. Consequently, the error rates may increase and the efficiency may decrease, and even the system may be destroyed.

Investigation of propagation characters of electromagnetic pulse, as well as coupling features of the pulses to susceptible systems, is the basis for grasping of electromagnetic interference and designing of electromagnetic compatibility. It will help for evaluation of complexity of electromagnetic environment and then benefit the tactic against the complex environment.

Difficulties may be encountered when computing electronic large targets and areas, the most of which are the large amount of dissect grids and the demand of the operation for computer software as well as the hardware. This will often result in the several days of computing time or even exceed the level of computer technology nowadays. A series of problems arise in this situation, that is, how to describe the electronic system in complex electromagnetic environment? How to describe the reciprocity between electromagnetic environment and electronic system? How to avoid large amount of repeated computation when the only part of environment changes? How to manage various computing methods reasonably to analyze problems in reality?

This brings the requirement for methodology. Electromagnet topology is used to guide the computing process of propagation and coupling of electromagnetic pulse. Reasonable layering is carried out on computing targets, and proper method will be chosen and pertinence settings will be undertaken before the computing.

2. PRINCIPAL IDEA OF ELECTROMAGNETIC TOPOLOGY AND ITS IMPLEMENTATION PROCEDURES

In 1970's, Professor Carl E. Baum from AFRL (Airforce Research Lab of USA) put forward the concept of electromagnetic for the first time [1]. The principal idea of it is to make use of theory of topology and the methods of it, and divide the research object into many areas. Contact among the areas is realized by topology structures. Integration is attained based on separate researches, thus, the whole complex electromagnetic coupling problem is decomposed into fairly simple problems, then analysis and evaluation of performance of devices in the electromagnetic environment [2].

Flow of analyzing problems with method of electromagnetic topology mainly consists of the following four steps.

2.1. To Carry on Topology Decomposition of the System

Here, we find out all the points which may have relations with outside environment in the system, and determine the range, and then decompose the system. All the outside electromagnetic energy that wants to bring interference on certain components in the system must propagate through multiple layers of shielding of the system. We define transmission functions for each shielding surface to describe the penetration of energy.

To describe the topology structure of the system, we can delaminate and fix number for the shielding surfaces. Simple systems can use signal subscript to fix number. Real systems (for example, airplanes and ships) usually have shielding surfaces with different structural order and types. Some shielding surfaces appear to be alveolate (for example, comparting cabins in ships), and some embed in shielding surfaces of the system and compart from each other (for example, cases of electronic equipments). So, the second subscript is introduced. For example, V_{ik} means the k th child space of the i th level. This can be found in Figure 1.

2.2. Conforming Sequence Graphs of Interactions

Set paths of the outside electromagnetic environment coupling into the system to be sides of the sequence graphs of interactions, and the volumes that correlate with each other as the nodes, thus the sequence graphs can be determined.

Topology leveling of the system will divide the energy coupling, take Figure 2 as an example, outside energy may couple to V_3 through $s_1s_2s_3$; Or it may penetrate s_1 first, and then couples to V_3 directly by cables. The former is called couple of order three, and the latter, couple of order two. Paths with different orders may bring different coupling effects. Difference of orders and real coupling paths will influence the shielding effects.

Transmitting function is numbered $T_{ikik'}$ for the one from V_{ik} to $V_{ik'}$. If the transmitting functions are not single, that is, there is more than one paths, then superscripts can be used to differentiate them.

2.3. Solving Child Problems

For problems after topology decomposition, we often use traditional computing methods of electromagnetic field, together with excitations and boundary conditions, to attain the solutions. Moreover, solutions of child problems can be achieved by measurement and computation of parse or experience equations or other non-numerical methods. Typical research achievements of coupling paths and experimental results of transmission functions of coupling paths are all the base for solving of child problems.

2.4. Integration of System Responses

Solving of child problems and integration of system response can be carried out at the same time, or be performed separately. Combination of results of child problems will achieve system response, and the responses of the system will finally fall onto certain devices or components inside the system.

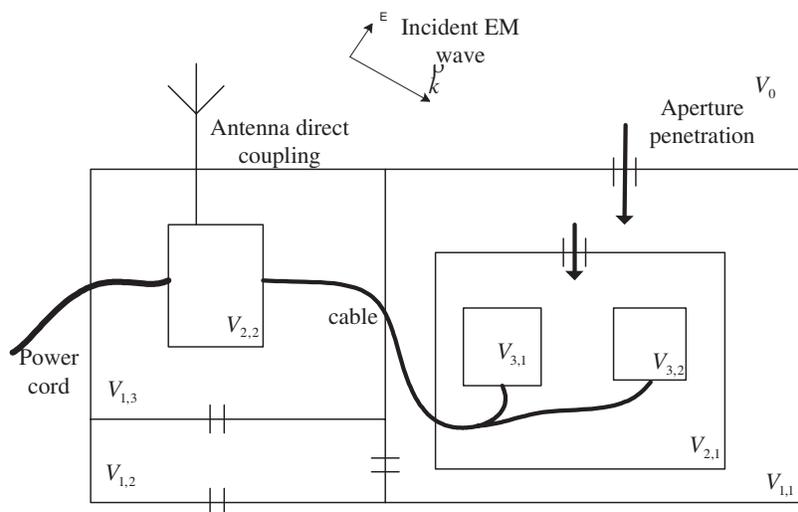


Figure 1: Example of topology structure.

Take the response of $V_{3,1}$ in Figure 2 as an example, let's suppose the couple paths after approximate remain only $T_{0,1;2,2} + T_{2,2;3,1}$, that is, energy from outside come to $V_{2,2}$ from the antenna, and then couples to $V_{3,1}$ by cables. This can be divided into two child problems, $T_{0,1;2,2}$ and $T_{2,2;3,1}$. The former uses numerical algorithms of electromagnetic field to model the antenna and outside surface of the system, considering the transmitting loss of cables, to compute response of $V_{2,2}$. Then, $V_{2,2}$ will be looked as a network with two ports, as one of the terminals of the transmission line, to achieve its scattering parameters by experimental or numerical methods. After that, we take the former results as an excitation, to attain response of $V_{3,1}$ with method of transmission line.

3. TIME DOMAIN INTEGRAL EQUATION AND ITS SPEEDING METHODS

Theory of electromagnetic topology provides methodology for analyzing of propagation characters of electromagnetic pulse. When the topology models sequence graphs of interaction have been built, the most important is to solve the child problems. Numerical algorithms of electromagnetic field are credible tools [3].

TDIE has advantages in analyzing propagation and coupling characters of electromagnetic pulses. It builds time domain integral equations based on Green functions and boundary conditions of the problem, and disperses integral areas of space variables and time variables. Linear equations are achieved from integral equations through matching in both space domain and time domain. One of the advantages of TDIE is that there is no need of setting boundary conditions manually.

MOT (Marching-on in-Time) is usually used when solving the integral equations. Flow of MOT consists of modeling, TDIE dispersing, solving of equations and post processing. Cost of the computation is high, and large amount of memory and CPU time are needed. Speeding methods of TDIE have tangible request.

PWTD (Plane Wave Time Domain) is widely used in speeding of MOT. The idea of PWTD is to unwrap vector potential to plane wave, and unite adjacent current elements into a unit. The direct reciprocity among elements is translated into reciprocity among units, thus the complexity of multiply of matrix vector can be decreased. Sources in the same units build their influences on others through three processes, that is, congregation, transformation and casting, which can avoid the trivial computing of historical interactions among each base function in traditional MOT [4].

TD-AIM (Time Domain Adaptive Integral Method). Fields of areas which are far from the source can be computed by simplifying distribution information of the source. Meeting the standard of fixed far field, we transform the interactions of fields and source base functions into interactions among regular grid currents. Then FFT and IFFT are used to increase the computation efficiency [5].

Mixed algorithms of PO-TDIE (Physical Optics-TDIE). When dividing the areas, we use the

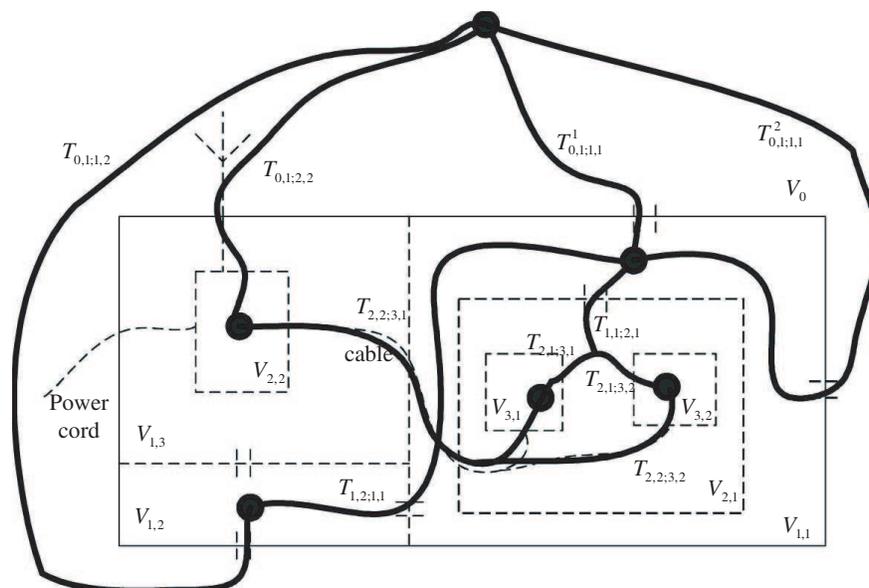


Figure 2: Example of coupling paths.

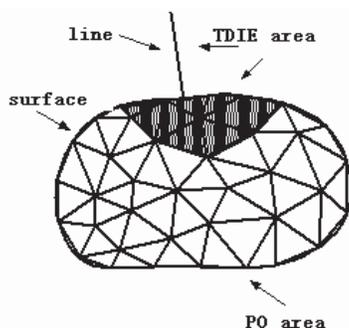


Figure 3: Sketch of PO-TDIE.

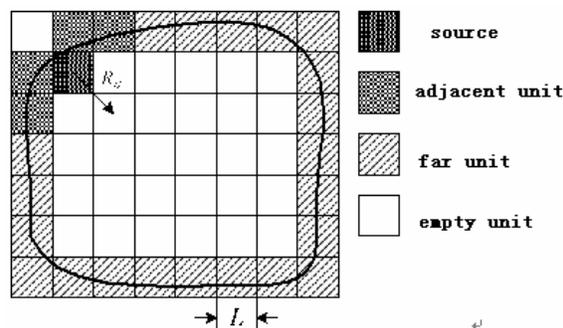


Figure 4: Sketch of UTD-TDIE.

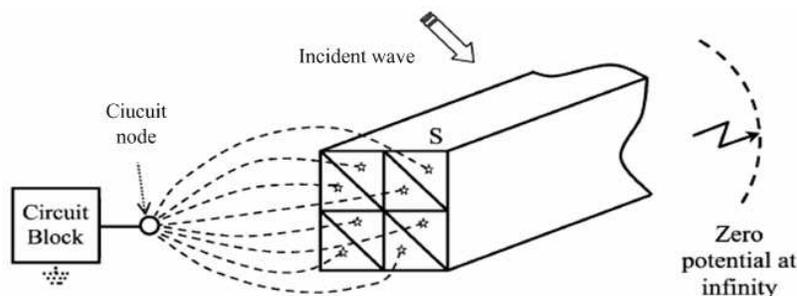


Figure 5: Sketch of TDCIE.

PO algorithm only on surface targets, and the line targets are all divided into TDIE areas. Regions that change sharply will be divided as TDIE areas, and smooth surfaces will be divided as PO areas. Triangle plane elements in common boundaries of PO and TDIE areas will be used by both PO and TDIE base functions. For assurance of precision, internal sides of common boundaries are all divided as TDIE areas. Reduction of computing amounts is embodied in the ignorance of the mutual-resistance of sources in PO areas.

UTD-TDIE (Uniform Geometrical Theory of Diffraction-TDIE). The main computation work focuses on the influence of the historical equivalent sources on current field, which is called sluggish integration. The areas needed to be computed in sluggish integration include live areas, adjacent areas, reflection areas and diffraction areas. Examples of those areas can be found in Figure 4.

TDCIE (Time Domain Coupling Integral Equation). Algorithms introduced in the above are mainly aimed at metal materials. In the field of combination of metal and dielectric, according to the mixed problems of interconnection of electromagnetic structures and circuits, we combine the TDIE in field analysis and MNA in circuit analysis to realize simultaneously simulation of coupling for complex electromagnetic structures and linear or nonlinear lumped circuits. Coupling current is introduced to use patulous equation of current continuity and generalized law of Kirchhoff in the coupling analysis of electromagnetic structures and circuits.

4. CONCLUSION

Theory of Electromagnetic Topology is an efficient methodology for analyzing and designing of electromagnetic compatibility and environment. Computing results proved that TDIE and its speeding methods is efficient for the solving of electromagnetic problems.

REFERENCES

1. Baum, C. E., "Electromagnetic topology: A formal approach to the analysis and design of complex electronic systems," Interaction Notes, IN-400, 1980.
2. Besnier, P., "Electromagnetic topology: An additional interaction sequence diagram for transmission line network analysis," *IEEE Trans. Electromagnetic Compatibility*, Vol. 48, No. 4, 685–692, 2006.
3. Chew, W. C., J. M. Jin, E. Michielssen, et al., *Fast and Efficient Algorithms in Computational Electromagnetics*, Artech House, London, 2001.

4. Ergin, A. A., B. Shanker, and E. Michielsen, "The plane wave time domain algorithm for the fast analysis of transient wave phenomena," *IEEE Ant. Prop. Mag.*, Vol. 41, 39–52, 1999.
5. Bleszynski, E., M. Bleszynski, and T. Jarozewicz, "AIM: Adaptive integral method for solving large-scale electromagnetic scattering and radiation problems," *Radio Sci.*, Vol. 31, 1225–1251, 1996.

Study of Sapphire Loaded H-Maser in Shanghai Observatory

Ke Dai^{1,2}, Wei Qun Zhang¹, Yan Jun Zhang¹, and Wen Ming Wang^{1,2}

¹Shanghai Astronomical Observatory, Chinese Academy of Sciences, Shanghai 200030, China

²Graduate School of Chinese Academy of Science, Beijing 100049, China

Abstract— Sapphire loaded cavity for active H-maser was designed to minimize the mass and volume of the traditional active H-maser. The ability of sapphire loaded cavity to achieve self-oscillation was calculated. Then we made the experiment for sapphire loaded cavity in traditional H-maser as bed. The signal was found with oscillation power of -102.93 dBm while the beam flux is about 1.1 mA. For the temperature coefficient of sapphire loaded cavity is too high, the way to compensate the temperature-coefficient was studied. The region of “zero temperature coefficient point” was discussed.

1. INTRODUCTION

Nowadays with the development of SLR, VLBI etc, H masers as the most stable frequency sources readily available were used abroad. However traditional active H masers are heavy and large in size. Can't be used as mobile instruments. Sapphire loaded cavity can solve this problem efficiently. It can reduce the size and weight of the traditional active H maser without degrading the excellent stability of the traditional masers. Many analyses and different design were reported [1–3]. For many years, ShangHai Observatory was studying on active and passive H masers. The study on sapphire loaded H maser has just begun. We solved the Maxwell's equation for TE₀₁₁ mode in sapphire loaded cavity calculated the Q , η' and temperature coefficient for the cavity, then we judged if it could achieve self-oscillation. With the sapphire loaded cavity we got a signal. To sapphire loaded cavity's high temperature coefficient problem we discussed the way to reduce the temperature coefficient. And we found the region to find “zero temperature coefficient”.

2. THEORY ANALYSIS AND SIMULATION

The structure of the sapphire loaded cavity is as Fig. 1. We can solve the following Eq. (1) to determine the dimensions of the cavity for 1.42 G in TE₀₁₁ mode. We need the optic axis of the sapphire crystal just parallel to Z axis. Then only the permittivity perpendicular to optic axis of the sapphire crystal could judge the frequency of the cavity for TE₀₁₁ mode. The permittivity of sapphire crystal which perpendicular to optic axis is 9.36 (note $\varepsilon_r = 9.36$). We can write \mathbf{a}/\mathbf{b} as ρ_1 . When the out side metal cavity size \mathbf{a} , \mathbf{h} was fixed, then by $\rho_1\mathbf{b}$ was fixed, For the cavity frequency was fixed, by Eq. (1) \mathbf{c} was fixed. The whole cavity dimension was fixed.

$$\frac{\gamma_2 [A_2 J_0(\gamma_2 b) + B_2 N_0(\gamma_2 b)]}{[A_2 J_1(\gamma_2 b) + B_2 N_1(\gamma_2 b)]} = \frac{\gamma_0 [A_3 J_0(\gamma_0 b) + B_3 N_0(\gamma_0 b)]}{[A_3 J_1(\gamma_0 b) + B_3 N_1(\gamma_0 b)]} \quad (1)$$

$$\begin{cases} A_2 = (\gamma_2/\gamma_0) J_1(\gamma_0 c) N_0(\gamma_2 c) - J_0(\gamma_0 c) N_1(\gamma_2 c) \\ B_2 = J_0(\gamma_0 c) J_1(\gamma_2 c) - (\gamma_2/\gamma_0) J_0(\gamma_2 c) J_1(\gamma_0 c) \end{cases} \quad (2)$$

$$\begin{cases} A_3 = -\frac{\pi}{2} \gamma_0 a N_1(\gamma_0 a) \\ B_3 = \frac{\pi}{2} \gamma_0 a J_1(\gamma_0 a) \end{cases} \quad (3)$$

$$\begin{aligned} \gamma_i^2 &= \omega^2 \mu_0 \varepsilon_0 \varepsilon_i - \pi^2/h^2 \\ \gamma_1 &= \gamma_3 = \gamma_0 \\ \gamma_2^2 &= \gamma_1^2 = \gamma_3^2 = \gamma_0^2 = \omega^2 \mu_0 \varepsilon_0 - \pi^2/h^2 \end{aligned} \quad (4)$$

We got a dimensions $a = 87.5$ mm $b = 43.75$ mm $c = 36.68$ mm $h = 162.9$ mm. Then we used FEM simulation software to simulate the sapphire loaded cavity. The E field in Fig. 2 and Fig. 3 was plotted to recognize the TE₀₁₁ mode. We got the $w = 1.4358$ GHz, unloaded $Q = 54727$.

The FEM software doesn't have the ability to recognize the mode TE₀₁₁, so we use the E field above to recognize the TE₀₁₁ mode.

After simulation we need to judge if the cavity could achieve H maser self-oscillation. We can use S parameter qualification. If $S > 5900$, the cavity could achieve self-oscillation. S parameter is scatter parameter. It is defined by Eq. (5)

$$S = Q_L \eta' \quad (5)$$

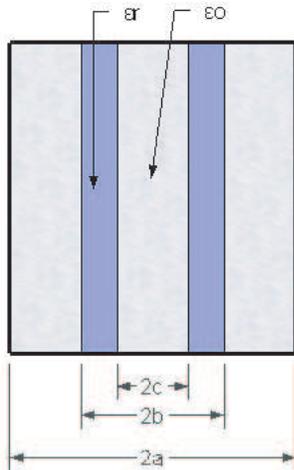


Figure 1: Structure of a sapphire loaded cavity.

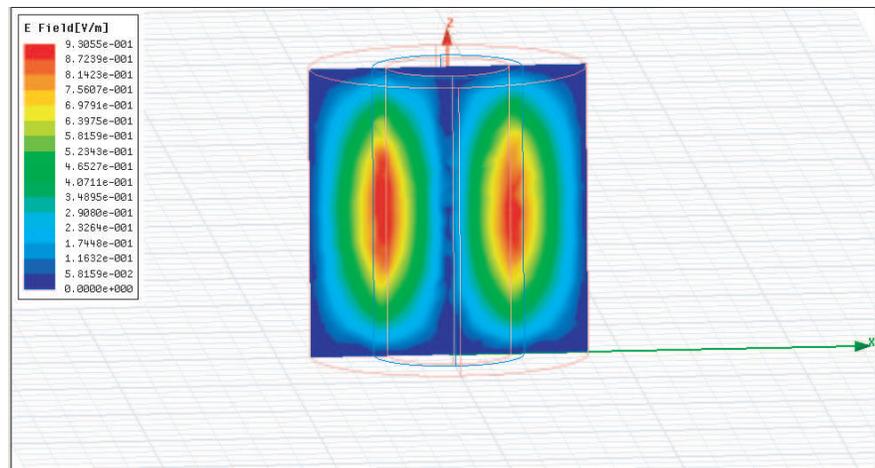


Figure 2: Recognized electric field of the cross section the cavity for TE₀₁₁.

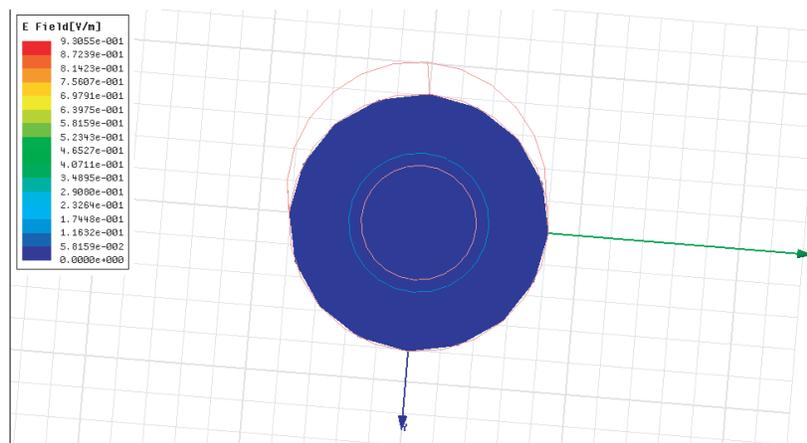


Figure 3: Recognized electric field of ends of the cavity for TE₀₁₁.

Q_L is loaded Q . η' is the filling factor of the cavity. The loaded Q is 45000 (Fig. 4) η' was calculated for the cavity by Eq. (5) using FEM software $\eta' = 0.523$. Then We can get the $S = 23535 > 5900$. The cavity can achieve self-oscillation

$$\eta' = \frac{V_b \langle H_z \rangle_b^2}{V_c \langle H^2 \rangle_c} = \frac{1}{V_b} \frac{\left[\int_b H_z dV \right]^2}{\int_c H^2 dV} \quad (6)$$

3. THE EXPERIMENT OF SAPPHIRE LOADED CAVITY

In order to confirm if the sapphire loaded cavity could sustain self-oscillation. We did a experiment on the cavity. We use traditional active H maser (shao-4) as our bed for experiment. That's a convenient way to reduce the cost of the experiment. We designed a equipment to fix the sapphire loaded cavity in the traditional H maser's vacuum system. The cavity frequency and loaded Q can be measured by net analyzer (Fig. 4) and with the flux 1.1 mA \sim 1.2 mA we got a signal of -102.9 dBm (Fig. 5).

And then we test the stability of the maser we got the 1s 9.8×10^{-13} , 10s 1.33×10^{-13} . For the equipment which fix the cavity still have stress problem, and the temperature control system of the traditional H maser can't meet the need of sapphire load cavity. (Temperature coefficient of the sapphire loaded cavity is much lager than the traditional cavity). We got a very bad long term stability.

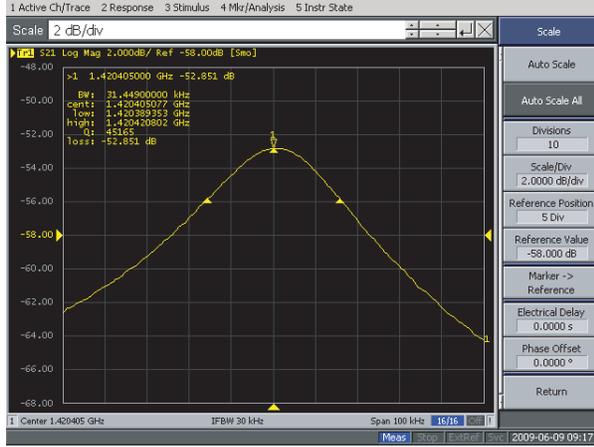


Figure 4: Measured loaded Q and Cavity frequency of the sapphire loaded cavity.

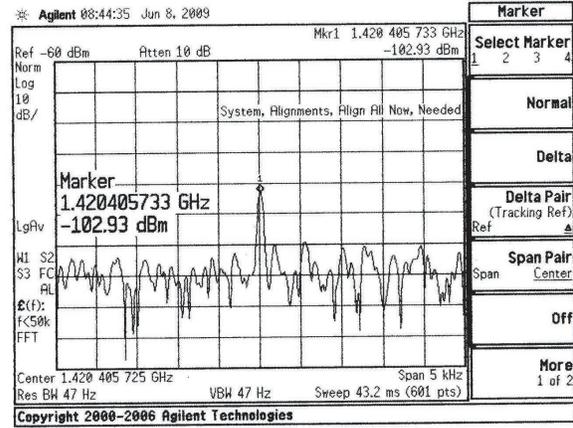


Figure 5: Self-oscillation signal of sapphire loaded H maser Temperature-compensation of the sapphire loaded cavity.

4. TEMPERATURE-COMPENSATION OF THE SAPPHIRE LOADED CAVITY

As we mentioned above, temperature coefficient of the sapphire loaded cavity is about $-55 \text{ kHz}/^\circ\text{C}$. It's much larger than the traditional cavity. There are two ways to reduce the bad effect on stability caused by high temperature coefficient. One is to build a very precisely temperature control system, the other is to compensate the temperature frequency shift of the sapphire loaded cavity. Our analysis is mainly on reducing the temperature coefficient of the cavity.

The sapphire loaded cavity's high temperature coefficient is the result of high temperature shift of the permittivity of sapphire (Al_2O_3). If we use another crystal which have the opposite temperature coefficient of permittivity to compensate the sapphire loaded cavity. We can get a low temperature coefficient cavity. We used the method mentioned in [4] to calculate the temperature coefficient. The equation of the temperature coefficient is as follows.

$$\frac{1}{f} \frac{\partial f}{\partial T} = \frac{1}{f} \left[\frac{\partial f}{\partial \varepsilon_{ts}} \frac{\partial \varepsilon_{ts}}{\partial T} + \frac{\partial f}{\partial \varepsilon_{tc}} \frac{\partial \varepsilon_{tc}}{\partial T} + \frac{\partial f}{\partial d} \frac{\partial d}{\partial T} + \frac{\partial f}{\partial h_1} \frac{\partial h_1}{\partial T} + \frac{\partial f}{\partial h_2} \frac{\partial h_2}{\partial T} \right] \quad (7)$$

$$\frac{1}{f} \frac{\partial f}{\partial T} = A_{r1} \tau_{r1} + A_{r2} \tau_{r2} + A_d \tau_{\alpha1} + A_{h1} \tau_{\alpha2} + A_a \tau_c + A_{h2} \tau_c \quad (8)$$

$$A_{r1} = \frac{\varepsilon_{ts}}{f} \frac{\Delta f}{\Delta \varepsilon_{ts}} \quad A_{r2} = \frac{\varepsilon_{tc}}{f} \frac{\Delta f}{\Delta \varepsilon_{tc}} \quad A_d = \frac{d}{f} \frac{\Delta f}{\Delta d} \quad A_a = \frac{a}{f} \frac{\Delta f}{\Delta a}$$

$$A_h = \frac{h}{f} \frac{\Delta f}{\Delta h} \quad \tau_{r1} = \frac{\Delta \varepsilon_{ts}}{\varepsilon_{ts} \Delta T} \quad \tau_{r2} = \frac{\Delta \varepsilon_{tc}}{\varepsilon_{tc} \Delta T} \quad \tau_{\alpha1} = \frac{\Delta d}{d \Delta T} \quad \tau_{\alpha2} = \frac{\Delta h_1}{h_1 \Delta T}$$

$$\tau_c = \frac{\Delta a_1}{a \Delta T} = \frac{\Delta h_2}{h_2 \Delta T}$$

τ_{r1} : the permittivity shift with temperature of sapphire (Al_2O_3),

τ_{r2} : the permittivity shift with temperature of compensation crystal,

$\tau_{\alpha1} \tau_{\alpha2}$: the radial and axes, expansion coefficient of the sapphire (Al_2O_3),

τ_c : the expansion coefficient of Ti.

We used SrTiO_3 to compensate the sapphire crystal for its high permittivity, which is 300, low dielectric loss, which is 5×10^{-4} and opposite permittivity shift with temperature. The Fig. 5 shows the compensation results of our calculation for sapphire loaded cavity with different dimensions.

From Fig. 6, we notice that the effect of compensation arose with ρ_1 . And with $a = 87.5$ we could find a zero temperature coefficient point between $\rho_1 = 0.54$ and 0.56 . Also we can find another region where zero temperature coefficient point could be found by fixing ρ_1 to 0.56 , taking a between 87.5 mm and 85 mm . We took ρ_1 range from 0.4 to 0.6 because Q and η' of this region is better.

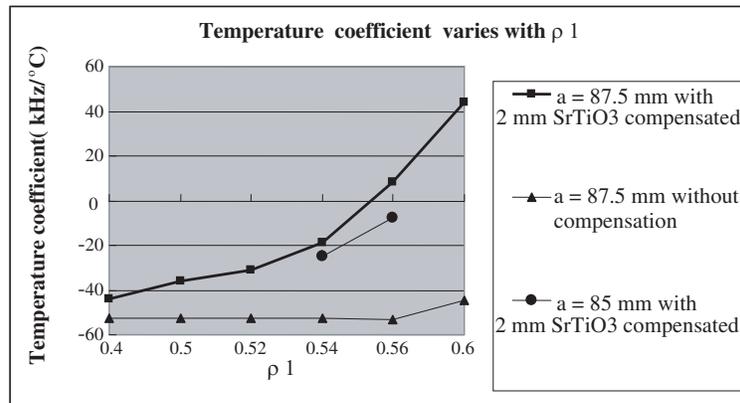


Figure 6: Temperature coefficient varies with ρ_1 .

5. CONCLUSION

Small size sapphire loaded cavity for H maser has been analyzed. Its ability to sustain self-oscillation has been confirmed both on theory and experiment. For the temperature coefficient of the sapphire loaded cavity is too high, method of compensation with SrTiO₃ was discussed.

REFERENCES

1. Morikawa, T., K. Takahei, M. Uehara, K. Mori, and M. Tsuda, "Design optimization of a sapphire loaded cavity for a spaceborne hydrogen maser," *Proc. of 14th EFTF*, 462–465, 2000.
2. Wang, N.-R., T.-Z. Zhou, and L.-S. Gao, "Recent progress on a sapphire loaded microwave cavity for compact hydrogen masers," *Frequency Control Symposium, China*, 144–147, 2008.
3. Zivanov, S., H. Schweda, D. Goujon, D. Gritti, and G. Perruchoud, "Physics package of the 35 kg active hydrogen maser for the ACES mission of ESA," *Frequency Control Symposium, Joint with the 21st European Frequency and Time Forum*, 637–641, 2007.
4. Yu, C., "Frequency-temperature analysis of sapphire ring dielectric-loaded cylindrical cavity in hydrogen maser atomic frequency standard," *Asia-Pacific Radio Science Conference*, 13–16, 2004.

Improvements on Phase-Shifted Distributed-Coupling-Coefficient Distributed Feedback Laser Structures for Single Longitudinal Mode Operation

J. Boavida¹, C. Fernandes¹, and J. Morgado^{1,2}

¹Instituto de Telecomunicacoes, Portugal

²Portuguese Air Force Academy, Portugal

Abstract— An optimized DFB laser structure is analyzed using an improved version of the Transfer-Matrix-Method. The proposed laser is specially useful in directly-modulated optical communication systems, since it presents extremely high selectivity and extremely low flatness at threshold regime, which results in an emitted spectrum above threshold that is very coherent and stable, with a side mode suppression ratio higher than 49 dB for a wide range of biasing current.

1. INTRODUCTION

Optical communication systems (OCS) have earned great attention from the scientific community, due to the increasing demand for high bit-rate data links [1]. In order to ensure such high bit-rates, the spectrum of the optical emitter should be as much stable and coherent as possible, which can be accomplished using Distributed-Feedback (DFB) lasers. However, not every DFB laser is considered an effective device. Therefore, there is the need to optimize such laser structures as much as possible, avoiding the deleterious effect of spatial hole burning (SHB), which is particularly important in directly modulated lasers. Many structures have been proposed to replace the conventional DFB laser, starting with the popular quarterly wavelength shifted (QWS)-DFB (Ref. [2]) and advancing to other complicated structures such as the corrugation-pitch-modulated (CPM) distributed-coupling-coefficient (DCC) DFB structure proposed in Ref. [3]. The aim of this paper is to introduce an optimized DFB laser which outperforms the structures referred elsewhere in literature.

2. METHODOLOGY

The outputs of this work are the specifications of a DFB laser structure with high figures of merit. Such structure is defined throughout an optimization process considering the restrictions that are initially imposed, and ultimately define the structural constraints associated with the DFB laser that has optimal performance. Such process is carried out at threshold regime so that the computational workload can be sustained. Afterwards, in order to assess the impact of the SHB effect, it is confirmed whether this DFB laser structure also presents high figures of merit above threshold regime.

In order to ensure a flawless convergence on the simulation of DFB laser characteristics at threshold and above-threshold regimes, an improved version of the Transfer-Matrix-Method (TMM) is used, which is fully described in Ref. [4]. The TMM considers that the laser cavity is divided into several sections, being each one associated with a transfer-matrix that is defined by the constancy

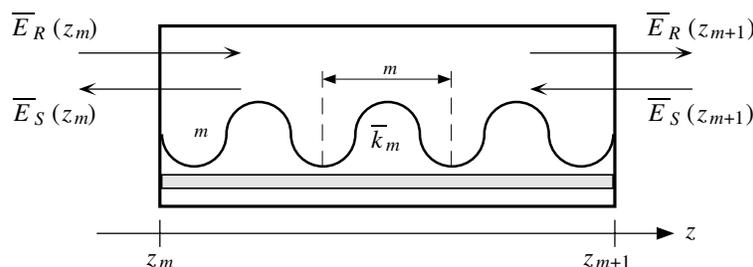


Figure 1: A simplified schematic diagram for a one-dimensional DFB laser structure section, placed between $z = z_m$ and $z = z_{m+1}$.

of its structural parameters. Each transfer-matrix correlates the two counter propagating waves, \overline{E}_R and \overline{E}_S , between both section ends (see Fig. 1).

The TMM oscillation conditions leads to pairs (δ, α) that are, respectively, the detuning and the gain of the oscillation modes. The figures of merit at the threshold regime are the normalized mode selectivity (\mathfrak{S}) and the flatness (\mathfrak{F}), which are defined, respectively, by [2, p. 131 and p. 128]

$$\mathfrak{S} = (\alpha \cdot L_{\text{cav}}) - (\alpha_{th} \cdot L_{\text{cav}}); \quad \mathfrak{F} = \frac{1}{L_{\text{cav}}} \int_0^{L_{\text{cav}}} [I(z) - \overline{I}]^2 dz \quad (1)$$

where L_{cav} is the cavity length and $I(z)$ is the normalised electric field intensity at an arbitrary position, z , given by

$$I(z) = |\overline{E}_R(z)|^2 + |\overline{E}_S(z)|^2, \quad (2)$$

and \overline{I} is its average value along the cavity. Besides, $(\alpha_{th} \cdot L_{\text{cav}})$ is the normalised threshold gain and $(\alpha \cdot L_{\text{cav}})$ is the normalised gain of the main side mode. The unchanged laser parameters assumed along the paper are summarized in Table 1.

Table 1: Laser parameters used in the simulation model.

Material parameters	Value
Spontaneous emission rate, A	$2.5 \times 10^8 \text{ s}^{-1}$
Bimolecular recombination coefficient, B	$1.0 \times 10^{-16} \text{ m}^3 \cdot \text{s}^{-1}$
Auger recombination coefficient, C	$3.0 \times 10^{-41} \text{ m}^6 \cdot \text{s}^{-1}$
Differential gain, A_0	$2.70 \times 10^{-20} \text{ m}^2$
Gain curvature, A_1	$1.50 \times 10^{19} \text{ m}^{-3}$
Differential peak wavelength, A_2	$2.70 \times 10^{-32} \text{ m}^4$
Internal loss, α_{loss}	$4.0 \times 10^3 \text{ m}^{-1}$
Effective index at zero injection, n_0	3.41351524
Carrier density at transparency, N_0	$1.23 \times 10^{24} \text{ m}^{-3}$
Differential index, dn/dN	$-1.8 \times 10^{-26} \text{ m}^3$
Group velocity, v_g	$8.33 \times 10^7 \text{ m} \cdot \text{s}^{-1}$
Nonlinear gain coefficient, ε	$1.5 \times 10^{-23} \text{ m}^3$
Structural parameters	Value
Active layer width, w	$1.5 \mu\text{m}$
Active layer thickness, d	$0.12 \mu\text{m}$
Cavity length, L_{cav}	$500 \mu\text{m}$
Optical confinement factor, Γ	0.35
Grating period, Λ	227.066 nm

3. STRUCTURE DEFINITION

The DFB structure under optimization is designated as 1PS-3DCC-DFB, since it has a single phase-shift (1PS) and a DCC profile with three coupling coefficients. The phase-shift is located in the center of the cavity ($z/L_{\text{cav}} = 0.5$) and its value is $\theta = 90^\circ$. The DCC profile is schematized in Fig. 2, where $\{K_{P_1}, K_{P_2}\}$ are the boundary normalized positions of the several sections and $\{\overline{k}_c L_{\text{cav}}, \overline{k}_{sc} L_{\text{cav}}, \overline{k}_s L_{\text{cav}}\}$ are the normalized coupling coefficients of these sections.

For the optimization procedure, the normalized average coupling coefficient and the coupling coefficient ratio are defined, respectively, as:

$$\overline{k}_{\text{av}} L_{\text{cav}} = 2 K_{P_1} \overline{k}_s L_{\text{cav}} + 2 (K_{P_2} - K_{P_1}) \overline{k}_{sc} + (1 - 2 K_{P_2}) \overline{k}_c, \quad (3)$$

$$\overline{k}_{\text{ratio}} = \frac{\overline{k}_c}{\overline{k}_s}. \quad (4)$$

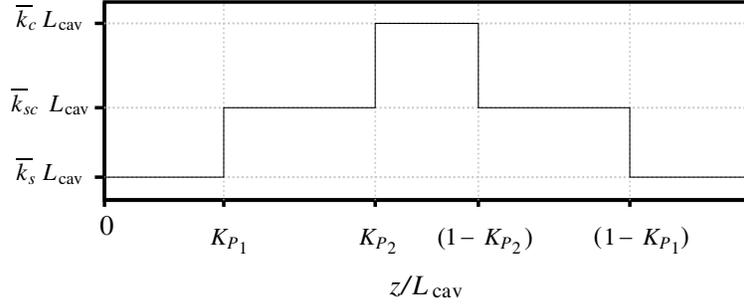


Figure 2: Schematic representation of the DCC profile under optimization.

The coupling coefficient of the intermediate section is defined as

$$\bar{k}_{sc} = \bar{k}_s \sqrt{\bar{k}_{\text{ratio}}}. \quad (5)$$

The decision variables for the optimization procedure are the following: $\{K_{P_1}, K_{P_2}, \bar{k}_{\text{ratio}}, \bar{k}_{\text{av}} L_{\text{cav}}\}$. In order to attain optimal values, the following procedure is carried out:

Step 1: $K_{P_1} = K_{P_2} = 0.25$, $\bar{k}_{\text{ratio}} = 5$ and $\bar{k}_{\text{av}} L_{\text{cav}} = 2$ are settled as a initial values;

Step 2: K_{P_1} and \bar{k}_{ratio} are varied simultaneously and independently, assuming $K_{P_2} = K_{P_1}$, until the best values of $\mathfrak{S}(K_{P_1}, \bar{k}_{\text{ratio}})$ and $\mathfrak{F}(K_{P_1}, \bar{k}_{\text{ratio}})$ are achieved;

Step 3: Assuming the previous optimized K_{P_1} and \bar{k}_{ratio} values, K_{P_2} and $\bar{k}_{\text{av}} L_{\text{cav}}$ are varied simultaneously and independently, until the best values of $\mathfrak{S}(K_{P_2}, \bar{k}_{\text{av}} L_{\text{cav}})$ and $\mathfrak{F}(K_{P_2}, \bar{k}_{\text{av}} L_{\text{cav}})$ are achieved;

Step 4: Assuming the previous optimized K_{P_2} and $\bar{k}_{\text{av}} L_{\text{cav}}$ values, K_{P_1} and \bar{k}_{ratio} are varied simultaneously and independently, until the best values of $\mathfrak{S}(K_{P_1}, \bar{k}_{\text{ratio}})$ and $\mathfrak{F}(K_{P_1}, \bar{k}_{\text{ratio}})$ are achieved;

Step 5: Steps 3–4 are repeated until no further improvements on \mathfrak{S} and \mathfrak{F} are achieved. At such stage the optimal solution is found.

At the end of the optimization process, the following decision variables are defined: $K_{P_1} = 0.1788$, $K_{P_2} = 0.2408$, $\bar{k}_{\text{ratio}} = 7$, $\bar{k}_{\text{av}} L_{\text{cav}} = 1.5$. Considering the Equations (3), (4), (5), these values correspond to a structure where $\bar{k}_s L_{\text{cav}} = 0.348$, $\bar{k}_{sc} L_{\text{cav}} = 0.920$ and $\bar{k}_c L_{\text{cav}} = 2.434$.

4. THRESHOLD ANALYSIS

After the definition of all the optimized structure constraints, it is possible to present its performance at threshold regime. Fig. 3(a) shows the normalized solutions for the threshold condition of the optimized structure. It is clear that there is a high mode selectivity between the main mode (the one with zero detuning and lowest gain) and the secondary modes. Fig. 3(b) compares the distribution of the normalized electrical field in the optimized 1PS-3DCC-DFB structure and two other structures, which are the standard QWS-DFB and the CPM-DCC-DFB from Ref. [3]. It is clear that the optimized structure has a flat field distribution similar to the one of the CPM-DCC-DFB structure.

These results near threshold regime are summarized in Table 2.

Table 2: \mathfrak{S} , \mathfrak{F} , $\alpha_{th} \cdot L_{cav}$ and I_{th} for several laser structures.

Laser structure	\mathfrak{S}	\mathfrak{F}	$\alpha_{th} \cdot L_{cav}$	I_{th} (mA)
Standard QWS-DFB	0.73	0.301	0.70	19.816
CPM-DCC-DFB from Ref. [3].	0.99	0.018	1.28	24.084
1PS-3DCC-DFB	2.34	0.021	1.52	26.104

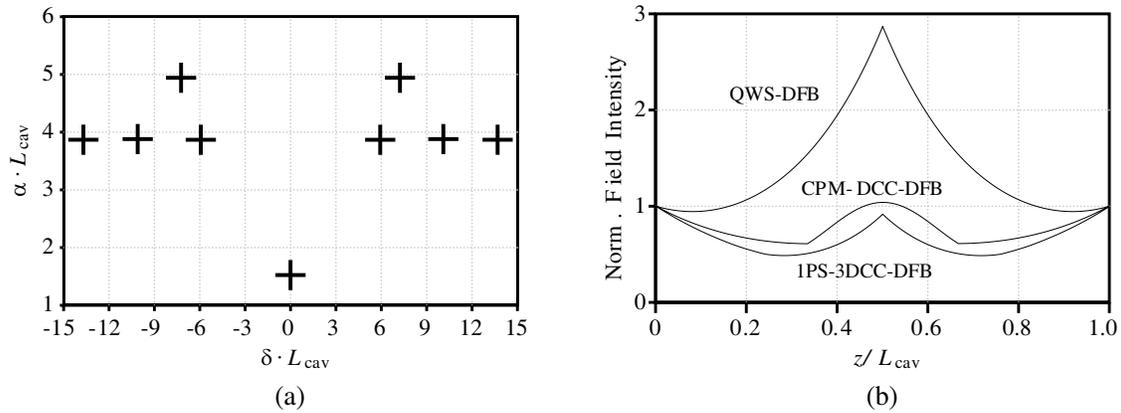


Figure 3: (a) Solutions for the threshold condition of the optimized 1PS-3DCC-DFB. (b) Normalized field intensity for the standard QWS-DFB, the CPM-DCC-DFB from Ref. [3] and the optimized 1PS-3DCC-DFB.

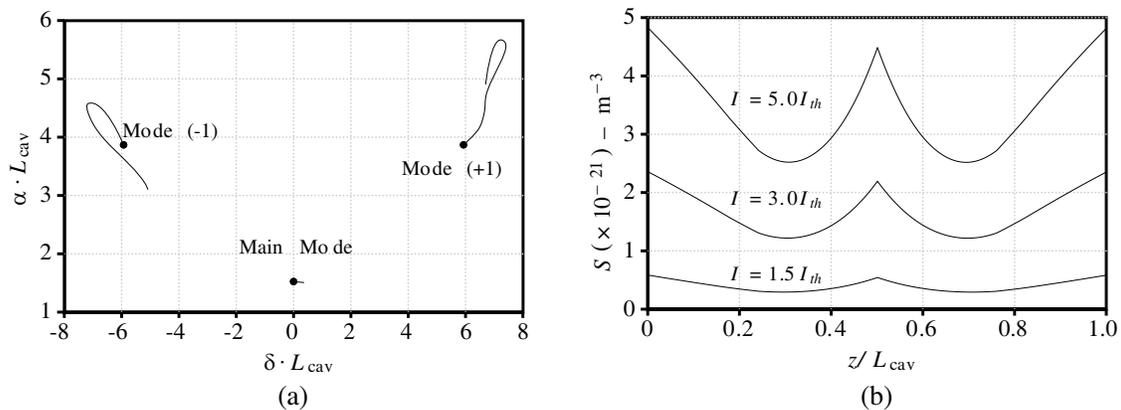


Figure 4: (a) Evolution of the modes $\{-1, 0, +1\}$ gain vs. detuning from $I = I_{th}$ (marked with a dot) up to $I = 5.0I_{th}$ and (b) Photon density distributions, from $I = 1.5I_{th}$ up to $I = 5.0I_{th}$.

5. ABOVE-THRESHOLD ANALYSIS

It is crucial to assess the extent of the SHB effect on the optimized 1PS-3DCC-DFB laser. For that, a biasing normalized current range of $I/I_{th} \in [1, 5]$ is analyzed. The corresponding above-threshold results are presented and discussed.

Figure 4(a) shows the evolution of the main mode and the closest secondary modes from $I = I_{th}$ (marked with a dot) up to $I = 5.0 I_{th}$, revealing that the normalized mode selectivity is not seriously harmed in this range, foreseeing a stable SMSR.

Figure 4(b) shows the evolution of the photon distribution for three normalized biasing currents. It is noticeable that this distribution remains flat even if the total number of photons in the cavity increases.

Figure 5(a) shows the emitted spectra of the optimized 1PS-3DCC-DFB laser for two biasing currents, revealing that a stable wavelength and also a stable SMSR are achieved.

Figure 5(b) shows the emitted lasing wavelength *vs.* normalized current injection for the optimized 1PS-3DCC-DFB laser, the standard QWS-DFB and the CPM-DCC-DFB from Ref. [3]. This figure reinforces the fact that the emitted wavelength of the optimized structure is very stable.

Figure 6(a) represents the SMSR *vs.* normalized current injection for the optimized 1PS-3DCC-DFB laser, the standard QWS-DFB and the CPM-DCC-DFB from Ref. [3]. This figure reinforces the fact that the SMSR is very high and stable for the 1PS-3DCC-DFB laser.

Figure 6(b) shows the $P(I/I_{th})$ characteristics for the optimized 1PS-3DCC-DFB laser, the standard QWS-DFB and the CPM-DCC-DFB from Ref. [3], where P is the optical power output measured at the right facet. For similar normalized current injections, the optimized 1PS-3DCC-DFB laser proposed in this paper shows larger values of P . Moreover, it has been checked that the optimized 1PS-3DCC-DFB laser has the highest external differential efficiency.

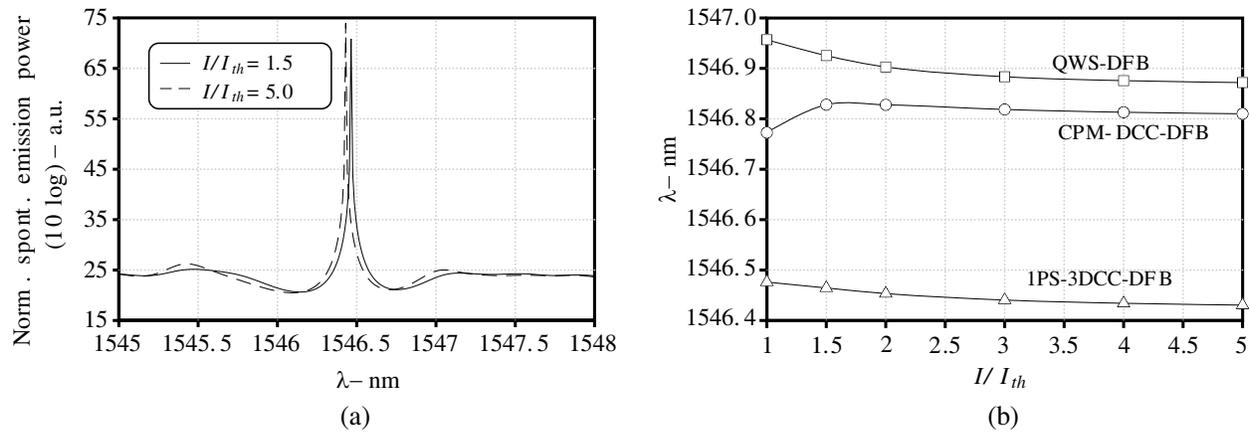


Figure 5: (a) Normalized spontaneous emitted spectra for the optimized laser and (b) $\lambda(I/I_{th})$ characteristic for the optimized 1PS-3DCC-DFB, the CPM-DCC-DFB from Ref. [3] and the QWS-DFB lasers.

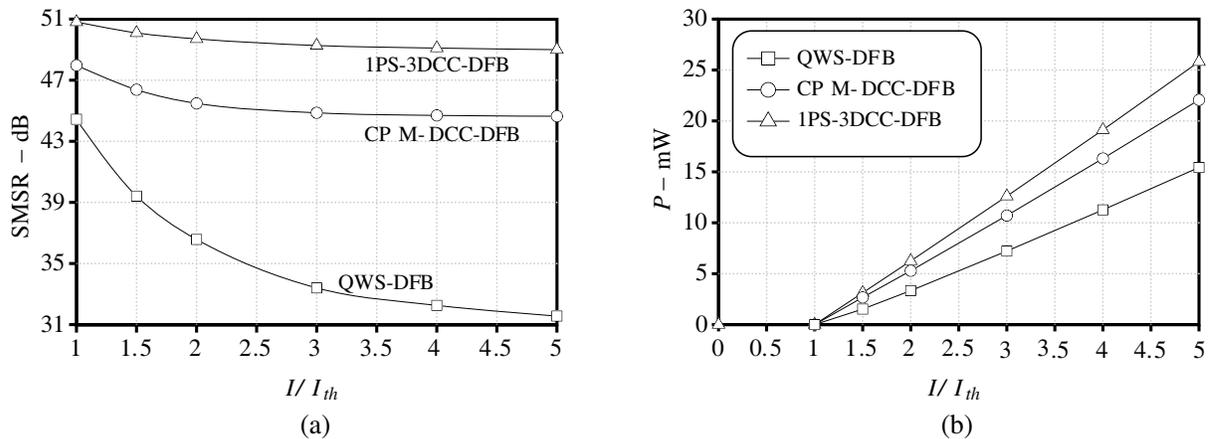


Figure 6: Above threshold results for the structures under analysis: (a) SMSR (I/I_{th}) and (b) $P(I)$.

6. CONCLUSION

An optimized 1PS-3DCC-DFB laser structure has been proposed and analysed. It reveals a very good immunity to SHB, being unmatched by other DFB lasers found in literature.

It can sustain a stable SMSR higher than 49 dB for a wide range of biasing current, as well as a power efficiency of about 0.256 W/A, combined with a wavelength stability ($\frac{\Delta\lambda}{\lambda}$) of 2.955×10^{-5} . Considering such high figures of merit, this laser structure can be very useful in the context of the modern high bit-rate optical communication systems using direct modulation.

REFERENCES

1. Bornholdt, C., U. Troppenz, J. Kreissl, W. Rehbein, B. Sartorius, M. Schell, and I. Woods, "40 Gbit/s directly modulated passive Feedback DFB laser for transmission over 320 km single mode fibre," *Proc. 34th European Conference on Optical Communication (ECOC'08)*, 43–44, Brussels, Belgium, September 2008.
2. Ghafouri-Shiraz, H., *Distributed Feedback Laser Diodes and Optical Tunable Filters*, J. Wiley & Sons, Chich., UK, 2003.
3. Fessant, T. "Threshold and above-threshold analysis of corrugation-pitch-modulated DFB lasers with inhomogeneous coupling coefficient," *IEE Proc. Optoelectron. — Pt. J*, Vol. 144, No. 6, 365–376, December 1997.
4. Fernandes, C., J. Morgado, and J. Boavida, "Optimization of an asymmetric three phase-shift distributed feedback semiconductor laser structure concerning the above-threshold stability," *The European Phys. J. Appl. Phys.*, Vol. 46, No. 3, 30701, June 2009.

On the Performance of DFB Laser Structures Specially Designed for Directly-Modulated Optical Communication Systems

J. Boavida¹, C. Fernandes¹, and J. Morgado^{1,2}

¹Instituto de Telecomunicacoes, Portugal

²Portuguese Air Force Academy, Portugal

Abstract— We propose and analyze two easily manufacturable optimized DFB laser structures that are suitable for directly-modulated optical communication systems. These structures present high power efficiency, a good immunity to spatial hole burning and high yield. The first DFB laser structure is extremely simple to fabricate whereas the second DFB laser structure has an enhanced performance at the expense of the inclusion of a slightly more complicated grating. Previously, in order to achieve a similar performance, hardly manufacturable lasers have been proposed elsewhere.

1. INTRODUCTION

Optical communication systems (OCS) have earned great attention from the scientific community, due to the increasing demand for high bit-rate data links [1]. In order to ensure such high bit-rates using directly-modulated (DM) lasers, the spectrum of the optical emitter should be as much stable and coherent as possible, which can be accomplished with Distributed-Feedback (DFB) lasers. Hence, there is the need to optimize DFB laser structures, to avoid the deleterious effect of spatial hole burning (SHB). Despite DFB lasers with both anti-reflection (AR) facets are immune to the undesirable effects of the phase randomness inserted during fabrication [2], they are considered power inefficient [3]. Some complicated laser structures with high-reflectivity (HR) facets have been recently proposed [3] in order to fulfil the OCS requirements. The aim of the present paper is to introduce optimized DFB lasers that outperform the structures referred elsewhere in literature.

2. METHODOLOGY

The outputs of this work are the structural specifications of DFB lasers that have high figures of merit in the DM-OCS context. Such structures are defined throughout an optimization process considering the restrictions that are initially imposed, and that ultimately define the structural constraints associated with the DFB lasers that have optimal performance. Such process is carried out at threshold regime so that the computational workload can be sustained. Afterwards, in order to assess the impact of the SHB effect, it is confirmed whether these DFB laser structures also present high figures of merit above threshold regime.

In order to ensure a flawless convergence in the simulation of DFB laser characteristics at threshold and above-threshold regimes, an improved version of the Transfer-Matrix-Method (TMM) is used, which is fully described in Ref. [4]. The TMM considers that the laser cavity is divided into several sections, being each one associated with a transfer-matrix that is defined by the constancy of its structural parameters. Each transfer-matrix correlates the two counter propagating waves, \bar{E}_R and \bar{E}_S , between both section ends (see Fig. 1).

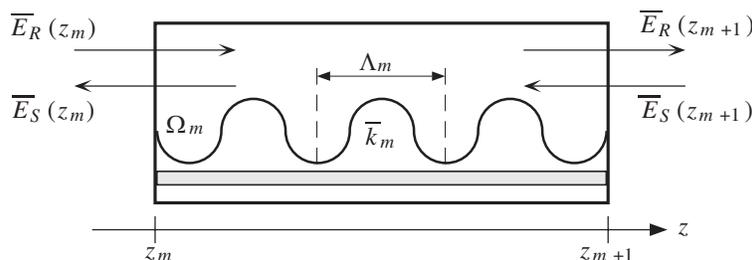


Figure 1: A simplified schematic diagram for a one-dimensional DFB laser structure section, placed between $z = z_m$ and $z = z_{m+1}$.

The TMM oscillation conditions leads to pairs (δ, α) that are, respectively, the detuning and the gain of the oscillation modes. The figures of merit at the threshold regime are the normalized mode selectivity (\mathfrak{S}) and the flatness (\mathfrak{F}), which are defined, respectively, by [5, p. 131 and p. 128]

$$\mathfrak{S} = (\alpha \cdot L_{\text{cav}}) - (\alpha_{\text{th}} \cdot L_{\text{cav}}); \quad \mathfrak{F} = \frac{1}{L_{\text{cav}}} \int_0^{L_{\text{cav}}} [I(z) - \bar{I}]^2 dz \quad (1)$$

where L_{cav} is the cavity length and $I(z)$ is the normalised electric field intensity at an arbitrary position, z , given by

$$I(z) = |\bar{E}_R(z)|^2 + |\bar{E}_S(z)|^2, \quad (2)$$

and \bar{I} is its average value along the cavity. Besides, $(\alpha_{\text{th}} \cdot L_{\text{cav}})$ is the normalised threshold gain and $(\alpha \cdot L_{\text{cav}})$ is the normalised gain of the main side mode. The unchanged laser parameters assumed along the paper are summarized in Ref. [5, p. 157].

3. STRUCTURES DEFINITION

The DFB structures under optimization are designated as HR-AR-DCC-DFB (see Fig. 2) and HR-AR-LDCC-DFB (see Fig. 3). They both have a left-end HR facet ($\hat{r}_1 = 1 \cdot e^{j\varphi_1}$) and a right-end AR facet ($\hat{r}_2 = 0$) in order to channel all the output power to the right side. Besides, they both have a symmetric corrugation with constant period.

According to Figs. 2 and 3, both DFB structures have a distributed-coupling-coefficient (DCC) profile, but the HR-AR-LDCC-DFB structure has a more elaborated DCC profile. The HR-AR-DCC-DFB structure has only two normalized coupling coefficients, designated by $\bar{k}_s L_{\text{cav}}$ and $\bar{k}_c L_{\text{cav}}$, and the transition occurs at the normalized boundary positions K_{P_1} and $1 - K_{P_1}$. The HR-AR-LDCC-DFB structure has a linear transition between $\bar{k}_s L_{\text{cav}}$ and $\bar{k}_c L_{\text{cav}}$ in order to smooth the DCC profile. Such transition occurs from K_{P_1} to K_{P_2} and again from $1 - K_{P_2}$ to $1 - K_{P_1}$.

Both structures undergo an optimization process in order to achieve the optimal values for $\bar{k}_s L_{\text{cav}}$, $\bar{k}_c L_{\text{cav}}$, K_{P_1} , K_{P_2} and φ_1 . The optimal parameters are summarized in Tables 1 and 2.

4. THRESHOLD ANALYSIS

After the definition of all the optimized structure constraints, it is possible to present its performance at threshold regime. Fig. 4(a) shows the normalized solutions for the threshold condition of the optimized structures. It is clear that there is a high mode selectivity between the main mode

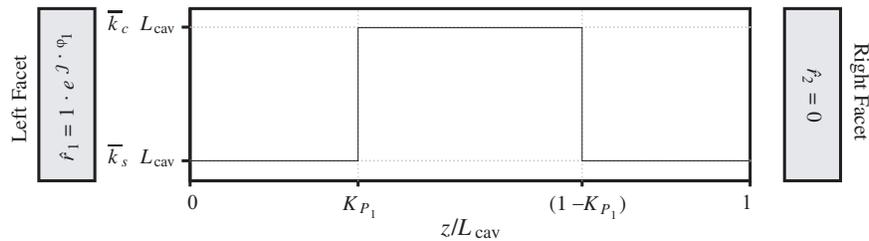


Figure 2: Schematic representation of the HR-AR-DCC-DFB structure under optimization.

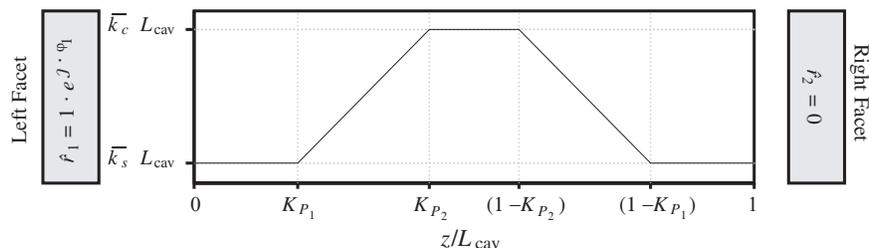


Figure 3: Schematic representation of the HR-AR-LDCC-DFB structure under optimization.

Table 1: Optimized HR-AR-DCC-DFB structure parameters.

Parameter	$\bar{k}_s L_{cav}$	$\bar{k}_c L_{cav}$	K_{P_1}	φ_1 (rad)
Value	0.130	1.300	0.280	$\pi/2$

Table 2: Optimized HR-AR-LDCC-DFB structure parameters.

Parameter	$\bar{k}_s L_{cav}$	$\bar{k}_c L_{cav}$	K_{P_1}	K_{P_2}	φ_1 (rad)
Value	0.174	1.413	0.190	0.435	$\pi/2$

 Table 3: \mathfrak{S} , \mathfrak{F} , $\alpha_{th} \cdot L_{cav}$ and I_{th} for several laser structures.

Laser structure	\mathfrak{S}	\mathfrak{F}	$\alpha_{th} \cdot L_{cav}$	I_{th} (mA)
Standard QWS-DFB	0.73	0.301	0.70	19.816
HR-AR-DCC-DFB	1.17	0.014	0.59	19.103
HR-AR-LDCC-DFB	1.63	0.012	0.59	19.098

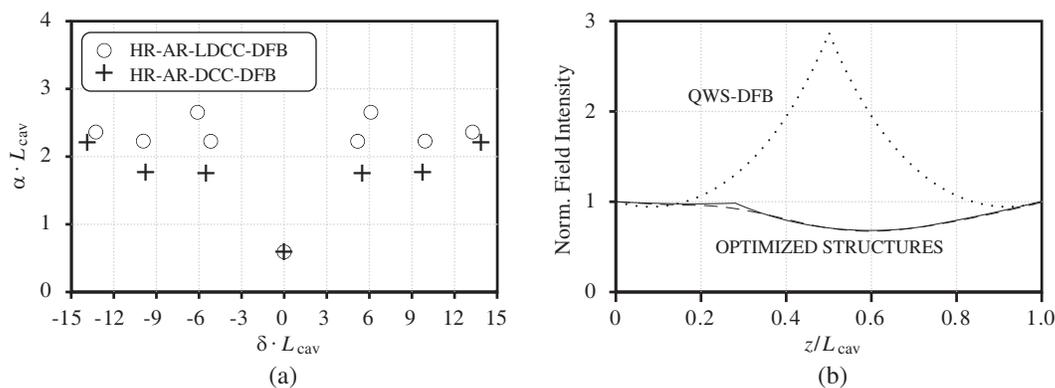


Figure 4: (a) Solutions for the threshold condition of the optimized structures, (b) normalized field intensity for the standard QWS-DFB and the optimized structures. The solid line stands for the HR-AR-DCC-DFB structure whereas the dashed line stands for the HR-AR-LDCC-DFB structure.

(the one with zero detuning and lowest gain) and the secondary modes. It is also possible to see that the HR-AR-LDCC-DFB structure has a higher selectivity than HR-AR-DCC-DFB structure. Fig. 4(b) compares the distribution of the normalized electrical field in the optimized structures and the standard QWS-DFB. It is apparent that the optimized structures have a flat field distribution.

These results near threshold regime are summarized in Table 3.

5. YIELD ANALYSIS

Given the lack of manufacturing control over φ_1 , it is crucial to test its influence on the laser performance. Thus, the main figures of merit of the optimized structures were assessed for every value of φ_1 . For the yield analysis we have considered \mathfrak{S} , \mathfrak{F} , $\Delta P/\Delta I$ and SMSR_{av} as the relevant laser figures of merit. SMSR_{av} is calculated as the average between SMSR ($I = 1.5 I_{th}$) and SMSR ($I = 5.0 I_{th}$). The SLM yield reaches a value higher than 50% for a preset condition of $\mathfrak{S} > 0.5$ and $\mathfrak{F} < 0.05$, corresponding closely to the range $\varphi_1 \in [0, \pi]$. Similar yield is indicated in Ref. [3], though for a laser associated with a rather complicated fabrication process. The results within this domain are illustrated in Figs. 5 and 6.

Figure 5(a) shows that, for $\varphi_1 \in [0, \pi]$, both optimized structures have \mathfrak{S} higher than the required minimum value ($\mathfrak{S} > 0.5$) for SLM operation. Besides, it is noticeable that the HR-AR-DCC-DFB structure has an overall increase of \mathfrak{S} over the HR-AR-LDCC-DFB structure, which is more intense for $\varphi_1 = \pi/2$.

Figure 5(b) shows that, for $\varphi_1 \in [0, \pi]$, both optimized structures have \mathfrak{F} considerably lower than the required maximum value ($\mathfrak{F} < 0.05$) for SLM operation. The HR-AR-DCC-DFB structure has an overall improvement of \mathfrak{F} over the HR-AR-LDCC-DFB structure.

Figures 6(a) and 6(b) show, respectively, the power efficiency and the $SMSR_{av}$ variations with φ_1 , for the optimized DFB structures. It is noteworthy that, within the range $\varphi_1 \in [0, \pi]$, the power efficiency is the lowest (highest) approximately where the $SMSR_{av}$ is the highest (lowest). Therefore, the fabrication of the optimized structures with the random facet phase $\varphi_1 \in [0, \pi]$ will benefit either the power efficiency or the $SMSR_{av}$.

6. ABOVE-THRESHOLD ANALYSIS

It is crucial to assess the extent of the SHB effect on the performance of the optimized lasers in the high power regime. For that, the biasing normalized current range $I/I_{th} \in [1, 5]$ is analyzed. The corresponding above-threshold results are presented and discussed.

Figure 7(a) shows the normalized spontaneous emitted spectra of the optimized lasers, greatly resembling each other. The main difference lies on the value of the emitted power associated with the main mode, thus the HR-AR-LDCC-DFB laser is expected to have a higher SMSR.

Figure 7(b) highlights the wavelength of the main modes associated with the HR-AR-LDCC-DFB, HR-AR-DCC-DFB and the QWS-DFB lasers. The optimized lasers follow the same characteristic and have a much more stable wavelength ($\frac{\Delta\lambda}{\lambda} \leq 1.3 \times 10^{-5}$) than the QWS-DFB wavelength.

Figure 8(a) plots the $P(I)$ characteristic for the optimized structures as well as for the QWS-

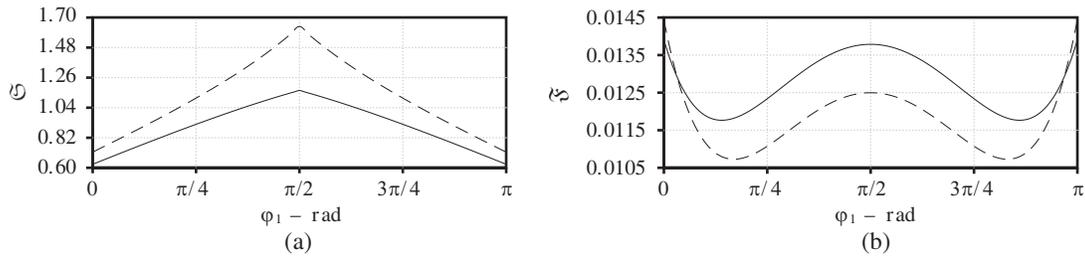


Figure 5: The dashed line stands for the HR-AR-LDCC-DFB structure while the solid line stands for the HR-AR-DCC-DFB structure. (a) $\mathfrak{S}(\varphi_1)$ and (b) $\mathfrak{F}(\varphi_1)$ characteristics.

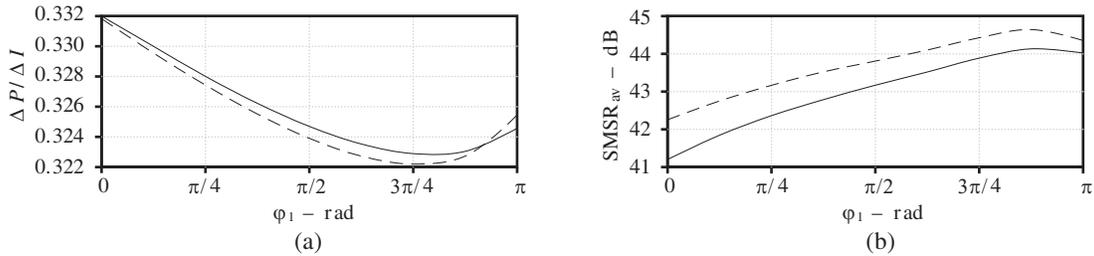


Figure 6: The dashed line stands for the HR-AR-LDCC-DFB structure while the solid line stands for the HR-AR-DCC-DFB structure. (a) $\Delta P/\Delta I(\varphi_1)$ and (b) $SMSR_{av}(\varphi_1)$ characteristics.

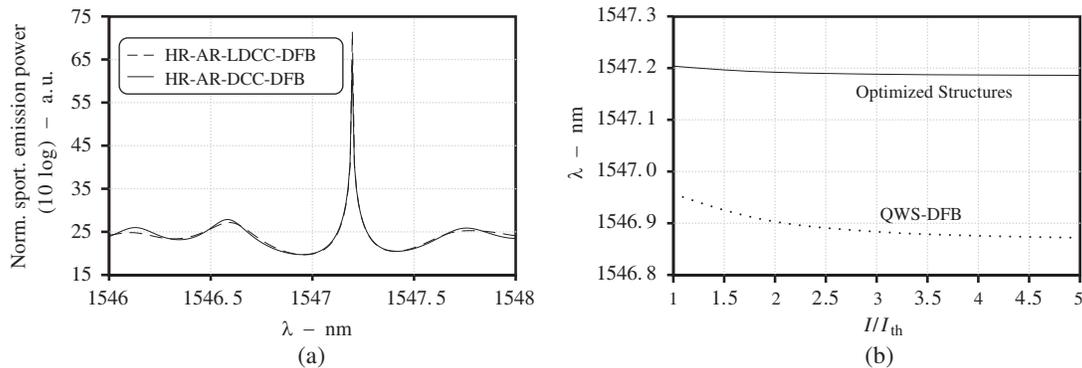


Figure 7: (a) Normalized spontaneous emitted spectra for the optimized lasers and (b) $\lambda(I/I_{th})$ characteristic for the HR-AR-LDCC-DFB, the HR-AR-DCC-DFB and the QWS-DFB lasers. Notice that the optimized structures (HR-AR-DCC-DFB and HR-AR-LDCC-DFB) have exactly the same $\lambda(I/I_{th})$ characteristic.

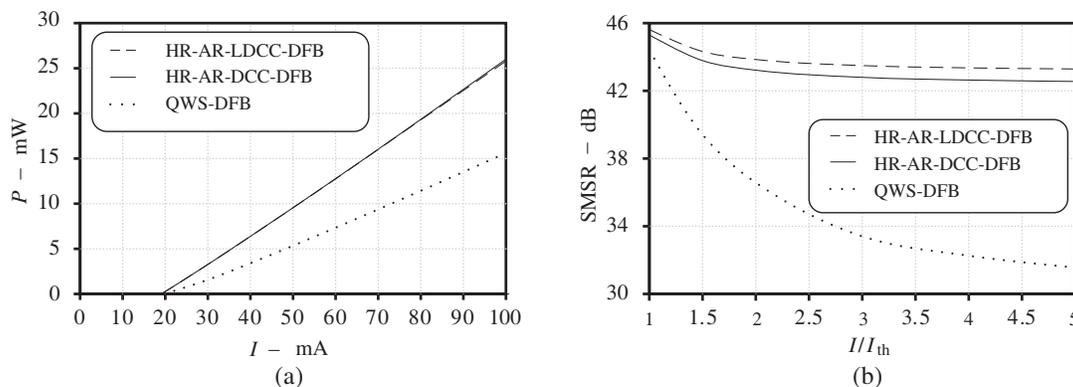


Figure 8: (a) $P(I)$ and (b) SMSR (I/I_{th}) characteristics for the HR-AR-LDCC-DFB, the HR-AR-DCC-DFB and the QWS-DFB lasers.

DFB structure. The HR-AR-LDCC-DFB and HR-AR-DCC-DFB lasers have roughly the same $P(I)$ characteristic, obtaining a power efficiency of about 0.33 W A^{-1} , which is similar to the efficiency reported in Ref. [3] for a laser that is much harder to manufacture. This figure shows that the QWS-DFB laser has a much lower power efficiency than the optimized structures.

Figure 8(b) shows a small difference (about 1 dB) between the HR-AR-LDCC-DFB laser SMSR and the HR-AR-DCC-DFB laser SMSR. These optimized structures present a much higher and more stable SMSR than the QWS-DFB laser, foreseeing its potential usage in the OCS context.

7. CONCLUSION

Two optimized cost-effective DFB lasers have been proposed and analyzed. Considering the compromise between easy-feasibility and high-performance, HR-AR-DCC-DFB laser is considered the most suitable for easy-feasibility whereas HR-AR-LDCC-DFB laser is considered the most suitable for high-performance, even if they are both easy to manufacture and they both present commendable results.

Both proposed structures present, simultaneously, high power efficiency ($\Delta P/\Delta I \geq 0.32 \text{ W} \cdot \text{A}^{-1}$), a good SHB immunity ($\text{SMSR} \geq 41$ and $\frac{\Delta\lambda}{\lambda} \leq 1.3 \times 10^{-5}$) and high yield ($\geq 50\%$). Therefore, the proposed structures fulfil all the requirements concerning the usage of DM-DFB lasers in the OCS context. Similar performances have been recently reported [3] but at the expense of laser structures demanding a much more intricate fabrication process than the ones associated with the laser structures proposed in this paper.

REFERENCES

1. Bornholdt, C., U. Troppenz, J. Kreissl, W. Rehbein, B. Sartorius, M. Schell, and I. Woods, "40 Gbit/s directly modulated passive feedback DFB laser for transmission over 320 km single mode fibre," *Proc. 34th European Conference on Optical Communication (ECOC'08)*, 43–44, Brussels, Belgium, Sep. 2008.
2. Grillot, F., "On the effects of an antireflection coating impairment on the sensitivity to optical feedback of HR/AR semiconductor DFB lasers," *IEEE J. Quantum Electron.*, Vol. 45, No. 6, 720–729, Jun. 2009.
3. Yoffe, G., S. Zou, B. Pezeshki, S. Rishton, and M. Emanuel, "High-precision high-yield laser source using DFB array," *IEEE Photon. Technol. Lett.*, Vol. 16, No. 3, 735–737, Mar. 2004.
4. Fernandes, C., J. Morgado, and J. Boavida, "Optimization of an asymmetric three phase-shift distributed feedback semiconductor laser structure concerning the above-threshold stability," *The European Phys. J. Appl. Phys.*, Vol. 46, No. 3, 30701, Jun. 2009.
5. Ghafouri-Shiraz, H., *Distributed Feedback Laser Diodes and Optical Tunable Filters*, J. Wiley & Sons, Chich., UK, 2003.

Fabrication of Separately Formed Electro-spun Fibers

Hirohisa Tamagawa

Department of Human and Information Systems, Faculty of Engineering, Gifu University, Japan

Abstract— Simply placing a paper mesh between the syringe needle and counter electrode of electrospinning unit, the straight fibers were separately formed between the paper mesh and counter electrode. Using this simple technique employing a paper mesh, even the composite fibers consisting of polymer and metal powder were easily formed separately.

1. INTRODUCTION

Electrospinning displays its high efficiency for fabricating a two-dimensional mat consisting of fine fibers [1–4]. By the use of electrospinning, for example, a two-dimensional filter with superior functionality can be easily fabricated. However, individually fabricating fibers is a difficult task for the electrospinning technology, since electro-spun fibers are inevitably entangled and stick together during its formation process. Extraction of single fiber is quite difficult task, and establishing a single fiber extraction technique is a challenging theme in this field.

In this paper, a quite simple technique of separately fabricating electro-spun short fibers is introduced. This technique simply employs a paper mesh, which is to be placed between the needle tip and counter electrode of electrospinning unit. Furthermore, usability of this technique was examined by try fabricating fine fibers consisting of different kinds of polymers.

2. BACKGROUND

Figure 1 illustrates the setup of electrospinning apparatus. Highly charged polymer solution or melt loaded in a syringe is spewed out from the syringe needle and travels toward the counter electrode which serves as a fiber collectors because of the high electric field. Consequently, fine fibers are formed on the counter electrode. Those fibers formed are always laid on the counter electrode horizontally, and it is a benefit for fabricating a dense nonwoven fiber mat as illustrated in Figure 2. On the other hand, it becomes a drawback, when it comes to extracting fine fibers separately, since the fibers collected on the counter electrodes stick one another.

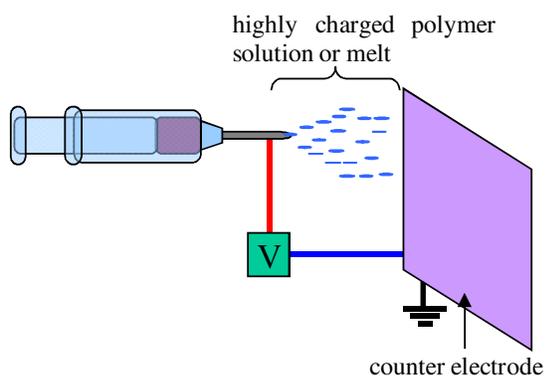


Figure 1: Setup of electrospinning.

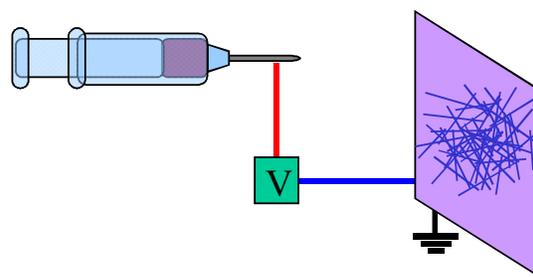


Figure 2: Fine fibers horizontally formed on the counter electrode forming a non-woven mat.

It was observed in our preliminary experiment that partially blocking the traveling path of highly charged polymer between the needle tip and counter electrode with some material, say cardboard frame, resulted in the formation of fibers between the cardboard frame and counter electrode vertically to the counter electrode. Figure 3 illustrates such a situation. The highly charged polymer spewed out from the needle tip of syringe formed fine fibers between the cardboard and counter electrode as indicated in the dotted circle. This simple technique was slightly modified for separately fabricating electro-spun fibers, and it is detailed in this paper.

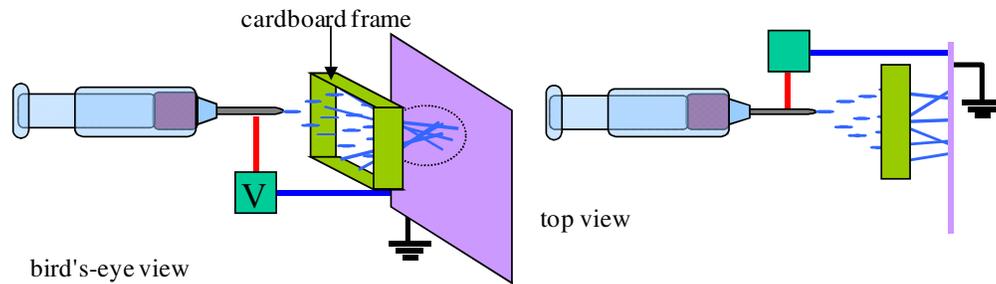


Figure 3: Separately formed electro-spun fibers by partially blocking the traveling path of highly charged polymer between the needle tip and counter electrode with a cardboard frame.

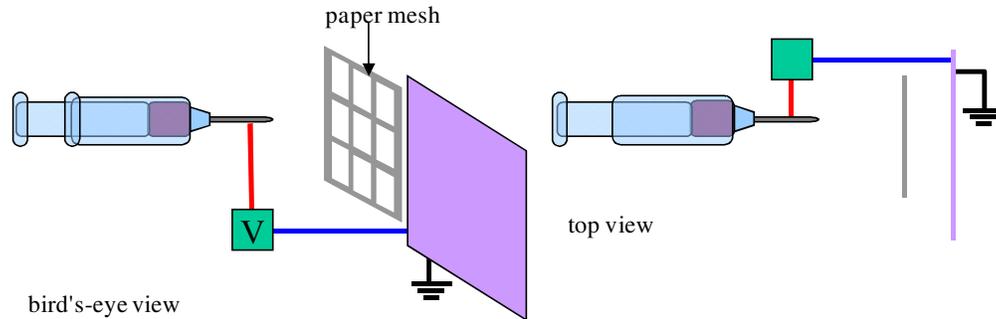


Figure 4: Setup for fabricating electro-spun fibers separately using a paper mesh.

3. EXPERIMENTAL

3.1. Polymer Solution

Two kinds of polymer solutions were prepared to be electro-spun.

- **Polynorbornen** 2.5 g of polynorbornen — hereafter called PN for short — was dissolved into 50 g of THF in the water bath at 70 degree C. Once it was fully dissolved, it was cooled down to the room temperature.
- **Polyvinylacetate** 11.5 g of polyvinyl acetate — hereafter called PVAc for short — was dissolved into 100 g of DMF in the water bath at 70 degree C. Once it was fully dissolved, it was cooled down to the room temperature.

3.2. Electrospinning Setup

Conventional setup of electrospinning was already shown in Figure 1. For separately fabricating electro-spun fibers, a paper mesh was simply placed between the needle tip and counter electrode as illustrated in Figure 4.

4. RESULTS AND DISCUSSION

4.1. Fiber Fabrication without a Paper Mesh

Using the setup without a paper mesh illustrated in Figure 1, the electrospinning of PN and PVAc was carried out, where the voltage was ~ 10 kV and the gap between the needle tip and counter electrode was ~ 10 cm. Figures 5(a) and 5(b) respectively show the fibers of PN and PVAc formed on the counter electrode. Concerning PN fibers, they are highly entangled and seen as dark islands in Figure 5(a). The PN fibers were formed horizontally to the counter electrode surface and stick together. These fibers could not be extracted separately. PVAc fibers were also formed on the counter electrode surface horizontally. Although those fibers are not seen so clearly in Figure 5(b), relatively a large number of fine PVAc fibers are seen at the right bottom of Figure 5(b). They stuck together and could not be extracted separately.

4.2. Separately Fabricated Short Fibers

Using the setup with a paper mesh illustrated in Figure 4, electrospinning was carried out. The highly charged polymer solution in the syringe was spewed out toward counter electrode under the high voltage, where the voltage was ~ 10 kV and the gap between the needle tip and counter

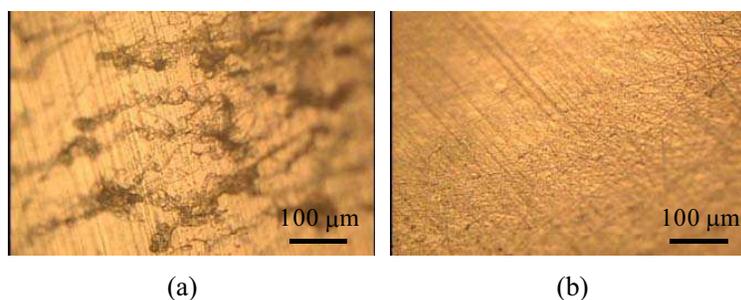


Figure 5: Electro-spun (a) PN and (b) PVAc fibers formed horizontally on the counter electrode.



Figure 6: PN fibers formed between the paper mesh and counter electrode.

electrode was ~ 10 cm. It passes through the paper mesh, resulted in the formation of fine fibers separately between the paper mesh and counter electrode. Use of the setup illustrated in Figure 4 does not result in the formation of infinitely long fibers but in the formation of short fibers. However, it is one of steps toward establishing a technique of forming long fine fibers separately. Figure 6 shows the PN fibers formed between the paper mesh and counter electrode. Compared with PN fibers shown in Figure 5(a), the fiber diameter was quite large. Fiber diameter heavily depends on the condition of electrospinning such as voltage, gap between the needle tip and counter electrode, the ratio of PN and solvent and so on. Since it was difficult to precisely control the experimental condition, it was difficult to fabricate the same diameter fibers. However, it was repeatedly confirmed, this paper mesh method could be definitely used for fabricating the PN fibers separately.

4.3. Formation Process of Separately Formed Short Fibers

Formation process of the separately formed short fibers is considered. Undoubtedly a quite simple process is brought to our mind as illustrated in Figure 7. Fiber ingredient is spewed out from the needle tip, and it travels toward paper mesh (Figure 7(a)). It is trapped with the mesh (Figure 7(b)), and immediately it is stretched toward the counter electrode, resulting in the short fiber formation (Figure 7(c)). But actual process was a bit different. Using a high speed camera (ES Kodak EKTAPRO HS Motion Analyzer Model 4540), the fiber formation process was analyzed.

Figure 8 shows the time history of PN fiber formation process from 0 ms through 160 ms, where time $t = 0$ ms in Figure 8(a) was arbitrarily defined. Once the voltage was applied between the syringe needle tip and counter electrode, the fiber ingredient was continuously supplied to the paper mesh as shown in the encircled area with solid line in Figure 8, where such a situation is roughly illustrated in the upper left of Figure 8 for better understanding. The polymer supply continued from beginning to end in this experiment. But fiber formation between the needle tip and counter electrode was not induced for a while even after the start of the polymer supply at $t = 0$ ms. Fiber formation suddenly started at $t = 40$ ms as shown in the encircled area with dotted line in Figure 8(b). The fiber continued to grow fat and it completed at $t = 160$ ms as shown in Figure 8(e). During fiber growing process, the fiber initially kept on waving quite largely but it gradually subsided down to stillness. Figure 9 shows the waving motion of fiber. Figures 9(a) and 9(b) were taken at $t = 56$ ms and 59 ms, respectively. The position of fiber in Figure 9(a) indicated by an arrow is obviously different from that in Figure 9(b) because of rampant wavy motion of fiber.

The fiber formation was not continuous process. Firstly, the fiber ingredient accumulated on the paper mesh. Secondly, the fiber jet was suddenly spewed out from the paper mesh to the counter electrode, resulting in a fiber. Since there were a number of holes on the paper mesh, fiber

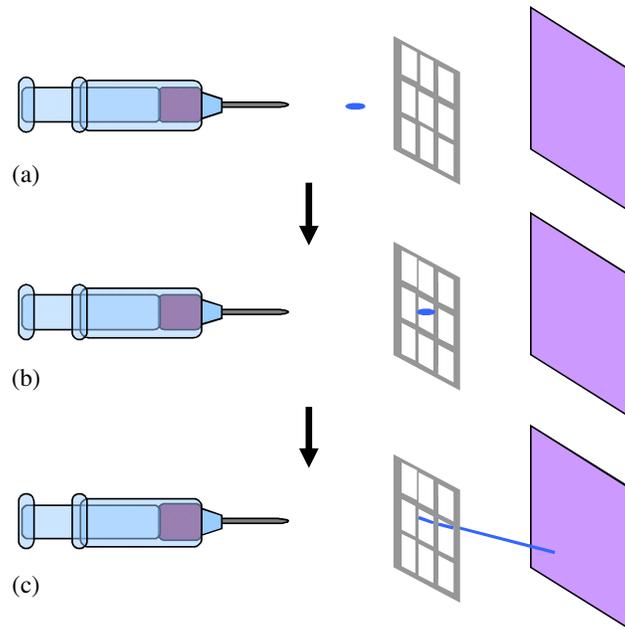


Figure 7: Formation process imagined for separately formed fibers between the paper mesh and counter electrode.

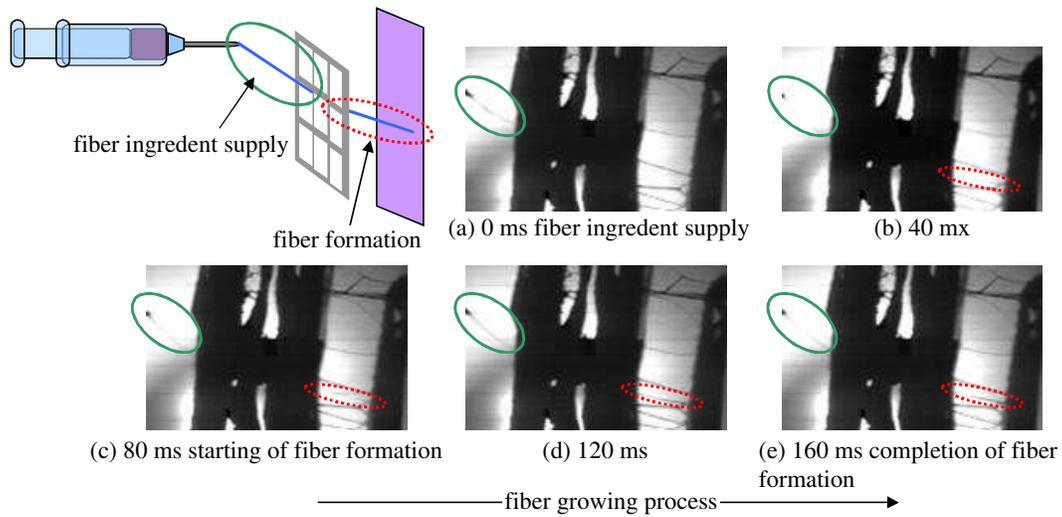


Figure 8: The time history of actual PN fiber formation process from 0 ms through 160 ms, where time $t = 0$ ms was defined arbitrarily.

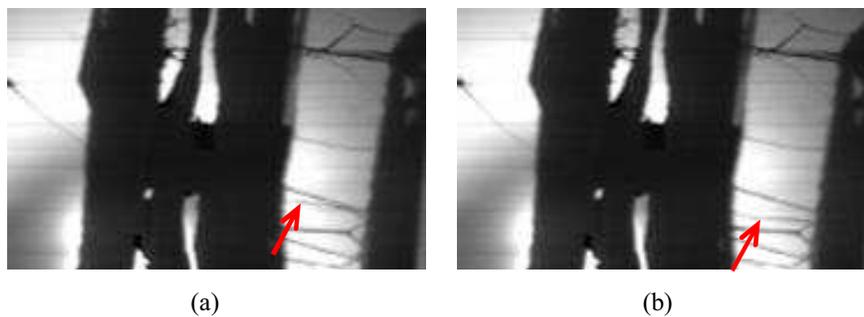


Figure 9: The waving motion of fiber at (a) $t = 56$ ms and (b) $t = 59$ ms.

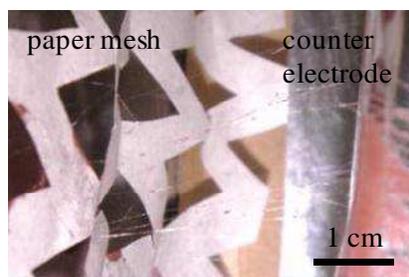


Figure 10: Electro-spun composite fibers consisting of PVAc and Cu powder formed between the paper mesh and counter electrode.

formation was induced not only at one hole but at multiple holes. Even though the supply of fiber ingredient continued at single position of paper mesh as shown in Figure 8, the accumulated fiber ingredient flew to the multiple holes of paper mesh, eventually resulting in the fiber formation from those multiple holes. As clearly seen in Figure 8, a number of short fibers were formed between the multiple holes of paper mesh and counter electrode.

Use of this paper mesh technique did not result in so successful formation of PVAc fibers. It must be due to the less volatile property of DMF compared with THF not due to the ineffectiveness of this paper mesh technique. In fact, even a composite fiber consisting of PVAc dissolved in THF instead of DMF and Cu powder was well formed under the condition that the voltage was ~ 9 kV and the gap between the needle and counter electrode was ~ 4 cm. The ingredient of this composite was prepared by dissolving 2.5 g of PVAc into 12.5 g of THF and subsequent addition of 1 g Cu powder to it. Figure 10 shows those fibers formed between the paper mesh and counter electrode, where the paper mesh is removed in this photo. Not all the fibers were separately formed, but we could come by a certain quantity of separately formed composite fibers.

5. CONCLUSIONS

Simply placing a paper mesh between the needle tip and collector resulted in the separately formed electro-spun polymer fibers. This technique was even applicable for the fabrication of polymer-metal composite fibers. Hence, this technique must be useful for separately fabricating the broad range of different kinds of fibers.

ACKNOWLEDGMENT

This research was performed by the financial support of the Air Force Research Laboratory under the agreement number, AOARD-07-4012, FA4869-08-1-4044 and FA2386-09-1-4033. I'd like to express my gratitude Dr. Tadayoshi Ihara (Gifu Univ.) for his teaching and helping us operate the high-speed camera.

REFERENCES

1. Kowalewski, T. A., S. Blonski, and S. Barrel, "Experiments and modeling of electrospinning process," *Bulletin of the Polish Academy of Sciences Technical Sciences*, Vol. 53, 385–394, 2005.
2. Supaphol, P. and S. Chuangchote, "On the electrospinning of poly (vinyl alcohol) nanofiber mats: A revisit," *Journal of Applied Polymer Science*, Vol. 108, 969–978, 2008.
3. Tomaszewski, W. and M. Szadkowski, "Investigation of electrospinning with the use of a multi-jet electrospinning head," *Fibres & Textiles in Eastern Europe October/December*, Vol. 13, 22–26, 2005.
4. Zhang, Y. Z., J. Venugopal, Z.-M., Huang, C. T. Lim, and S. Ramakrishna, "Crosslinking of the electrospun gelatin nanofibers," *Polymer*, Vol. 47, 2911–2917, 2008.

Profile Measurement for Micro-optical Component Using Lensless Fourier Digital Holography

Yunxin Wang, Dayong Wang, Yan Li, Jie Zhao, Yizhuo Zhang,
Yuhong Wan, and Zhuqing Jiang

College of Applied Sciences, Beijing University of Technology, Beijing 100124, China

Abstract— Digital holography imaging method is presented for three-dimensional profile detection of the micro-optical component. Firstly, digital holography configuration based on off-axis lensless Fourier transform (LFT) is designed, and then the profile of the hexagonal micro-lens array is detected by Fresnel reconstruction and phase-unwrapping algorithms. Meanwhile, the white light interferometer is applied to measure the profile parameters of the micro-lens array. The results indicate that the profile parameters obtained above are in good agreement, which demonstrate that the LFT based digital holography provides an effective method for 3D profile detection of the micro-optical component.

1. INTRODUCTION

Three-dimensional (3D) profile measurement for micro-optical component has attracted remarkable attention recently for its applications in many fields such as laser beam shaping, optical imaging and optical transmission. The topography characteristics of the micro-optical component determine its optical performance directly, so the 3D profile measurement plays an important role in the quality evaluation and machining guidance. The traditional imaging methods like differential interference contrast (DIC), Zernike phase contrast microscope can't acquire the quantitative information of the phase. Therefore, many full-field and quantitative phase imaging techniques are developed recently including Fourier phase microscopy (FPM), Hilbert phase microscopy (HPM), diffraction phase microscopy (DPM) and digital holographic microscopy (DHM) [1–5]. Among them, digital holography has several special advantages. DHM doesn't need complex scanning configuration and possesses a simple structure. DHM can retrieve quantitative information of object wavefront from a single digital hologram, which allows it to achieve the real-time detection. Furthermore, since the numerical focalization algorithm can be implemented by the wave propagation theory, DHM does not demand to record hologram in the focus image plane of the object.

In this contribution, digital holography based on lensless Fourier transform (LFT) is applied to the profile inspection of the micro-optical component. The hexagonal micro-lens array is taken as the test sample, and its profile parameters are acquired by Fresnel reconstruction and phase-unwrapping algorithms. Moreover, the white light interferometer is used for the profile measurement of the hexagonal micro-lens array. The experimental results from the above methods are compared to verify the feasibility of the LFT based digital holography.

2. LFT BASED DIGITAL HOLOGRAPHY

LFT based digital holography can make full use of the bandwidth of the detector, and the configuration structure is quite simple. The schematic of the experimental configuration for the inspection of micro-optical component is illustrated in Fig. 1, which is based on the Mach-Zehnder interferometer. The input laser is divided into two parts by a polarization beam splitter (PBS), one beam goes through the transparent or semi-transparent test sample as the object beam, and another beam is expanded and filtered to produce a finer reference point source. The object and reference beams are combined at BS, the interference pattern is recorded by CCD detector, and then the image data are sent to computer by a collection card. The intensity ratio between the reference and object beams can be adjusted by the combination of $\lambda/2$ half-wave plate and PBS to improve the image quality of digital hologram.

In lensless Fourier transform digital holography, the cross section of the test sample and the reference point source should be located in the same plane [6]. The hologram with $N \times N$ pixels in the CCD plane can be expressed by the combination of the four terms:

$$I(x_H, y_H) = |R(x_H, y_H)|^2 + |O(x_H, y_H)|^2 + R(x_H, y_H)O^*(x_H, y_H) + R^*(x_H, y_H)O(x_H, y_H) \quad (1)$$

where $O(x_H, y_H)$ and $R(x_H, y_H)$ are the complex amplitude distributions for object and reference waves in the hologram plane, and $*$ denotes the complex conjugate operator.

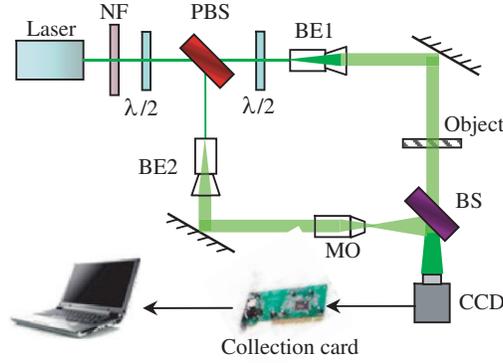


Figure 1: Schematic of imaging configuration.

Numerical reconstruction of the off-axis lensless Fourier hologram is performed by the Fresnel reconstruction algorithm, and the Fresnel diffraction integral is defined as:

$$O(x_I, y_I) = \frac{\exp(jkz_0)}{j\lambda z_0} \iint_{-\infty}^{\infty} I(x_H, y_H) R(x_H, y_H) \exp\left\{ \frac{jk}{2z_0} \left[(x_I - x_H)^2 + (y_I - y_H)^2 \right] \right\} dx_H dy_H \quad (2)$$

where $O(x_I, y_I)$ is the reconstructed wave field for the test sample in the image plane, z_0 is the distance between the hologram plane and the reconstructed image plane.

Equation (2) can be represented as:

$$O(x_I, y_I) = \frac{\exp(jkz_0)}{j\lambda z_0} \exp\left\{ \frac{jk}{2z_0} (x_I^2 + y_I^2) \right\} F \left\{ I(x_H, y_H) R(x_H, y_H) \exp\left[\frac{jk}{2z_0} (x_H^2 + y_H^2) \right] \right\} \quad (3)$$

where F is the two-dimensional Fourier transform.

The discrete representation of Eq. (3) can be calculated by:

$$O(m, n) = \frac{\exp(jkz_0)}{j\lambda z_0} \exp\left\{ \frac{jk}{2z_0} (m^2 \Delta x_I^2 + n^2 \Delta y_I^2) \right\} FFT \left[I(t, s) R(t \Delta x_H, s \Delta y_H) \exp\left\{ \frac{jk}{2z_0} (t^2 \Delta x_H^2 + s^2 \Delta y_H^2) \right\} \right] \quad (4)$$

where FFT is the fast Fourier transform, Δx_H , Δy_H are the sampling intervals of the hologram plane, and m, n, t, s are integers ($N/2 \leq m, n, t, s \leq N/2 - 1$). The relations of the sampling intervals between the image and hologram planes can be described as $\Delta x_I = \lambda z_0 / N \Delta x_H$ and $\Delta y_I = \lambda z_0 / N \Delta y_H$.

Therefore, the intensity and phase distribution of the test sample can be acquired by

$$I(m, n) = |O(m, n)|^2 = \text{Re}^2 |O(m, n)| + \text{Im}^2 |O(m, n)| \quad (5)$$

and

$$\phi(m, n) = \arctan \frac{\text{Im} [O(m, n)]}{\text{Re} [O(m, n)]} \quad (6)$$

where $\text{Re}[O(m, n)]$ and $\text{Im}[O(m, n)]$ denote the real and imaginary parts of the object complex amplitude.

The phase values obtained by Eq. (6) is limited in the range of $(-\pi, +\pi)$ for the principle of the arctan function, so the phase image will contain 2π discontinuities when the optical depth of test sample is greater than the wavelength λ . The least-squares phase-unwrapping algorithm is applied to acquire the accurate phase information [7].

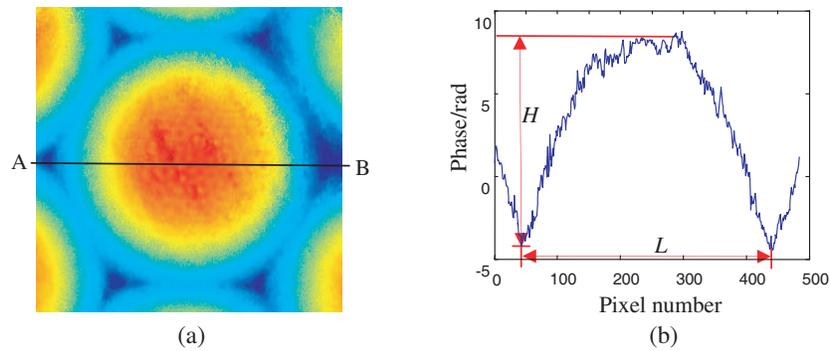


Figure 2: Measurement result of hexagon micro-lens array using LFT based digital holography configuration. (a) Phase image. (b) Phase distribution along the line AB.

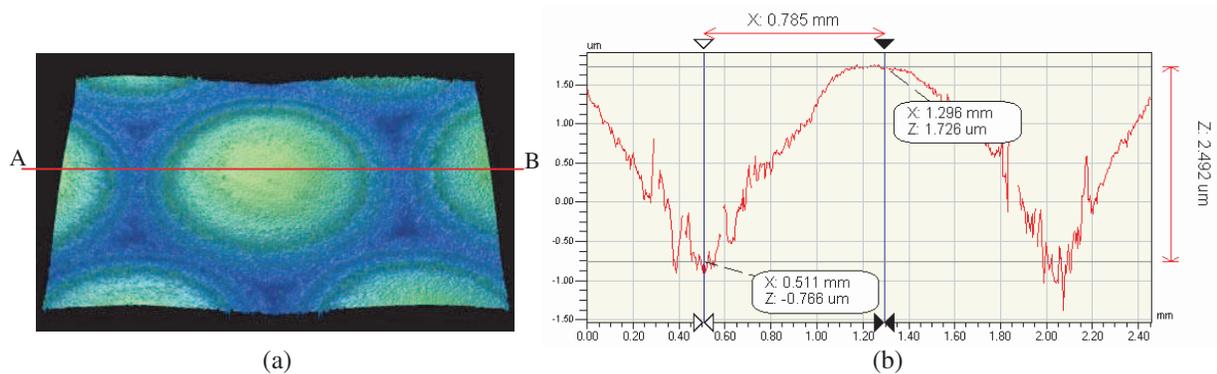


Figure 3: Measurement result of hexagon micro-lens array using white light interferometer. (a) 3D profile image. (b) Profile distribution along the line AB.

3. EXPERIMENTS AND RESULTS

The micro-optical component mainly includes diffractive optical element (DOE), micro-lens array (MLA) and binary optics element (BOE). MLA based on the refraction or diffraction principle is composed by a series of micro lens with the size from several microns to hundred microns, which is characteristic by small size, little transmission loss and convenient large-scale manufacture. We take the hexagon micro-lens array as an example to verify the effectiveness of the digital holography technology in the profile detection.

When the light wave passes through a semi-transparent or transparent sample, its phase will carry the profile information. The height of the object H can be calculated by:

$$H = \frac{\lambda\phi}{2\pi(n_l - n_a)} \quad (7)$$

where ϕ is the detected phase, n_l is the specimen refractive index, and n_a is the medium refractive index. The medium is commonly air, thus n_a is approximate to 1 in our experiment.

The profile of MLA, especially the micro-lenses height H and length L , influenced its imaging quality severely. The measurement result of LFT based digital holography configuration is shown in Fig. 2, and the acquired height H and length L of the tested hexagon micro-lens array are $2.424 \mu\text{m}$ and 1.594 mm respectively.

Meanwhile, the profile of the same hexagon micro-lens array is measured using the white light interferometer, and the result is illustrated in Fig. 3. The according height H and length L are $2.492 \mu\text{m}$ and 1.570 mm respectively. Therefore, the measurement results of above two methods show a good consistency.

4. CONCLUSION

Digital holography imaging based on off-axis lensless Fourier transform (LFT) is presented for 3D profile measurement of the micro-optical component. The hexagon micro-lens array is taken as a test object and its profile is acquired. Meanwhile, the white light interferometer is applied to

detect the character of the same hexagon micro-lens array, and the results indicate that the profile parameters obtained by these two methods are in good agreement. This demonstrates that LFT based digital holography is feasible and effective for the 3D profile detection, and it will have wide applications in quality evaluation and machining guidance for the micro-optical component.

ACKNOWLEDGMENT

This research was supported partly by the Science Foundation of Education Commission of Beijing under grant No. KZ200910005001 and Funding Project for Academic Human Resources Development in Institutions of Higher Learning under the Jurisdiction of Beijing Municipality.

REFERENCES

1. Lue, N., W. Choi, G. Popescu, T. Ikeda, R. R. Dasari, et al., "Quantitative phase imaging of live cells using fast Fourier phase microscopy," *Applied Optics*, Vol. 46, No. 10, 1836–1842, 2007.
2. Ikeda, T., G. Popescu, R. R. Dasari, and M. S. Feld, "Hilbert phase microscopy for investigating fast dynamics in transparent systems," *Optics Letters*, Vol. 30, No. 10, 1165–1168, 2005.
3. Popescu, G., T. Ikeda, R. R. Dasari, and M. S. Feld, "Diffraction phase microscopy for quantifying cell structure and dynamics," *Optics Letters*, Vol. 31, No. 6, 775–777, 2006.
4. Carl, D., B. Kemper, G. Wernicke, et al., "Parameter-optimized digital holographic microscope for high resolution living-cell analysis," *Applied Optics*, Vol. 43, No. 36, 6536–6544, 2004.
5. Colomb, T., F. Charrière, J. Kühn, et al., "Advantages of digital holographic microscopy for real-time full field absolute phase imaging," *Proc. of SPIE*, Vol. 6861, 1–10, 2008.
6. Wang, H., D. Wang, J. Zhao, et al., "Simple and robust digital holography for high-resolution imaging," *Chinese Optics Letters*, Vol. 6 No. 3, 165–167, 2008.
7. Quiroga, J. A. and E. Bernabeu, "Phase-unwrapping algorithm for noisy phase-map processing," *Applied Optics*, Vol. 33, No. 29, 6725–6731, 1994.

A Novel Data Transmission Security via a Noisy Channel Using a Microring Resonator System

T. Threepak¹, S. Mitatha¹, and P. P. Yupapin²

¹Hibrid Computing Research Laboratory, Faculty of Engineering
King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand
²Advance Research Center for Photonics, Department of Applied Physics
King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand

Abstract— A novel encryption methods using microring resonator is presented. By using the proposed system, the confidentiality of information in optical communication can be formed. In this paper, chaotic noise sophisticates irreversible and unpredictable in nonlinear microring resonator that is used as the encrypted function. In our successfully simulation results, the noisy signals/channels created by the microring resonator are combined with the encoded information to protect the man-in-the-middle attacks (decipher). By using the extremely small processing device, the propose design can be easily applied to secure any form of communication in wireless network, mobile communication network and military applications with low power consumption and very high-speed procedures.

1. INTRODUCTION

Chaotic behavior has been studied in several areas in either theory or applications such as mathematics [1], physics [2], electronics [3], and communications [4]. Furthermore, the chaotic noise has been found useful in several areas of applications such as electronic communication [5], switching and control [6], and optical communication [7]. Where the present use of the benefit of such a nonlinear behavior, especially, in the military purpose for when the information is required to be kept confidential. In general, the nonlinearity of the system involves the behaviors such as chaos, bistability and bifurcation, which can be generated in the electronic circuit and optical fiber [8, 9], laser system [10] and optical waveguide [11]. One application is the use of the device known as a micro ring resonator, which can be formed by a waveguide or a fiber optic which has shown a very promising application when the noisy signal is generated and multiplexed to the required information, where the information can be retrieved by the required users. Moreover, this technique becomes great interest when such a device is fabricated within the range of a micrometer scale [12, 13]. The published works of the secure communication systems based on chaos in a micro ring resonator were proposed by references [14, 15]. In this paper, we utilize chaotic effect in microring resonator as encrypted modules in symmetric key encryption.

2. NOISY SIGNAL GENERATION

A simple device schematic diagram is as shown in Fig. 1, when the light from a monochromatic light source is launched into a ring resonator. The relation of input light field and output light field is shown in Equation (1) that given by [16]. In addition to, the optical fields E_{r1} and E_{r2} represent the right and left hands circulations in a ring resonator, correspondingly. This equation indicates that a ring resonator in the particular case is very similar to a Fabry-Perot cavity, which has an input and output mirror with a field reflectivity, $1 - \kappa$, and a fully reflecting mirror. Where n_o

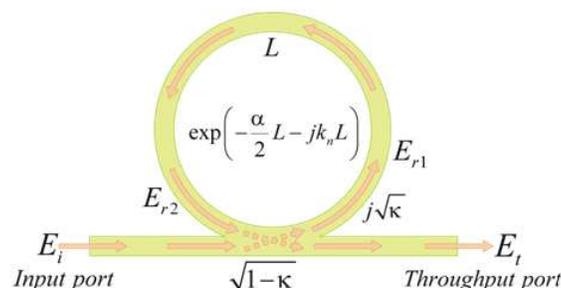


Figure 1: A Schematic diagram of the micro ring resonator.

and n_2 are the linear and nonlinear refractive indices, and the coupling coefficient is κ . Where $x = \exp^{-\frac{\alpha L}{2}}$ represents a roundtrip losses coefficient, $\phi_0 = kLn_0$ and $\phi_{nl} = kLn_2|E_1|^2$ are the linear and nonlinear phases that shift respectively; $k = \frac{2\pi}{\lambda}$ is the wave propagation number in the vacuum and γ is loss in medium.

$$\left| \frac{E_{\text{out}}(t)}{E_{\text{in}}(t)} \right|^2 = (1 - \gamma) \left[1 - \frac{(1 - (1 - \gamma)x^2)\kappa}{(1 - x\sqrt{1 - \gamma}\sqrt{1 - \kappa})^2 + 4x\sqrt{1 - \gamma}\sqrt{1 - \kappa}\sin^2\left(\frac{\phi}{2}\right)} \right] \quad (1)$$

This nonlinear behavior of light traveling in a single ring resonator (SRR) is described by Yupapin et al. [18]. Where the parameters of the system were fix to be $\lambda_0 = 1.55 \mu\text{m}$, $n_0 = 1.54$, A_{eff} is the effective mode core area of the fiber, $A_{\text{eff}} = 30 \mu\text{m}^2$, the waveguide ring resonator loss (α) is 0.02 dBmm^{-1} . The fractional coupler intensity loss (γ) is 0.1, and $R_1 = 10 \mu\text{m}$. The coupling coefficient of the fiber coupler is fixed to $\kappa = 0.0225$. The nonlinear refractive indices range from $n_2 = 1.4 - 2.2 \times 10^{-15} \text{ m}^2/\text{W}$ [10], and the 20,000 iterations of roundtrips inside the optical fiber ring was calculated by assume that $\phi_L = 0$ for simplicity; By these configurations in previous research, bifurcation and chaotic signals are generated [17], which can be use as hashing and encryption box in any cryptographic model as shown in next section.

3. ENCRYPTION AND DECRYPTION

In secure communication, encryption is method to ensure that transmission data between sender and receiver still unrevealed. This prevents attackers from tapping information, exploit, modify, or obtain undesirable outcome from communication channel. Encryption algorithms accomplish this by accepting as input both message and secret key to produce a ciphertext that unable to interpret without correct secret key and encryption algorithm. These secret messages are transmitted to the other party, who can then decrypt the ciphertext into original data by using similar decryption algorithm and same secret key. Otherwise, the ciphertext is unable to decipher into original message. Because characteristic of encrypting algorithms are much like characteristic of chaos such as sophisticates, irreversible, and unpredictable, encryption method can be buit by using microring resonator.

To secure transmission via noisy channel, chaos signals are generated by passing key signal into microring resonator and then compose with encoded data to generate ciphertext. At the receiver, users who know secret key are able to both generate the same chaos pattern and filter chaos signal from ciphertext to get encoded data. Subsequently, by using suitable decoder, original data signal are produced. Schemetic of these methods are shown in Fig. 2. In this process, suitable procedures of encoder and decoder module for continuous signal are differentiation and integration function, respectively, to maintain the chaotic pattern of encrypted data in communcation channel. When pass through communication channel, data signal always scrammble within noisy signal. Any sniffed data between communication links are unable to interpret without same key information, correct microring specification, and suitable encoding method. This method always maintains confidentiality along the transmission channel.

In our simulation, 0.25 MHz, 0.05 mW amplitude sine wave with 4 mW average power key signal, as shown in Fig. 3(a), come into waveguide of $10 \mu\text{m}$ microring resonator to generate hash signal. In addition, 1 W, 0.25 MHz sine wave with 4 W average power data signal, as shown in Fig. 3(b), come

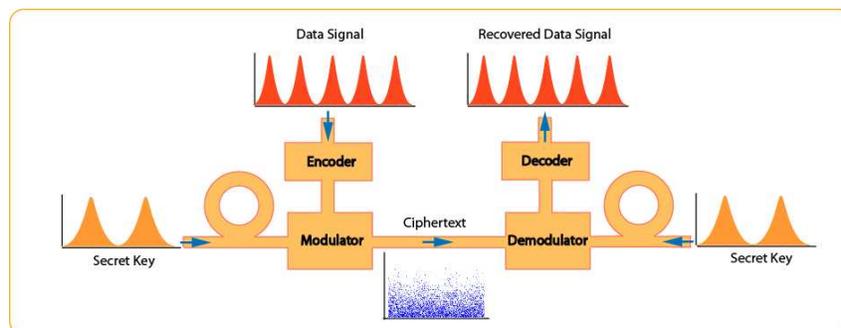


Figure 2: Symmetric key cryptography method.

into encoder to produce encoded signal. Lambda of waveguide is $1.55 \mu\text{m}$. Coupling coefficient is equal to 0.212. Linear and nonlinear coefficient are equal to 1.2 and $3.8 \times 10^{-13} \text{ cm}^2/\text{W}$. Effective core area is equal to $30 \mu\text{m}^2$. Both encoded signal and chaos is combine to produce encrypted signal as shown in Fig. 3(c). At receiver side, both similar key signal and microring resonator configuration used to generate same chaos for decipher the ciphertext and obtain encoded data signal. Then, launch encoded data into decoding module to find original data, as shown in Fig. 3(d).

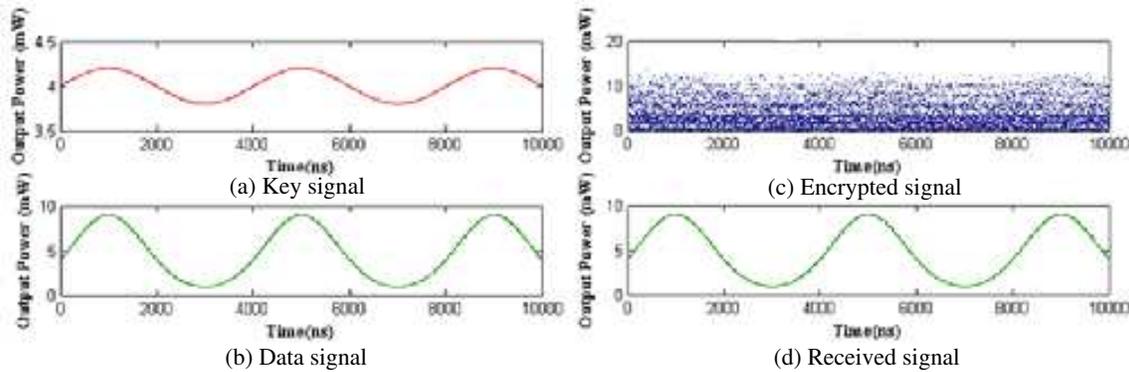


Figure 3: Encryption simulation result. (a) Key signal. (b) Data signal. (c) Encrypted data. (d) Received data.

4. IMAGE STEGANOGRAPHY IN NOISY CHANNEL

Steganography is security technique to hide information in other data for specific purposes such as owner identification, integrity inspection, or data verification. To conceal image into noisy channel, image information must be encoded into proper structure and couple with noisy signal before transmit to receiver. Outcome signals are still in noisy pattern that can be used in other purposes such as scramble other information, random number generation, generate message authentication code, or encryption. Along the communication channel, sophisticated pattern that hard to identify of noisy signal protects composed information. At receiver side, chaos signal are extracted by similar noise pattern, generated by using equivalent secret key and microring resonator configuration, to obtain encoded image. Afterward, decoder is used to release hiding image.

In our simulation, following configurations are used in this research. 0.25 MHz, 0.05 mW amplitude sine wave with 4 mW average power key signal, as shown in Fig. 3(a), come into waveguide of $10 \mu\text{m}$ microring resonator to generate noisy channel. Lambda of waveguide is $1.55 \mu\text{m}$. Coupling coefficient is equal to 0.212. Linear and nonlinear coefficient are equal to 1.2 and $3.8 \times 10^{-13} \text{ cm}^2/\text{W}$. Effective core area is equal to $30 \mu\text{m}^2$. 8-bits-gray-scale Lenna image with 250×250 pixels encodes into bit stream of Return-to-Zero signals and coupling with noisy signals which are generated by passing key signals into microring resonator. Then, noisy signals with image steganography are transfer into communication channel. At receiver side, both similar key signal and microring resonator configuration used to generate same noisy signal for filtering to obtain encoded data. Then,

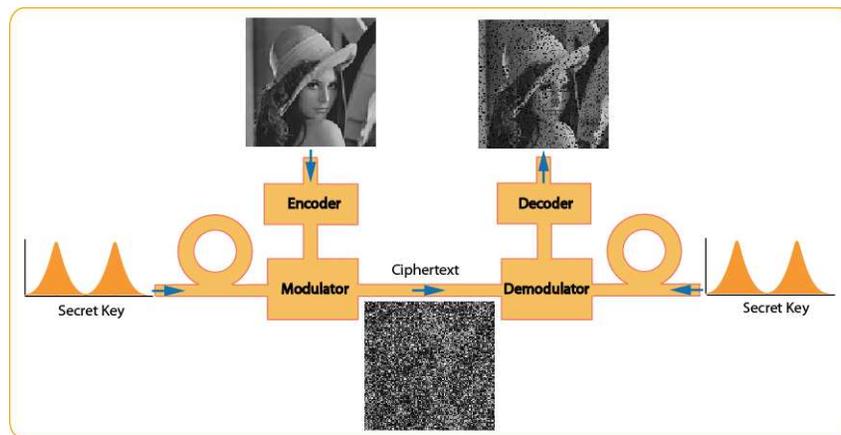


Figure 4: Image steganography in noisy channel.

launch encoded data into decoding module to find original image. The result of our simulation, Lena image at sender, communication link, and receiver, are illustrated in Fig. 4.

5. CONCLUSIONS

In conclusion, we have shown that secure channel and image steganography in optical communication can be implemented by using microring resonator. These operations are success by using chaos phenomenon in microring resonator as encryption function in cryptography mathematics. In this model, by using microring and key signal to generate noisy channel, data sending from sender are secure in transmission and also decipher at the receiver side. By using similar method with image data, steganographic process can be implemented. The key advantages of the proposed system are simple to implement in single chip and easily applied to secure any form of communication in wireless network, mobile communication network and military applications with low power consumption and very high-speed procedures.

REFERENCES

1. Ja, Y. H., "Multiple bistability in an optical-fiber double-ring resonator utilizing the Kerr effect," *IEEE J. Quantum Electron.*, Vol. 30, 329–333, 1994.
2. Company, J., F. J. Praile-Pelaez, and M. A. Muriel, "Optical bistability and differential amplification in nonlinear fiber resonators," *IEEE J. Quantum Electron.*, Vol. 30, 2578–2588, 1994.
3. Kittel, A., J. Parisi, and K. Pyragas, "Generalized synchronization of chaos in electronic circuit experiments," *Physica. D: Nonlinear Phenomena*, Vol. 112, No. 3–4, 459–471, 1998.
4. Ogusu, K. and S. Yamamoto, "Nonlinear fiber Fabry-Perot resonator using thermo-optic effect," *J. Lightwave Technol.*, Vol. 11, 1774–1781, 1993.
5. Steele, A. L., "Optical bistability, instabilities and power limiting behavior from a dual nonlinear optical fiber loop mirror resonator," *Optics Comm.*, Vol. 236, 209–218, 2004.
6. Felber, F. S. and J. H. Marburger, "Theory of nonresonant multistable optical devices," *Appl. Phys. Lett.*, Vol. 28, 731–733, 1976.
7. Marburger, J. H. and F. S. Felber, "Theory of a lossless nonlinear Fabry-Perot interferometer," *Phys. Rev. A*, Vol. 17, 335–342, 1978.
8. Bischofberger, T. and Y. R. Shen, "Theoretical and experimental study of the dynamics behavior of a nonlinear Fabry-Perot interferometer," *Phys. Rev. A*, Vol. 19, 1169–1176, 1979.
9. Ikeda, K., H. Daido, and O. Akimoto, "Optical turbulence: Chaotic behavior of transmitted light from a ring cavity," *Phys. Rev. Lett.*, Vol. 45, 709–712, 1980.
10. Nakatsuka, H., S. Asaka, H. Itoh, K. Ikeda, and M. Matusuoka, "Observation of bifurcation to chaos in an all-optical bistable system," *Phys. Rev. Lett.*, Vol. 50, 109–112, 1983.
11. Gibbs, H. M., F. A. Hopf, D. L. Kaplan, and R. L. Shoemaker, "Observation of chaos in optical bistability," *Phys. Rev. Lett.*, Vol. 46, 474–477, 1981.
12. Zhang, Y., B. Martin, K. P. Huy, D. Amans, and P. Benech, "Ultra-compact microdisk resonator filters on SOI substrate," *Optics Express*, Vol. 14, No. 26, 12814–12821, 2000.
13. Van, V., T. A. Ibrahim, P. P. Absil, J. V. Hryniewicz, F. G. Johnson, R. Grover, and P.-T. Ho, "Optical signal processing using nonlinear semiconductor microring resonators," *IEEE J. Sel. Top. Quant.: Nonlinear Optics*, Vol. 8, 705–713, 2002.
14. Garcia-Ojalvo, J. and R. Roy, "Spatiotemporal communication with synchronized optical chaos," *Phys. Rev. Lett.*, Vol. 86, 5204–5207, 2001.
15. Bollt, E., Y.-C. Lai, and C. Grebogi, "Coding, channel capacity, and noise resistance in communicating with chaos," *Phys. Rev. Lett.*, Vol. 79, 3787–3790, 1997.
16. Yupapin, P. P., W. Suwancharoen, and S. Suchat, "Nonlinearity penalties and benefits of light traveling in a fiber optic ring resonator," *Optik*, Vol. 120, 216–221, 2009.
17. Yupapin, P. P. and W. Suwancharoen, "Chaotic signal generation and cancellation using a micro ring resonator incorporating an optical add/drop multiplexer," *Opt. Commun.*, Vol. 280, 343–350, 2007.
18. Yupapin, P. P., P. Saeung, and C. Li, "Characteristics of complementary ring-resonator add/drop filters modeling by using graphical approach," *Opt. Commun.*, Vol. 272, 81–86, 2007.

An Experimental Design for Reversed Cherenkov Radiation in a Double-negative-metamaterial-loaded Waveguide

Zhaoyun Duan¹, Xutong Mao¹, Jucheng Lu¹, Yanyu Wei¹, Yubin Gong¹,
Wenxiang Wang¹, Bae-Ian Wu², and Min Chen³

¹Vacuum Electronics National Laboratory, School of Physical Electronics

University of Electronic Science and Technology of China, Chengdu 610054, China

²Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

³Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Abstract— We have proposed an experimental design for verifying the reverse Cherenkov radiation (RCR) in a circular waveguide partially filled by double-negative metamaterials (DNMs). These media are composed of different materials such as the metallic strips for the SRRs and rods/dielectric materials for holding the strips. They can be characterized by simultaneously negative permittivity and permeability tensors at a narrow frequency band. And they are suitable for quasi-TM mode. Thus, the electromagnetic waves can effectively interact with an electron beam and then the RCR can be produced and detected by two power output windows. Currently, the related work on the manufacture and assembling as well as measurement of the DNMs is in progress.

1. INTRODUCTION

A left-handed medium with simultaneously negative permittivity and permeability was first introduced by Veselago [1]. Several unique electromagnetic wave phenomena in this medium, such as the negative refractive index, the reversed Cherenkov radiation (RCR), the reversed Doppler effect, and even reversal of radiation pressure to radiation tension. However, unfortunately, up to now, the RCR has not been directly verified in the experiments [2–4]. We have been studying the RCR in DNMs theoretically [5–7], and are making an effort to design an experiment for proving the RCR in the double-negative metamaterials (DNMs). Now, we have proposed an experimental design for the RCR as follows.

2. EXPERIMENTAL DESIGN

In this design process, there are three main steps. Firstly, the DNMs which are suitable for the RCR have been designed. In order to illustrate the SRR structure more clearly, we show it in Figures 1(a) and (b), a metamaterial consisting of SRRs and wires is fabricated using a conventional printed both sides of an FR-4 dielectric substrate. The FR-4 board has a thickness of 0.25 mm and the relative permittivity is 4.4 and the conductance is 0.0068 S/m. The copper SRRs and wires are positioned on opposite sides of the substrate, modeling the structures typically produced by lithographic circuit board techniques. The copper thickness is 0.03 mm, the width of the wire is 1 mm, the inner ring radius of the SRR is 2.9 mm and both rings have a line width of 0.8 mm, the gap in each ring is

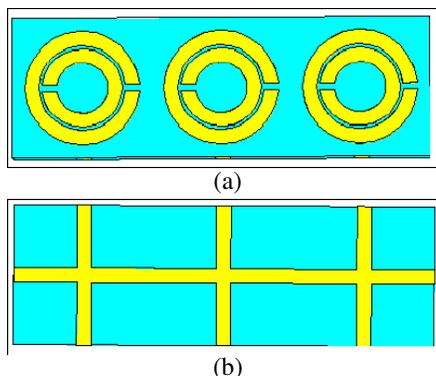


Figure 1: The (a) front and (b) back views of a unit cell.

0.3 mm, the gap between the inner and outer rings is 0.3 mm, and the periodicity of the array in three directions is 10 mm, 10 mm and 3 mm, respectively.

As we know [6, 7], this artificial medium partially loaded in the circular waveguide is demanded at least three negative parameters: $\varepsilon_{r\rho}(\omega)$, $\varepsilon_{rz}(\omega)$, and $\mu_{r\theta}(\omega)$ for the RCR. Therefore, we can design a kind of the DNMs shown in Figure 2. In order to keep the homogeneousness of the medium, we divide the material into three layers in the radial direction and keep the density of the unit cell every layer constant.

We have derived the RCR theoretically [5–7]: $\text{Re}(\varepsilon_{r\rho}\mu_{r\theta}) > 1/\beta^2 = 1/(v_0/c)^2$ where v_0 is the speed of electrons in vacuum and c is the light speed in free space. Due to the anode voltage of the electron gun less than 20 kv, we must design the special metamaterials ($\varepsilon_{r\rho} = -6.8$, $\mu_{r\theta} = -4.2$), which is a big challenge to us. On the other side, in order to match characteristic impedance with free space impedance well, $\varepsilon_{r\rho} = \mu_{r\theta}$ must be satisfied as possible as we can. In terms of above demands we have designed a metamaterial shown in Figure 2. We have obtained the needed electromagnetic parameters numerically which are shown in Figures 4 and 5. From the numerical results we find double negative region is between 10 to 11GHz, and we may observe the RCR in this region.

Secondly, we can design the two output windows for detecting the RCR, which are shown in Figure 3, just represented by the two holes in the waveguide. Third, an electron gun, a focusing system, and a collector should be designed. Thus, if the electron gun is placed in the left of Figure 3, then we can determine whether the RCR occur. If we detect the microwave at the double negative region in the left output window and can not find it at the same frequency band in the right output window, then we can confirm the RCR happen in the DNMs. In addition, the experimental setup should be in the vacuum state. All the material used in the experiment should be considered carefully.

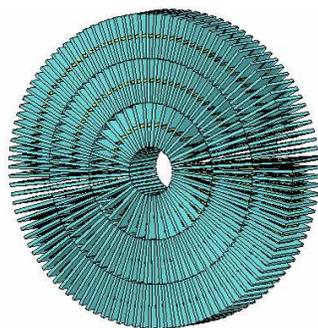


Figure 2: The DNMs suiting for the RCR in the circular waveguide.

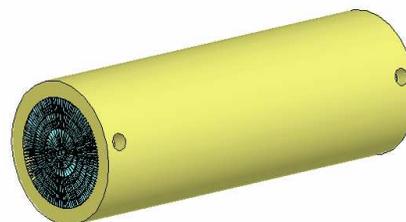


Figure 3: A scheme of verifying the RCR in the circular waveguide.

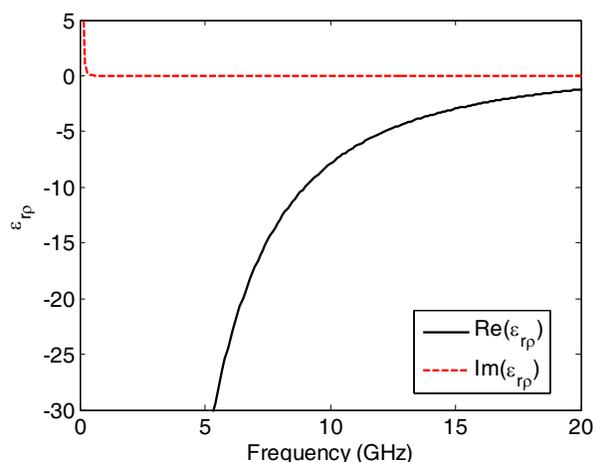


Figure 4: The relative permittivity for the proposed DNMs.

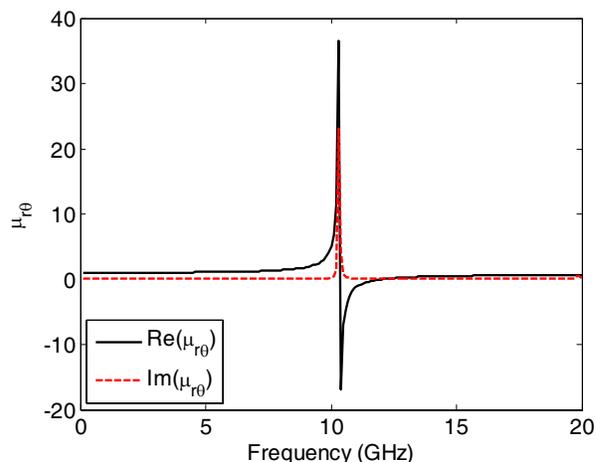


Figure 5: The relative permeability for the proposed DNMs.

3. CONCLUSIONS

In this paper, we have proposed an effective method to design the DNMs for verifying the RCR. The DNMs with at least three negative constitutive parameters ($\varepsilon_{r\rho}(\omega)$, $\varepsilon_{rz}(\omega)$ and $\mu_{r\theta}(\omega)$) are just suitable for the RCR. In order to be able to operate in vacuum the system must have a high obturation. For this reason we should be carefully in the experiment. In addition, we should increase the anode voltage of the electron gun as possible as we can, then we can have more opportunities to observe the RCR in practice. Due to the complexity of the experimental setup, we are designing and manufacturing the DNMs. Much more work is underway.

ACKNOWLEDGMENT

This work was supported by National Natural Science Foundation of China (Grant Nos. 60971031, 60601007, and 60532010) and Sichuan Youth Foundation.

REFERENCES

1. Veselago, V. G., "The electrodynamics of substances with simultaneously negative values of ε and μ ," *Soviet Physics Uspekhi*, Vol. 10, No. 4, 509–514, 1968.
2. Duan, Z. Y., B.-I. Wu, J. Lu, J. A. Kong, and M. Chen, "Research progress in reversed cherenkov radiation in double-negative metamaterials," *Progress In Electromagnetics Research*, PIER 90, 75–87, 2009.
3. Antipov, S., L. Spentzours, W. Gai, M. Conde, F. Franchini, R. Konecny, W. Liu, J. G. Power, Z. Yusof, and C. Jing, "Observation of wakefield generation in left-handed band of metamaterial-loaded waveguide," *Journal of Applied Physics*, Vol. 104, 014901, 2008.
4. Shehegolkov, D. Y., A. K. Azad, J. F. O'Hara, and E. I. Smirnov, "A proposed measurement of the reverse cherenkov radiation effect in a metamaterial-loaded circular waveguide," *34th International Conference on Infrared, Millimeter, and Terahertz Waves*, 21–25, Busan, Korea, Sep. 2009.
5. Duan, Z. Y., B.-I. Wu, J. Lu, J. A. Kong, and M. Chen, "Reversed cherenkov radiation in a waveguide filled with anisotropic double-negative metamaterials," *Journal of Applied Physics*, Vol. 104, 063303, 2008.
6. Duan, Z. Y., B.-I. Wu, J. Lu, J. A. Kong, and M. Chen, "Cherenkov radiation in anisotropic double-negative metamaterials," *Optics Express*, Vol. 16, No. 22, 18479–18484, 2008.
7. Duan, Z. Y., B.-I. Wu, J. Lu, J. A. Kong, and M. Chen, "Reversed cherenkov radiation in unbounded anisotropic double-negative metamaterials," *J. Phys. D: Appl. Phys.*, Vol. 42, 185102, 2009.
8. Li, Z. F., K. Aydin, and E. Ozbay, "Determination of the effective constitutive parameters of bianisotropic metamaterials from reflection and transmission coefficients," *Physical Review E*, Vol. 79, 026610, 2009.

A Novel Broadband Metamaterial Resonator with Negative Permittivity

J. Zhang and Z. R. Hu

Microwave and Communication Systems Group, School of Electrical and Engineering
Faculty of Engineering and Physical Sciences, University of Manchester, UK

Abstract— It is well known that the properties of a metamaterial are usually determined from the geometrical structure and material parameters rather than composition. One of the problems of resonant types of metamaterials is that they usually have narrow operating bandwidth. In this paper, a novel metamaterial resonator with desired broadband negative permittivity was proposed.

1. INTRODUCTION

Metamaterials has led to considerable interests in the electromagnetism area in the first decade of 21st century. This is due to some of its unique electromagnetic features that may provide solutions for solving current technological limitations. Volumetric epsilon-negative (ENG) metamaterials [1], metamaterials with negative permittivity and positive permeability, generally consist of anisotropic resonant particles. The ENG particles have to be planar such as that reported in [2]. Based on the resonant behavior of the particles, there is only a narrow frequency band in which the effective permittivity of the medium is negative. Compared to the non-resonant structures of metamaterials such as LH planar transmission lines, the frequency band of resonant types of metamaterials is much narrower. This is a limitation in applications of ENG metamaterials.

2. PLANAR BROADBAND RESONATORS FOR ENG METAMATERIALS

This paper deals with the evolution of an electric-field-coupled resonator for constructing ENG metamaterials and presents a new planar broadband resonator with negative permittivity. Figure 1(a) presents the geometric design of the planar broadband resonator. Figure 1(b) shows sample of a metamaterial built by a single layer of planar broadband resonators.

This unit cell resonator is composed of three different metal rings on top of a dielectric substrate. These rings are copper strips with 0.25-millimeter width and the outside radiuses are 0.5-millimeter, 1.0-millimeter and 1.5-millimeter respectively. The substrate is FR4 with 0.203-millimeter thickness. The resonator presented here exhibits a strong negative permittivity response at a wide band.

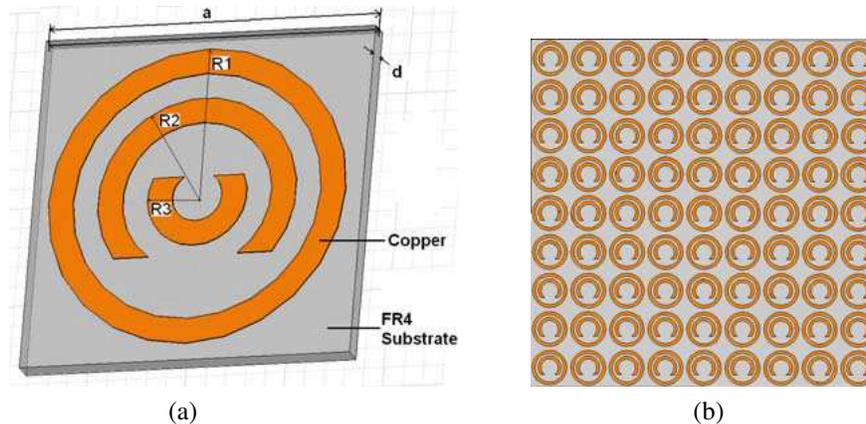


Figure 1: Structure of planar broadband resonator used in ENG metamaterials (Dimensions of the unit cell are: $a = 3.333$ mm, $d = 0.203$ mm, $R1 = 1.5$ mm, $R2 = 1.0$ mm, and $R3 = 0.5$ mm). (a) The planar broadband resonator unit cell. (b) A single layer of planar broadband resonators.

3. SIMMULATION AND RESULT

The unit cell resonator was simulated using two Floquet ports on master and slave boundaries conditions by Ansoft HFSS (High Frequency Structural Simulator). The effective permittivity and permeability of this metamaterial was extracted from simulated S -parameters S_{11} and S_{21} . The effective permittivity and permeability of metamaterials can be determined from transmission and reflection coefficients [3]. In this case, transmission coefficient is represented by S_{11} and reflection coefficient is equal to S_{21} . The equations for the determination are listed as below.

$$n = \pm \frac{\cos^{-1} \left(\frac{1 - S_{11}^2 + S_{21}^2}{2S_{21}} \right)}{kd} \quad (1)$$

$$z = \pm \sqrt{\frac{(1 + S_{11})^2 - S_{21}^2}{(1 - S_{11})^2 - S_{21}^2}} \quad (2)$$

$$\varepsilon_{eff} = n/z \quad (3)$$

$$\mu_{eff} = n \cdot z \quad (4)$$

Here n is the refractive index, z stands for the impedance, k represents wave number ($k = \omega/c$, $\omega = 2\pi f$), and d is equal to the length of the unit cell.

The full wave simulated results show that the medium constructed of a single layer of these resonators can provide negative permittivity in rather broad bandwidth, which is from 21 GHz to 32 GHz in this case. Its fractional bandwidth is of 41.54%, which is about 2.5 times larger than ELC resonator that reported in [2]. The ELC resonator structure is able to provide negative permittivity within the frequency range of 13.8 ~ 16.3 GHz and the fractional bandwidth is of 16.67%. The results of planar broadband resonators simulation is showed in Figure 2. The left figure describes the transmission coefficient S_{11} and reflection coefficient S_{21} in dB from simulation of planar broadband resonators. The right figure represents the real part of the effective permittivity and permeability of a single layer of planar broadband resonators.

4. DISCUSSIONS

It is known that plane waves can not pass the medium with opposite signs of permittivity and permeability. Figure 2(a) shows the transmission coefficient (S_{21}) and refraction coefficient (S_{11}) of the simulation of planar broadband resonators. The transmission coefficient was down to -33 dB at 26.5 GHz when the permittivity was negative and permeability was positive at this operating frequency. From Figure 2(b) that showed the real parts of effective permittivity and permeability from 1 GHz to 50 GHz, the planar broadband resonators structure is able to provide desired positive or negative permittivity. This structure gave negative permittivity within the frequency range of 21.2 ~ 32 GHz. In addition the fractional bandwidth was about 41.54%, which is nearly 2.5 times than that produced by ELC resonators structure. There is a strong electric resonance with the real

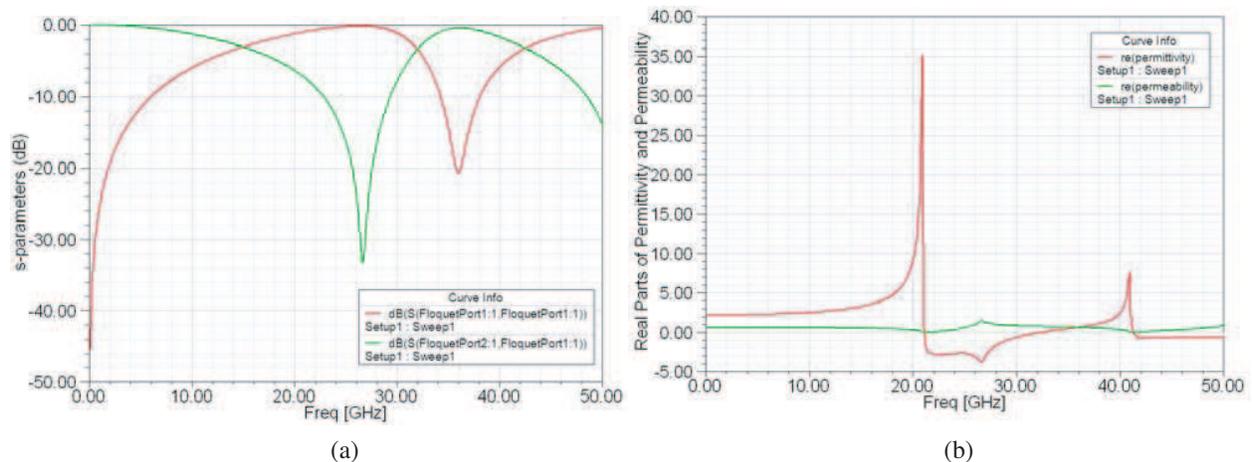


Figure 2: Results of planar broadband resonators simulation: (a) S_{11} and S_{21} in dB from simulation of planar broadband resonators. (b) The real parts of the effective permittivity and permeability of a single layer of planar broadband resonators.

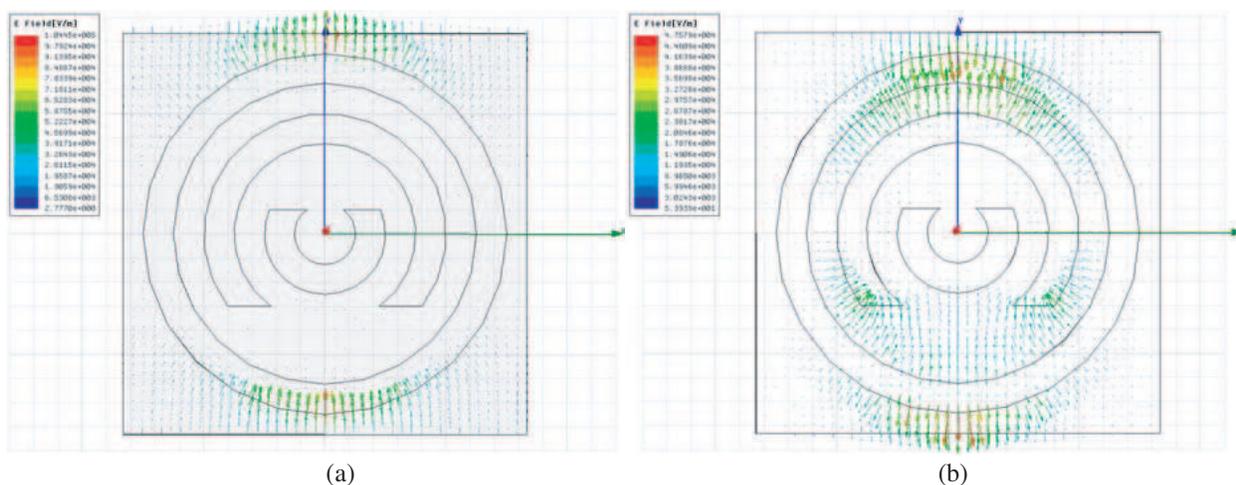


Figure 3: The distribution of the electric field: (a) distributions of the electric field on the surface of one planar broadband resonator at frequency of 5 GHz. (b) Distributions of the electric field on the surface of one planar broadband resonator at frequency 26.7 GHz.

part of the permittivity varying from nearly 35 down to -4 . Meanwhile a magnetic anti-resonance occurred and it is associated with the electric resonance. In addition there is a higher order of resonance around 41 GHz.

Figure 3(a) presents the distribution of the electric field on the surface of one planar broadband resonator at frequency of 5 GHz, where the resonance has not happened. Figure 3(b) presents the distribution of the electric field on the surface of one planar broadband resonator at frequency 26.7 GHz, where resonance was happened and the negative permittivity appeared. Here showed that the upside of new metamaterial resonator was involved in the electric resonance, at where formed as a capacitor-like structure. It is worth to mention that the symmetry of the structure also caused that the permeability did not achieve negative values.

By comparing results of simulations of the ELC resonators and planar broadband resonators, it showed that reducing coupling strength will lead to increase of the bandwidth for negative permittivity. However as the resonance is also narrowed by reducing coupling, the values of negative permittivity are becoming greater (magnitude decreasing).

5. CONCLUSIONS

One of the problems of resonant types of metamaterials is that they usually have narrow operating bandwidth. Based on the design of electric-field-coupled resonators for metamaterials, the effective permittivity of a plane layer of ELC resonators achieved negative values within the frequency range of 13.8 ~ 16.3 GHz. For the planar broadband resonators structure, it is able to provide negative permittivity within the frequency range of 21.2 ~ 32 GHz. For a plane layer of planar broadband resonators, the fractional bandwidth was about 41.54%, which is nearly 2.5 times than that produced by ELC resonators structure (16.67%). However, the magnitude of negative permittivity is smaller that produced by ELC resonators structure.

Both resonators have inductive and capacitive elements, but only the capacitive type is applied in coupling the fundamental mode to electric field. The simulations have shown that increasing the capacitance of elements leads to reducing coupling strength, and then causes narrow resonance and greater bandwidth. In addition, the symmetry of structure on inductive elements counterbalances influences produced in magnetic field.

REFERENCES

1. Engheta, N. and R. W. Ziolkowski, *Metamaterials: Physics and Engineering Explorations*, Wiley & Sons, ISBN 9780471761020, May 2006.
2. Schurig, D., J. J. Mock, and D. R. Smith, "Electric-field-coupled resonators for negative permittivity metamaterials," *Appl. Phys. Lett.*, Vol. 88, 041109, 2006.
3. Smith, D. R., S. Schulte, P. Markos, and C. M. Soukoulis, "Determination of effective permittivity and permeability of metamaterials from reflection and transmission coefficients," *Physical Review B*, Vol. 65, 195104, 2002.

Maxwell Equation in Electromagnetic and Gravitational Fields

Zi-Hua Weng

School of Physics and Mechanical & Electrical Engineering, Xiamen University
Xiamen 361005, China

Abstract— The paper studies the validity of Maxwell equation in the case for coexistence of electromagnetic field and gravitational field. In the description with the algebra of quaternions, Newton’s law of gravitation is the same as that in classical gravitational theory. While Maxwell equation is identical with that in conventional electromagnetic theory with vector terminology. And the related conclusions can be spread to the case for coexistence of electromagnetic field and gravitational field by the algebra of octonions. The study claims that Maxwell equation keeps unchanged in the case for coexistence of gravitational field and electromagnetic field, except for the direction of displacement current.

1. INTRODUCTION

In the field theory, Maxwell equation is believed to be correct in the electromagnetic field. But this validity is confined only to the range of electromagnetic field. When there exist the electromagnetic field and gravitational field simultaneously, the people doubt whether Maxwell equation is still correct or not. The validity of Maxwell equation is being disputed all the time in the gravitational field with electromagnetic field. And this validity remains as puzzling as ever. The existing theories do not explain why Maxwell equation keeps unchanged, and then do not offer compelling reason for this unique situation. The paper attempts to find out why Maxwell equation keeps the same in most cases, even in the gravitational field.

The electromagnetic field can be described with the quaternion, which was invented by W. R. Hamilton in 1843 [1]. The algebra of quaternions was first used by J. C. Maxwell to represent field equations and properties of electromagnetic field in 1861 [2]. H. R. Hertz in 1883 and O. Heaviside in 1884 recast Maxwell equation in terms of vector terminology and electromagnetic forces [3], thereby reduced the original twenty equations down to the four differential equations.

Applications of quaternion will extend through to the gravitational theory. In 1687, I. Newton [4] published the mechanical theory to describe the three laws of motion and the universal gravitation. In 1812, S.-D. Poisson reformulated Newton’s law of gravitation [5] in terms of the scalar potential. Recently, the algebra of quaternions can be used to represent the gravitational field [6].

In the paper, by means of the quaternion feature, we obtain Maxwell equation in the electromagnetic field, and Newton’s law of gravitation in the gravitational field. Rephrasing with the algebra of octonions, we can achieve Maxwell equation, wave equation, and Helmholtz equation in the case for coexistence of gravitational field and electromagnetic field.

2. GRAVITATIONAL FIELD

The feature of gravitational field can be described by the algebra of quaternions. In the quaternion space, the coordinates are $r_0, r_1, r_2,$ and $r_3,$ with the basis vector $\mathbb{E}_g = (1, \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3).$ The radius vector is $\mathbb{R}_g = \Sigma(r_i \mathbf{i}_i),$ and the velocity $\mathbb{V}_g = \Sigma(v_i \mathbf{i}_i),$ with $\mathbf{i}_0 = 1.$ Here $r_0 = v_0 t.$ v_0 is the speed of light, and t is the time. $i = 0, 1, 2, 3.$ $j = 1, 2, 3.$

The gravitational potential is

$$\mathbb{A}_g = \Sigma(a_i \mathbf{i}_i), \tag{1}$$

and the strength $\mathbb{B}_g = \Sigma(b_i \mathbf{i}_i)$ of gravitational field is

$$\mathbb{B}_g = \diamond \circ \mathbb{A}_g, \tag{2}$$

where the \circ denotes the quaternion multiplication, the operator $\diamond = \Sigma(\mathbf{i}_i \partial_i),$ $\partial_i = \partial / \partial r_i.$

The gravitational strength \mathbb{B}_g covers two components, $\mathbf{g}/v_0 = \partial_0 \mathbf{a} + \nabla a_0$ and $\mathbf{b} = \nabla \times \mathbf{a}.$

$$\mathbf{g}/v_0 = \mathbf{i}_1(\partial_0 a_1 + \partial_1 a_0) + \mathbf{i}_2(\partial_0 a_2 + \partial_2 a_0) + \mathbf{i}_3(\partial_0 a_3 + \partial_3 a_0), \tag{3}$$

$$\mathbf{b} = \mathbf{i}_1(\partial_2 a_3 - \partial_3 a_2) + \mathbf{i}_2(\partial_3 a_1 - \partial_1 a_3) + \mathbf{i}_3(\partial_1 a_2 - \partial_2 a_1), \tag{4}$$

where the gauge equation is, $b_0 = \partial_0 a_0 + \nabla \cdot \mathbf{a} = 0.$ $\mathbf{a} = \Sigma(a_j \mathbf{i}_j),$ $\nabla = \Sigma(\mathbf{i}_j \partial_j).$

The linear momentum density, $\mathbb{P} = m\nabla_g$, is the source density \mathbb{S}_g of gravitational field. The latter one is defined from the gravitational strength \mathbb{B}_g and the operator \diamond .

$$\diamond^* \circ \mathbb{B}_g = -\mu_g \mathbb{S}_g, \quad (5)$$

where m is the mass density. $*$ denotes the quaternion conjugate. $\mu_g = 4\pi G/v_0^2$ is a coefficient, and G is the gravitational constant.

The strength \mathbf{b} is too weak to detect recently. For example, the mass m_1, m_2 move with the speed v_1, v_2 respectively. They move parallel, and there will be two forces f_1, f_2 between them. Here $f_1 = m_2 g_1$, with $g_1 = Gm_1/r^2$. And $f_2 = m_2 b_1 v_2$, with $b_1 < (\mu_g/4\pi)m_1 v_1/r^2$. In most cases, $f_2/f_1 \approx v_1 v_2/c^2 \ll 1$, therefore, we will neglect f_2 generally. As an extreme case, there are $\mathbf{a} = 0$, $\mathbf{b} = 0$, and $f_2 = 0$ specially in the Newtonian gravity theory.

Table 1: The quaternion multiplication table.

	1	\mathbf{i}_1	\mathbf{i}_2	\mathbf{i}_3
1	1	\mathbf{i}_1	\mathbf{i}_2	\mathbf{i}_3
\mathbf{i}_1	\mathbf{i}_1	-1	\mathbf{i}_3	$-\mathbf{i}_2$
\mathbf{i}_2	\mathbf{i}_2	$-\mathbf{i}_3$	-1	\mathbf{i}_1
\mathbf{i}_3	\mathbf{i}_3	\mathbf{i}_2	$-\mathbf{i}_1$	-1

2.1. Newton's Law of Gravitation

In the gravitational field, the scalar part s_0 of source \mathbb{S}_g in Eq. (5) is rewritten as follows

$$\nabla^* \cdot \mathbf{h} = -\mu_g s_0, \quad (6)$$

where $\mathbf{h} = \Sigma(b_j \mathbf{i}_j)$. $s_0 = p_0 = mv_0$.

Further, the above is reduced to

$$\nabla^* \cdot (\mathbf{g}/v_0 + \mathbf{b}) = -\mu_g mv_0. \quad (7)$$

Equations (2) and (4) yield the equation

$$\nabla \cdot \mathbf{b} = 0, \quad (8)$$

and then, we have Newton's law of gravitation from Eqs. (7) and (8),

$$\nabla^* \cdot \mathbf{g} = -m/\varepsilon_g, \quad (9)$$

where the coefficient $\varepsilon_g = 1/(\mu_g v_0^2)$.

The above states that the gravitational potential \mathbf{a} has an influence on the Newton's law of gravitation, although the term $\nabla \cdot \partial_0 \mathbf{a}$ is very tiny. Meanwhile Newton's law of gravitation is a scalar invariant, and the definition of gravitational strength can be extended from the steady state to movement state.

2.2. Ampere's Law of Gravitation

In the quaternion space, the vectorial part \mathbf{s} of linear momentum density \mathbb{S}_g can be decomposed from Eq. (5).

$$\partial_0 \mathbf{h} + \nabla^* \times \mathbf{h} = -\mu_g \mathbf{s}, \quad (10)$$

where $\mathbf{s} = \Sigma(s_j \mathbf{i}_j)$. $s_j = p_j = mv_j$.

The above can be rewritten as follows

$$\partial_0 (\mathbf{g}/v_0 + \mathbf{b}) + \nabla^* \times (\mathbf{g}/v_0 + \mathbf{b}) = -\mu_g \mathbf{s}. \quad (11)$$

Equations (2)–(4) yield the equation

$$\partial_0 \mathbf{b} + \nabla^* \times \mathbf{g}/v_0 = 0, \quad (12)$$

and then, we obtain Ampere’s law of gravitation from Eqs. (11) and (12),

$$\partial_0 \mathbf{g}/v_0 + \nabla^* \times \mathbf{b} = -\mu_g \mathbf{s}. \tag{13}$$

The above means that the Newton’s law of gravitation in the quaternion space is the same as that in classical gravitational theory. In the quaternion space, the mass in either steady state or movement state can exert the gravity on other objects. The linear momentum yields the gravitational strength \mathbf{b} , which may be quite weak. And the strength \mathbf{b} and \mathbf{g} can be induced each other in the gravitational field from Eq. (12). In the paper, we append the condition $\mathbf{a} = 0$ and $\mathbf{b} = 0$ to accord with Newtonian gravitational theory.

Table 2: The operator and multiplication of the physical quantity in the quaternion space.

definition	meaning
$\nabla \cdot \mathbf{a}$	$-(\partial_1 a_1 + \partial_2 a_2 + \partial_3 a_3)$
$\nabla \times \mathbf{a}$	$\mathbf{i}_1(\partial_2 a_3 - \partial_3 a_2) + \mathbf{i}_2(\partial_3 a_1 - \partial_1 a_3) + \mathbf{i}_3(\partial_1 a_2 - \partial_2 a_1)$
∇a_0	$\mathbf{i}_1 \partial_1 a_0 + \mathbf{i}_2 \partial_2 a_0 + \mathbf{i}_3 \partial_3 a_0$
$\partial_0 \mathbf{a}$	$\mathbf{i}_1 \partial_0 a_1 + \mathbf{i}_2 \partial_0 a_2 + \mathbf{i}_3 \partial_0 a_3$

3. ELECTROMAGNETIC FIELD AND GRAVITATIONAL FIELD

The feature of gravitational field and electromagnetic field can be described simultaneously by the octonion space, which consists of two quaternion spaces.

In the quaternion space for the gravitational field, the basis vector is \mathbb{E}_g , the radius vector is \mathbb{R}_g , and the velocity is \mathbb{V}_g . In the quaternion space for the electromagnetic field, the basis vector is $\mathbb{E}_e = (\mathbf{I}_0, \mathbf{I}_1, \mathbf{I}_2, \mathbf{I}_3)$, the radius vector is $\mathbb{R}_e = \Sigma(R_i \mathbf{I}_i)$, and the velocity is $\mathbb{V}_e = \Sigma(V_i \mathbf{I}_i)$. The \mathbb{E}_e is independent of the \mathbb{E}_g , with $\mathbb{E}_e = \mathbb{E}_g \circ \mathbf{I}_0$.

These two quaternion spaces can be combined together to become an octonion space [7], with the octonion basis vector $\mathbb{E} = (1, \mathbf{i}_1, \mathbf{i}_2, \mathbf{i}_3, \mathbf{I}_0, \mathbf{I}_1, \mathbf{I}_2, \mathbf{I}_3)$. The radius vector in the octonion space is $\mathbb{R} = \Sigma(r_i \mathbf{i}_i + R_i \mathbf{I}_i)$. And that the octonion velocity is $\mathbb{V} = \Sigma(v_i \mathbf{i}_i + V_i \mathbf{I}_i)$. When the electric charge is combined with the mass to become the electron or the proton etc., we obtain the relation, $R_i \mathbf{I}_i = r_i \mathbf{i}_i \circ \mathbf{I}_0$, and $V_i \mathbf{I}_i = v_i \mathbf{i}_i \circ \mathbf{I}_0$. Here the symbol \circ denotes the octonion multiplication.

The potential of gravitational field and electromagnetic field are $\mathbb{A}_g = \Sigma(a_i \mathbf{i}_i)$ and $\mathbb{A}_e = \Sigma(A_i \mathbf{I}_i)$ respectively. They are combined together to become the field potential in the octonion space.

$$\mathbb{A} = \mathbb{A}_g + k_{eg} \mathbb{A}_e. \tag{14}$$

The field strength \mathbb{B} consists of the gravitational strength \mathbb{B}_g and electromagnetic strength \mathbb{B}_e . The selecting gauge equations are $b_0 = 0$ and $B_0 = 0$.

$$\mathbb{B} = \diamond \circ \mathbb{A} = \mathbb{B}_g + k_{eg} \mathbb{B}_e, \tag{15}$$

where k_{eg} is a coefficient. $B_0 = \partial_0 A_0 + \nabla \cdot \mathbf{A}$, with $\mathbf{A} = \Sigma(A_j \mathbf{i}_j)$.

The electromagnetic strength \mathbb{B}_e covers two parts, $\mathbf{E} = (B_{01}, B_{02}, B_{03})$ and $\mathbf{B} = (B_{23}, B_{31}, B_{12})$.

$$\mathbf{E}/V_0 = \mathbf{I}_1(\partial_0 A_1 + \partial_1 A_0) + \mathbf{I}_2(\partial_0 A_2 + \partial_2 A_0) + \mathbf{I}_3(\partial_0 A_3 + \partial_3 A_0), \tag{16}$$

$$\mathbf{B} = \mathbf{I}_1(\partial_3 A_2 - \partial_2 A_3) + \mathbf{I}_2(\partial_1 A_3 - \partial_3 A_1) + \mathbf{I}_3(\partial_2 A_1 - \partial_1 A_2). \tag{17}$$

The electric current density, $\mathbb{S}_e = q \mathbb{V}_e$, is the source of electromagnetic field in the octonion space. Meanwhile the linear momentum density is that of gravitational field. And the source \mathbb{S} was devised to describe consistently the field source of electromagnetism and gravitation.

$$\diamond^* \circ \mathbb{B} = -\mu \mathbb{S} = -(\mu_g \mathbb{S}_g + k_{eg} \mu_e \mathbb{S}_e), \tag{18}$$

where μ is a coefficient, and μ_e is the electromagnetic constant, with $k_{eg}^2 = \mu_g/\mu_e$. q is the electric charge. And $*$ denotes the conjugate of octonion.

From Eq. (18), we have

$$\diamond^* \circ \mathbb{B}_g + k_{eg} \diamond^* \circ \mathbb{B}_e = -(\mu_g \mathbb{S}_g + k_{eg} \mu_e \mathbb{S}_e), \tag{19}$$

According to the basis vectors, the above can be decomposed further as follows,

$$\diamond^* \circ \mathbb{B}_g = -\mu_g \mathbb{S}_g, \quad \diamond^* \circ \mathbb{B}_e = -\mu_e \mathbb{S}_e. \tag{20}$$

where the former equation is the same as Eq. (5) for the gravitational field, while the latter equation is for the electromagnetic field.

The above description states that the electromagnetic field is independent of the gravitational field absolutely, when the operator \diamond is only dealt with the quaternion coordinate but the octonion coordinate. In the octonion space, Maxwell equation of electromagnetic field remains the same as that in conventional theory of electromagnetic field.

Table 3: The octonion multiplication table.

	1	\mathbf{i}_1	\mathbf{i}_2	\mathbf{i}_3	\mathbf{I}_0	\mathbf{I}_1	\mathbf{I}_2	\mathbf{I}_3
1	1	\mathbf{i}_1	\mathbf{i}_2	\mathbf{i}_3	\mathbf{I}_0	\mathbf{I}_1	\mathbf{I}_2	\mathbf{I}_3
\mathbf{i}_1	\mathbf{i}_1	-1	\mathbf{i}_3	$-\mathbf{i}_2$	\mathbf{I}_1	$-\mathbf{I}_0$	$-\mathbf{I}_3$	\mathbf{I}_2
\mathbf{i}_2	\mathbf{i}_2	$-\mathbf{i}_3$	-1	\mathbf{i}_1	\mathbf{I}_2	\mathbf{I}_3	$-\mathbf{I}_0$	$-\mathbf{I}_1$
\mathbf{i}_3	\mathbf{i}_3	\mathbf{i}_2	$-\mathbf{i}_1$	-1	\mathbf{I}_3	$-\mathbf{I}_2$	\mathbf{I}_1	$-\mathbf{I}_0$
\mathbf{I}_0	\mathbf{I}_0	$-\mathbf{I}_1$	$-\mathbf{I}_2$	$-\mathbf{I}_3$	-1	\mathbf{i}_1	\mathbf{i}_2	\mathbf{i}_3
\mathbf{I}_1	\mathbf{I}_1	\mathbf{I}_0	$-\mathbf{I}_3$	\mathbf{I}_2	$-\mathbf{i}_1$	-1	$-\mathbf{i}_3$	\mathbf{i}_2
\mathbf{I}_2	\mathbf{I}_2	\mathbf{I}_3	\mathbf{I}_0	$-\mathbf{I}_1$	$-\mathbf{i}_2$	\mathbf{i}_3	-1	$-\mathbf{i}_1$
\mathbf{I}_3	\mathbf{I}_3	$-\mathbf{I}_2$	\mathbf{I}_1	\mathbf{I}_0	$-\mathbf{i}_3$	$-\mathbf{i}_2$	\mathbf{i}_1	-1

3.1. Gauss’s Law

In the electromagnetic field, the part \mathbf{S}_0 of the source \mathbb{S}_e in Eq. (20) is rewritten as,

$$\nabla^* \cdot \mathbf{H} = -\mu_e \mathbf{S}_0, \tag{21}$$

where $\mathbf{S}_0 = S_0 \mathbf{I}_0$, $S_0 = qV_0$, $\mathbf{H} = \Sigma(B_j \mathbf{I}_j)$.

Further, the above is reduced to

$$\nabla^* \cdot (\mathbf{E}/V_0 + \mathbf{B}) = -\mu_e qV_0 \mathbf{I}_0. \tag{22}$$

Equations (15) and (17) yield the Gauss’s law of magnetism

$$\nabla \cdot \mathbf{B} = 0, \tag{23}$$

and then, we have the Gauss’s law from Eqs. (22) and (23),

$$\nabla^* \cdot \mathbf{E} = -(q/\varepsilon_e) \mathbf{I}_0, \tag{24}$$

where the coefficient $\varepsilon_e = 1/(\mu_e V_0^2)$.

The above states that the electromagnetic potential has an influence on Gauss’s law of electromagnetism. By far we have two equations, Eqs. (23) and (24), for Maxwell equation. In contrast to classical Maxwell equation with vector terminology, we find that Eqs. (23) and (24) are the same as that in conventional electromagnetic theory respectively, although the definition of gauge equation $B_0 = 0$ is different to that in conventional electromagnetic theory.

3.2. Ampere-Maxwell Law

The vectorial part \mathbf{S} of electromagnetic source \mathbb{S}_e can be decomposed from Eq. (20),

$$\partial_0 \mathbf{H} + \nabla^* \times \mathbf{H} = -\mu_e \mathbf{S}, \tag{25}$$

where $\mathbf{S} = \Sigma(S_j \mathbf{I}_j)$, $S_j = qV_j$.

The above can be rewritten as follows

$$\partial_0 (\mathbf{E}/V_0 + \mathbf{B}) + \nabla^* \times (\mathbf{E}/V_0 + \mathbf{B}) = -\mu_e \mathbf{S}. \tag{26}$$

Equations (14), (16), and (17) yield the Faraday’s law

$$\partial_0 \mathbf{B} + \nabla^* \times \mathbf{E}/V_0 = 0, \tag{27}$$

Table 4: The operator and multiplication of the physical quantity in the octonion space.

definition	meaning
$\nabla \cdot \mathbf{S}$	$-(\partial_1 S_1 + \partial_2 S_2 + \partial_3 S_3)\mathbf{I}_0$
$\nabla \times \mathbf{S}$	$-\mathbf{I}_1(\partial_2 S_3 - \partial_3 S_2) - \mathbf{I}_2(\partial_3 S_1 - \partial_1 S_3) - \mathbf{I}_3(\partial_1 S_2 - \partial_2 S_1)$
$\nabla \mathbf{S}_0$	$\mathbf{I}_1 \partial_1 S_0 + \mathbf{I}_2 \partial_2 S_0 + \mathbf{I}_3 \partial_3 S_0$
$\partial_0 \mathbf{S}$	$\mathbf{I}_1 \partial_0 S_1 + \mathbf{I}_2 \partial_0 S_2 + \mathbf{I}_3 \partial_0 S_3$

and then, we obtain Ampere-Maxwell law in the electromagnetic field from Eqs. (26) and(27),

$$\partial_0 \mathbf{E}/V_0 + \nabla^* \times \mathbf{B} = -\mu_e \mathbf{S}. \tag{28}$$

The above means that the electric charge in either steady state or movement state can exert the electric force and magnetic force on other charges. The strength \mathbf{B} and \mathbf{E} can be induced each other in the electromagnetic field from Eq. (27). And Eqs. (27) and (28) are combined with Eqs. (23) and (24) to become Maxwell equation in the octonion space. These equations can be transitioned to Maxwell equation in the quaternion space, by means of the field potential $\mathbb{A}_q = -\mathbb{A}_e \circ \mathbf{I}_0$. By contrast with Maxwell equation in vector terminology, it is found that Eqs. (23), (24), (27), and (28) are the same as that in conventional electromagnetic theory respectively, except for the direction of displacement current. In most cases, we neglect this direction and related influences. Moreover, as the inferences of Maxwell equation, the wave equation and Helmholtz equation are the same as that in conventional electromagnetic theory with vector terminology too.

4. CONCLUSIONS

In the electromagnetic theory, some features can be rephrased with the algebra of quaternions, such as Maxwell equation, wave equation, and Helmholtz equation etc. In contrast to conventional electromagnetic theory with vector terminology, we find the definition of gauge equation is different, and the direction of displacement current in Ampere-Maxwell equation is opposite.

With the octonion algebra, the gravitational field and electromagnetic field can be described simultaneously. The inferences in either gravitational field or electromagnetic field can be spread to the case for coexistence of electromagnetic field and gravitational field. We achieve same conclusions as that in conventional electromagnetic theory, including Maxwell equation and so on.

It should be noted that the study for the validity of Maxwell equation in the electromagnetic and gravitational fields examined only some simple cases in quaternion spaces with neglecting the field energy. Despite its preliminary characteristics, this study can clearly indicate that Maxwell equation of electromagnetic field can be deduced with the algebra of quaternions. In the octonion space, some equations of the two fields can be drawn out simultaneously. For the future studies, the research will concentrate on only some predictions about the displacement current.

ACKNOWLEDGMENT

This project was supported partially by the National Natural Science Foundation of China under grant number 60677039.

REFERENCES

1. Hamilton, W. R., *Elements of Quaternions*, Longmans, Green & Co., London, 1866.
2. Maxwell, J. C., *A Treatise on Electricity and Magnetism*, Dover Publications Inc., New York, 1954.
3. Rothman, T. and S. Boughn, "The lorentz force and the radiation pressure of light," *American Journal of Physics*, Vol. 77, No. 2, 122–127, 2009.
4. Newton, I., *The Mathematical Principles of Natural Philosophy*, Dawsons of Pall Mall, London, 1968.
5. Wu, N., "Classical gravitational interactions and gravitational lorentz force," *Communication of Theoretical Physics*, Vol. 44, No. 11, 883–886, 2005.
6. Weng, Z.-H., "Electromagnetic forces on charged particles," *PIERS Proceedings*, 361–363, Moscow, Russia, August 18–21, 2009.
7. Gogberashvili, M., "Octonionic electrodynamics," *Journal of Physics A*, Vol. 39, No. 22, 7099–7104, 2006.

Study on Description of Electromagnetic Wave

Yelin Xu

Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China

Abstract— Classic electromagnetic wave theory, which is called ring-ring theory by the author, describes electromagnetic wave as electricity producing magnetism and magnetism producing electricity. . . . The author points out that ring-ring theory is not the best description of electromagnetic wave because it is completely opposite to experimental result and it decays too fast. The author also shows us that if the solution of Maxwell Equation is used to describe electromagnetic wave, the above deficiencies can be overcome. New description of electromagnetic wave brings us new knowledge: electromagnetic wave has the function of continuously adjusting its velocity automatically.

1. RAISING OF THE PROBLEM

Ring-ring theory is shown as Figure 1(a), experimental result [1–6] is shown as Figure 1(b), and apparently these two illustrations are contradictory. Ring-ring theory regards that both magnetic line ring and electric line ring are fixed and immovable; electromagnetic wave has longitudinal component; electric field component E and magnetic field component H appear successively; electricity and magnetism have different energy. Yet experiments reveal that electromagnetic wave is traveling wave; electromagnetic wave is transverse wave; electricity and magnetism have the same phase; electricity and magnetism have the same energy. Evidently ring-ring theory is completely opposite to experimental result, it can not describe electromagnetic wave correctly.

2. CLASSIC RING-RING THEORY DECAYS TOO FAST

Figure 2(a) is the enlarged diagram of ring-ring electromagnetic wave. Where letter A in the figure expresses the antenna; B_1 is the magnetic ring produced by A ; B_1 produces electric ring E_1 ; E_1 produces magnetic ring B_2

The following analyses explain that the above ring-ring theory is the connection chart of many transformers. We regard antenna A as the “primary coil” of the first transformer, B_1 as the “iron-core” of the first transformer, the first half of E_1 as the “secondary coil” of the first transformer, then B_1 is a complete transformer.

Different from the ordinary transformer, the lacquered wire and iron-core here consist of air.

For the same reason, B_2 , B_3 . . . are all complete transformers. Based on the above analyses, it's reasonable for us to calculate the relationship between B_1 and B_2 by using the method of calculating a transformer which will simplify the calculating method.

B_1 is known, what are the values of E_1 and B_2 ? Assume that the three rings B_1 , E_1 and B_2 are all round with the same size, the diameter of the center of the ring is D , the cross section diameter

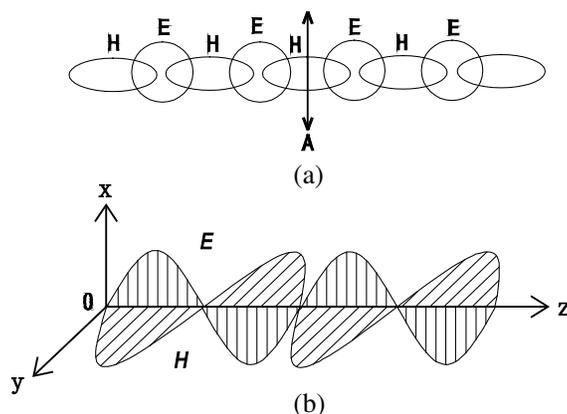


Figure 1: (a) is ring-ring theory; (b) is experimental result. Apparently, ring-ring theory is not the best description of electromagnetic wave.

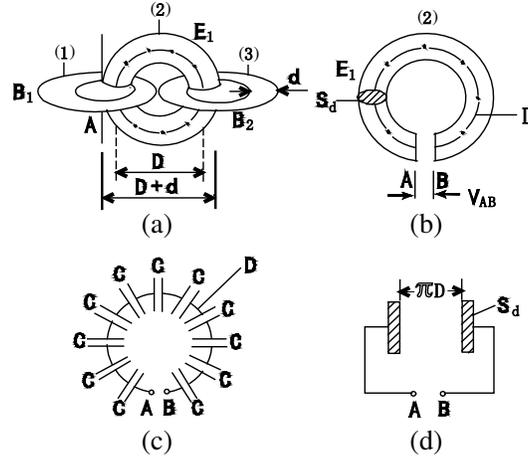


Figure 2: According to ring-ring theory, electromagnetic wave decays from 100% to 2% within one wavelength.

of the ring is d , the frequency of the current is f , wave length is λ , there is:

$$\begin{aligned} \lambda &= D + d = 4d; & D &= 3d; & D &= \frac{3}{4}\lambda, & (\pi D)^2 &= \frac{9}{16}\pi^2\lambda^2 \\ d &= \frac{1}{4}\lambda, & S_d &= \frac{\pi}{64}\lambda^2, & S_d^2 &= \frac{\pi^2}{4096}\lambda^4 \end{aligned} \tag{1}$$

A small opening is kept at the lower part of ring (2) (Figure 2(b)), we will calculate the electric potential difference V_{AB} between the two ends of the opening. We regard that the distribution of the force lines in the ring is uniform, according to Faraday Induction Law, we have:

$$V_{AB} = -\frac{d\varphi}{dt} \tag{2}$$

$$B_1 = B_{10} \cos \omega t \tag{3}$$

$$V_{AB} = S_d B_{10} \omega \sin \omega t \tag{4}$$

We will then work out the capacitance C_{AB} between point A and B along the loop D . The method is to connect countless thin capacitors in series along loop D (Figure 2(c)), the difference between the inner and outer radius along loop D is neglected. When the wire of the series connected capacitors is so short that it can be ignored, Figure 2(c) can be simplified as Figure 2(d), there is:

$$C_{AB} = \frac{\varepsilon S_d}{\pi D} \tag{5}$$

The displacement current I_d along the loop D is:

$$I_d = V_{AB} \omega C_{AB} \tag{6}$$

I_d produces magnetic field B_2 in ring (3), B_2 is calculated according to Ampere Theorem:

$$\oint_D B_2 dL = \mu I_d \tag{7}$$

$$B_2 = \frac{\mu}{\pi D} I_d = (\varepsilon\mu) \frac{S_d^2 \omega^2}{(\pi D)^2} B_{10} \sin \omega t = 0.017(\varepsilon\mu)(f^2 \lambda^2) B_{10} \sin \omega t \tag{8}$$

The transmission speed of electricity is:

$$v = f\lambda = \frac{1}{\sqrt{\varepsilon\mu}} \tag{9}$$

So, there is:

$$B_2 = 0.017 B_{10} \sin \omega t \tag{10}$$

The average value of formula (10) can be expressed as:

$$\frac{1}{\frac{1}{4}T} \int_0^{\frac{T}{4}} B_2 dt = \frac{1}{\frac{1}{4}T} \int_0^{\frac{T}{4}} 0.017B_{10} \sin \omega t dt \tag{11}$$

Let b_1 and b_2 be the average value, there is:

$$b_2 = 0.017b_1 \tag{12}$$

Calculation indicates that the above theory decays too fast. Furthermore, from the above calculation we can see that in ring-ring theory the energy of electric field is not equal to that of magnetic field, that is:

$$W_{B_1} \neq W_{E_1} \quad W_{E_1} \neq W_{B_2} \tag{13}$$

This is not in conformity with numerous research result. To sum up, it is difficult for ring-ring theory to describe electromagnetic wave.

3. DESCRIBE ELECTROMAGNETIC WAVE BY APPLYING THE SOLUTION OF MAXWELL EQUATION

Maxwell Equation, which is the summary of a large number of experiments, has been proved to be unquestionable. Investigating new describing method should depend on this equation which can be expressed as:

$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t} \tag{14}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{15}$$

$$\nabla \cdot \mathbf{D} = \rho \tag{16}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{17}$$

The solution of this equation is:

$$\mathbf{E}(z, t) = E_0 [\sin(\omega t - kz + \phi_0)] \cdot \mathbf{e}_x \tag{18}$$

$$\mathbf{H}(z, t) = H_0 [\sin(\omega t - kz + \phi_0)] \cdot \mathbf{e}_y \tag{19}$$

$$\mathbf{e}_E \times \mathbf{e}_H = \mathbf{e}_K \tag{20}$$

When we observe formula (18) to (20) carefully, we find that the relationship between electricity and magnetism here is completely identical with that in Faraday Generator Rule. Therefore, the author assumes that the electricity and magnetism in electromagnetic wave are actually those in Faraday Generator Rule.

Based on this thought, the formation of electromagnetic wave is analyzed as: the antenna is an independent wire (Figure 3(a)), in which there is alternating current I . Current I does work and radiates magnetic line ring H with light velocity as its velocity, so the magnetic line ring has kinetic mass and momentum, it has the feature of remaining its original state of motion, and it will keep going forward. Magnetic line ring H goes forward, when it is stretched, there is no need for the environment to do work. Therefore, the magnetic line ring H can keep going farther and farther. Analysis is as following:

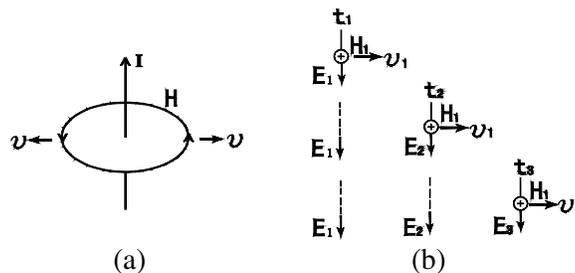


Figure 3: The emission and transmission of electromagnetic wave.

Under the situation in Figure 3(a), the Ampere's Law is valid. Here, we change its form into:

$$BL = \Sigma I \tag{21}$$

B is the magnetic induction intensity of the magnetic line ring, L is the length of the magnetic line ring, I is current. When the length of magnetic line changes, there is:

$$(B + \Delta B)(L + \Delta L) = \Sigma I \tag{22}$$

We can further get:

$$\frac{\Delta B}{B} = -\frac{\Delta L}{L} \tag{23}$$

That is: the relative change value of magnetic line ring length is equal to that of the magnetic induction intensity, but both values have opposite signs. Evidently, no work is needed for stretching the magnetic line ring H in Figure 3(a). The magnetic line ring can forever goes forward according to the Law of Conservation of Momentum.

When magnetic line H_1 moves to the right (Figure 3(b)), relatively, space moves to the left. According to Faraday Law, stretch out your right hand and let H_1 pass through your palm, if your thumb points to the left, the direction that your other four fingers point is just the direction of the electric line E_1 which is formed by induction.

The phenomenon of induction is a kind of relative motion, so E_1 is not possible to move forward with magnetic line H_1 . Magnetic line H_1 keep on moving forward, yet E_1 vanishes on the spot. At the moment of t_2 , E_2 is generated; at this time E_1 no longer exists (showed with dotted line). With the same reason, we can explain the figure at the moment of t_3 .

To sum up, the author explains electromagnetic wave as: transversely-moving magnetic line cuts the space to induct electric line, the two lines are vertical to each other and have the same phase, and this is just the real appearance of magnetic line and electric line in electromagnetic wave. Apparently, only with this explanation can electric field and magnetic field become an organic whole, and they have equal phases and energy.

Such description of electromagnetic wave is completely in accord with the current large deal of study results [7-12]. Furthermore, new description also brings us new knowledge. Analysis is as the next section.

4. ELECTROMAGNETIC WAVE HAS THE FUNCTION OF CONTINUOUSLY ADJUSTING ITS VELOCITY AUTOMATICALLY

The energy of magnetic line and electric line are W_m and W_e respectively. The motion velocity of magnetic line relative to antenna is v . According to Faraday Induction Law, we have:

$$\frac{W_e}{W_m} = \epsilon\mu v^2 \tag{24}$$

To make the new description identical with the result of numerous researches, magnetic line in the new description can be regarded as the driving force and electric line as the load. Take v as the abscissa and W_e/W_m as ordinate, we get Figure 4.

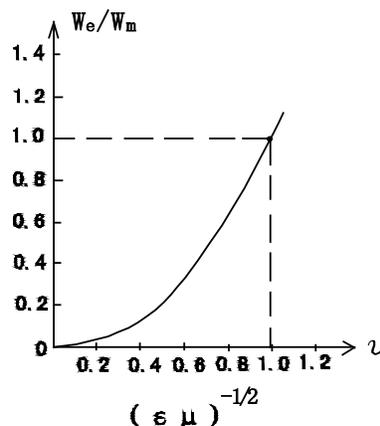


Figure 4: Electromagnetic wave can adjust its velocity automatically.

When $v < (\varepsilon\mu)^{-1/2}$, the energy of electric line is less than that of magnetic line, the load is too light, and the motion of magnetic line will gain speed, this speed-gaining process will last till $v = (\varepsilon\mu)^{-1/2}$.

When $v > (\varepsilon\mu)^{-1/2}$, magnetic line will slow down till $v = (\varepsilon\mu)^{-1/2}$.

When $v = (\varepsilon\mu)^{-1/2}$, electromagnetic wave becomes stable.

Above all, electromagnetic wave has the function of continuously adjusting its velocity automatically, so its velocity will keep a constant.

REFERENCES

1. Zuo, Z., et al., *Electric Wave and Antenna*, Hefei Industry University Press, Hefei, China, 2006.
2. Zhang, Y., *Electromagnetic Wave Propagation in Space*, Xi'an Electronic Science and Technology University Press, Xi'an, 2007.
3. Cao, X., et al., *Electromagnetic Field and Electromagnetic Wave*, Xi'an Electronic Science and Technology University Press, Xi'an, 2007.
4. Kong, J. A., *Electromagnetic Wave Theory [M]*, EMW Publishing, Cambridge, 2000.
5. Jackson, J. D., *Classical Electrodynamics*, John Wiley & Sons, New York, 1975.
6. Ulaby, F. T., *Fundamentals of Applied Electromagnetics [M]*, Prentice Hall Inc, New Jersey, 1997.
7. Griffiths, D. J., *Introduction to Electrodynamics*, Beijing World Publishing Corporation, Beijing, 2006.
8. Marshall, S. V. and G. G. Skitek, *Electromagnetic Concepts and Applications [M]*, Prentice Hall Inc, New Jersey, 1987.
9. Collin, R. E. and F. J. Zucker, *Antenna Theory [M]*, McGraw-Hill, New York, 1969.
10. Guru, B. S., et al., *Electromagnetic Field Theory Fundamentals [M]*, [S.I.]: PWS Publishing Company, 1998.
11. Wang, J., et al., *Electromagnetic Field and Electromagnetic Wave*, Xi'an Electronic Science and Technology University Press, Xi'an, 2005.
12. Xie, C., et al., *Electromagnetic Field and Electromagnetic Wave*, Higher Education Press, Beijing, 2008.

On 3D Potential Field Solutions for Atmospheric Charge Distributions

G. C. Dijkhuis

Convectron NV, Rotterdam, The Netherlands

Abstract— For atmospheric electricity including space charge we obtain static 3D solutions from quaternion space. From standard functions we calculate regular quaternion potentials validated by inductive proofs. We check plots of their iso-surfaces with parametric plots from the inverse function.

Quaternion arctangens models thundercloud electrification by equi-potential spheres sharing a dipolar ring source. Nested iso-surfaces for space charge converge on this ring as tangent tori. Intersecting iso-surfaces draw circular field lines around the same ring as Hopf links.

Quaternion inversion converges tangent iso-spheres for space charge on a central dipole source. Its alignment with a uniform field plots stable field equilibria as for dielectric or paramagnetic spheres. Opposed far and near field sources plot cusped iso-surfaces as for the Earth's magnetosphere or for Meissner states in superconductors. Meissner-type field cusps explain why ball lightning usually avoids metallic or dielectric objects.

1. INTRODUCTION

Charge on dust particles and aerosols endows atmospheric electricity with a degree of freedom not included in conventional plasma and MHD theory for ionized gases. Significant space charge is manifest in thundercloud electrification and lightning discharge, with stratospheric analogues seen as blue jets, sprites and elves forming at ionospheric altitudes. Enigmatic properties of ball lightning also involve atmospheric space charge distributions [1, 2]. For such phenomena H. Kikuchi formulates electro-hydro-dynamic theory featuring field line reconnection at critical ionization velocities in electric cusps [3]. D. Callebaut's non-linear Fourier method for solar flares and ball lightning preserves convergence at large energy storage in high-order terms [4]. For atmospheric electricity including space charge we obtained spherical solutions as charged eigenstates of the 3D Poisson equation from hyper-complex quaternion potentials [5]. A photo sequence taken during a recent thunderstorm in Germany documents tracks with stable luminosity resembling a long-lived ball lightning event [6].

High-energy emissions from atmospheric electricity are now established as common feature of rocket-triggered lightning in Florida, USA, and winter thunderclouds above the Sea of Japan [7, 8]. They associate gamma radiation at breakdown with runaway electrons. Ref. [9] derives an electric field threshold for generating runaway electrons in air. We compare cross section formulas for Thomson ionization and Rutherford scattering. The high-energy limit of their energy dependence favors impact ionization over dissipation by Coulomb forces. For air the threshold field E_{th} at sea-level such scattering cannot thermalize free electrons until impact ionization achieves equal concentrations for ions and neutral particles through:

$$\left. \begin{array}{l} \text{ion density: } n = 2.5 \times 10^{25} \text{ m}^{-3} \\ \text{nitrogen ion: } V_i = 14 \text{ eV} \end{array} \right\} \Rightarrow E_{th} = \frac{ne^2}{80\pi\epsilon_0^2 V_i} = 2.3 \frac{MV}{m} \quad (1)$$

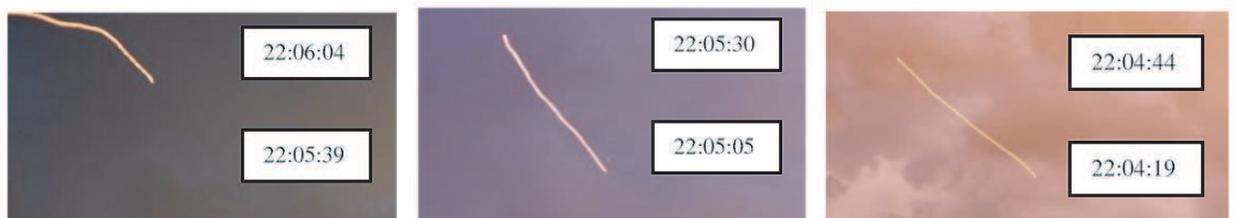


Figure 1: Tracks on photo series of thunderstorm on 13 August 2009 in Bad Hindelang, Germany, resemble a long-lived ball lightning event. Courtesy of Oliver Francks.

Here fundamental constants e and ε_o , and ionization potential V_i for nitrogen quantify the standard linear scaling with particle density. At sea level threshold (1) will form dense and cool electron plasma near the bosonic borderline between classical Maxwell-Boltzmann and quantum-mechanical Fermi-Dirac regimes. Equation (1) downscales stratospheric thresholds orders of magnitude below the sea-level reference value. This decrease enhances non-thermal effects in stratospheric discharge phenomena above dipolar thunderclouds.

2. QUATERNION POTENTIAL

From recent quaternion literature we summarize our method as follows [10,11]. Let the partial derivatives of a quaternion matrix Q define the identity matrix u and three elementary matrices i, j, k as:

$$Q \equiv \begin{pmatrix} w & -x & -y & -z \\ x & w & -z & y \\ y & z & w & -x \\ z & -y & x & w \end{pmatrix}, \quad \partial_w Q \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \equiv u, \quad \partial_x Q \equiv \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \equiv i,$$

$$\begin{cases} \partial_y Q \equiv j, \quad \partial_z Q \equiv k, \quad i^2 = j^2 = k^2 = ijk = -u, \quad Ir \equiv ix + jy + kz \\ \det(Q) = (w^2 + x^2 + y^2 + z^2)^2 \equiv (w^2 + r^2)^2, \quad I^2 = -u \end{cases} \quad (2)$$

whereby i, j, k copy the product table of Hamilton’s imaginary triplet, including their anti-commutation rules $ij = -ji$ etc. The non-zero determinant in (2) ensures inverses for all quaternion functions as needed for their parametric plots. Scalar radius r and matrix I serve as shorthand notations writing quaternion functions in complex form. Three underlying complex planes $w + ix$ etc. expand the $w + Ir$ -plane into 4D quaternion space as shown graphically in Fig. 2.

A regular complex function $W + iX$ has a derivative defined by two partial derivatives as $W_w + iX_w$, subject to two Cauchy-Riemann conditions $W_w = X_x, W_x = -X_w$ connecting potentials W with X for field lines. Equations (3a) and (3b) define analogous quaternion functions and derivatives, subject to twelve symmetry conditions on the entries of Jacobian matrix J through:

$$J \equiv \begin{pmatrix} W_w & W_x & W_y & W_z \\ X_w & X_x & X_y & X_z \\ Y_w & Y_x & Y_y & Y_z \\ Z_w & Z_x & Z_y & Z_z \end{pmatrix}, \quad \begin{cases} f(q) \equiv W + iX + jY + kZ, \quad \partial_q f \equiv W_w + iX_w + jY_w + kZ_w & (3a) \\ \vec{X}_w = -\nabla W, \quad \nabla \times \vec{X} = 0, \quad \vec{x} \times \vec{X} = 0, \quad (\vec{x} \cdot \nabla)\vec{X} = W_w \vec{x} & (3b) \\ W_w \det(J) = (W_w^2 + W_x^2 + W_y^2 + W_z^2) \det(S) & (3c) \end{cases}$$

ensuring non-zero determinants in (3c) for regular inverse functions by symmetry of sub-matrix S from the *curl*-condition in (3b). Appendix A validates the regularity rules in (3b) by inductive proofs for quaternion polynomials and power laws with real coefficients.

3. SPACE CHARGE IN THUNDERCLOUDS

As first application of quaternion potentials to atmospheric electricity we model horizontal thundercloud electrification driven by winter storms above the Sea of Japan [8]. Sea water mirrors

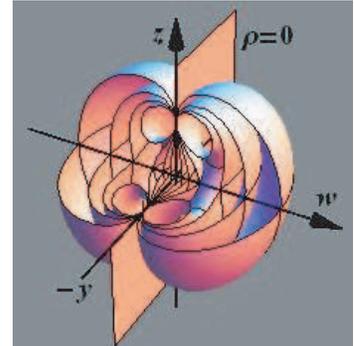
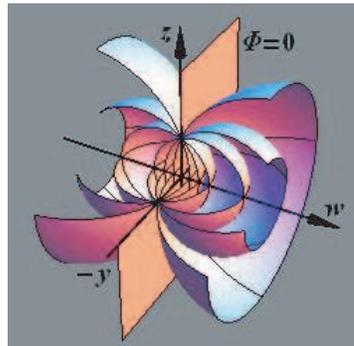
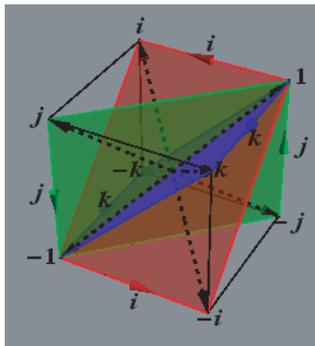


Figure 2: Three complex planes resolve 4D quaternion space.

Figure 3: (a) Potential surfaces of $\arctan(q)$ share unit circle as dipolar ring source; (b) Space charge surfaces nest as tangent tori near the ring singularity.

their charge distributions as virtual sub-surface images with opposite polarity. Taken together as a square quadrupole configuration, their central singularity lies at sea level where field lines with opposite directions come arbitrarily close [3]. In 2D geometry Ref. [12] solves the Poisson equation numerically for quadrupolar sources with charge confined to elliptic sub-regions. Successive plots localize bifurcation points for equipotential lines and electric cusps for the field strength. And likewise for field patterns preceding lightning triggered by wire-trailing rockets.

Any complex function that models a planar potential field also models a 3D potential field via quaternion space. Constant values of its real component define equipotential surfaces that prove normal to field surfaces in three complementary sets defined by constant values of each imaginary component. Thus field lines follow as intersection of any complementary pair of such surfaces. As complementary surface pairs prove non-orthogonal, mere swapping or cycling of their real and imaginary components does not provide conjugate solutions of our quaternion potentials.

The Cauchy-Riemann conditions allow regular complex functions only harmonic potentials that solve the 2D Laplace equation, keeping the complex plane source-free except at boundaries or singular points. But the 3D Laplace operator in Poisson's equation reveals space charge distributions as additional field source for our quaternion potentials. This fundamental merit marks their relevance for atmospheric electricity and other non-neutral fluids or plasma regimes.

Let us model thundercloud potentials by $\arctan(w + Ir)$ in quaternion space. Instead of $W + iX + jY + kZ$ we write $\Phi + I\Psi$ for electric potential and field components. Real functions $\Phi = \arctan(w)$ pass through the origin of a (w, Φ) -plane, where horizontal asymptotes at $\pm 1/2\pi$ confine its principal values. The complex version $\arctan(w + ix)$ locates a pair of singular points with opposite polarity at $\pm i$ in a (w, x) -plane. Extension along the vertical axis gives 3D electrostatic solutions for parallel line sources with opposite charge. Merely reading equi-potentials for field lines and vice versa presents the conjugate magnetostatic solution for opposite currents flowing through parallel conductors.

The substitutions $Ir = ix + jy + kz$, $r^2 = x^2 + y^2 + z^2$ expand complex planes $w + Ir$ into quaternion space $q = w + ix + jy + kz$, and likewise for functions written as $\Phi + I\Psi$. For contour plots of equi-potential surfaces we need the real component Φ of $\arctan(q)$ explicitly:

$$\phi + I\psi = \arctan(w + Ir), \tag{4a}$$

$$\phi = \frac{1}{2} \arctan\left(\frac{2w}{1 - w^2 - r^2}\right), \tag{4b}$$

$$\rho = -\nabla^2\phi \tag{4c}$$

where the charge density ρ follows from real potential Φ by Poisson's equation. Our 3D perspective sets $x = 0$, and computes Cartesian co-ordinates (w, y, z) respectively for fixed Φ and for fixed ρ from (4). Contour plots in Fig. 3(a) show dipolar equi-potential surfaces as coaxial spheres sharing the unit circle of their symmetry plane. This unit circle serves in Fig. 3(b) as singularity shared by apparently tangent iso-surfaces for bipolar space charge. Near the circle nested tori add multiply connected topology to these iso-surfaces. As singular field source the circle itself constitutes a dipolar ring. Evaluating the same quaternion potential in parametric form confirms the overall picture and adds further details.

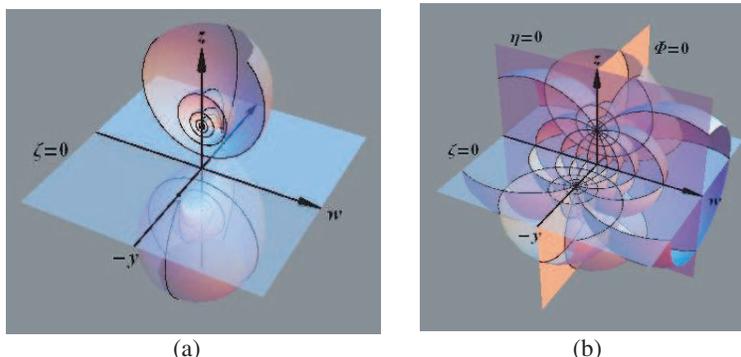


Figure 4: (a) Field surfaces of $\arctan(q)$ nest as crescent shapes near the singular ring. (b) intersections of field surfaces trace field lines orthogonal to equi-potential surfaces.

Table 1: Quaternion co-ordinates used for iso-surface plots in Figs. 3 and 4.

Equation	Plot type	Iso-surfaces for	Variables	Shown in
(4b)	Direct	Potential Φ	w, y, z	Fig. 3(a)
(4c)	...	Charge density ρ	w, y, z	Fig. 3(b)
(5)	Parametric	Potential Φ	η, ζ	Fig. 4(b)
(5)	...	Field lines 1 ζ	Φ, η	Figs. 4(a) and (b)
(5)	...	Field lines 2 η	Φ, ζ	Fig. 4(b)

In parametric plots potential and field $\Phi + I\Psi$ serve as independent variables controlling position co-ordinates $w + Ir$ as in Table 1. Inversion of arctangens in (4) readily gives the explicit co-ordinate form as:

$$w + Ir = \tan(\phi + I\psi) = \frac{\sin 2\phi + I \sinh 2\psi}{\cos 2\phi + \cosh 2\psi}, \begin{cases} I\psi \equiv i\xi + j\eta + k\zeta \\ \psi^2 \equiv \xi^2 + \eta^2 + \zeta^2 \end{cases} \quad (5)$$

with three field variables ξ, η, ζ resolving Ψ along quaternion axes (i, j, k) . Now our 3D perspective sets $\xi = 0$, and calculates co-ordinates (w, y, z) of iso-surfaces for potential Φ and field components η, ζ . Parametric plots in Fig. 4(b) draw field lines as intersection of congruent iso-surface bundles. Remarkably their nested surfaces shrink into crescent shapes around the singular circle. As before this circle fully determines the equi-potentials as a coaxial sphere bundle. Also apparent are their normal angles with planar field lines nested as Hopf links around the singular circle.

4. SPACE CHARGE IN BALL LIGHTNING

Extensive data bases connect formation of ball lightning worldwide with thunderstorm activity. Being linear our regularity rules cover of sums of regular quaternion functions. Let us model potentials for ball lightning by summing the linear potential $w + Ir$ and its inverse $(w + Ir)^{-1}$. As before we write $\Phi + I\Psi$ for its electric potential and field components. The real function $\Phi = w + w^{-1}$ draws two hyperbolas in the (w, Φ) -plane without zeroes, and confined by the vertical axis $w = 0$ and an oblique asymptote $w = \Phi$. Its complex version $w + ix + (w + ix)^{-1}$ centres a dipolar singularity inside the unit circle of the (w, x) -plane, pointing opposite to the far field. Alignment of far and near field sources results from subtraction of their linear potential and its inverse. Then vertical extension gives us 3D fields outside dielectric or para-magnetic cylinders placed normal to the far field. Merely reading equi-potentials for field lines and vice versa returns the opposed source solutions as enabled by ferromagnetism and superconductivity.

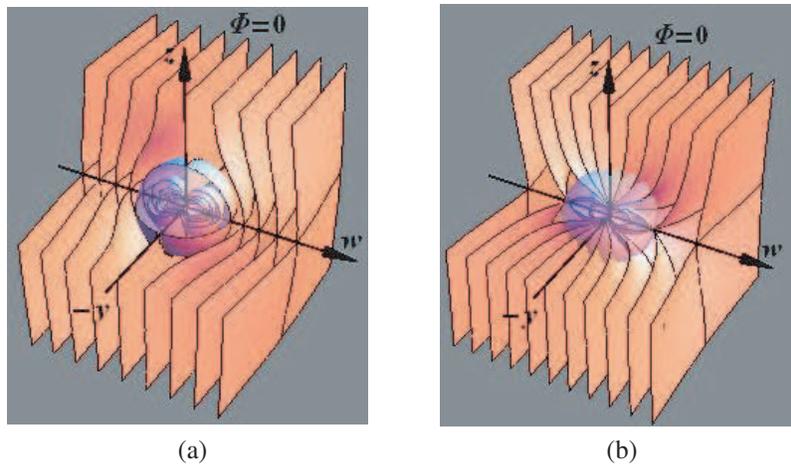


Figure 5: (a) Contour plot shows isosurfaces of quaternion potential for dipole aligned with a uniform field. The equilibrium is stable for dielectric or paramagnetic spheres. (b) Contour plot has cusped iso-surfaces of the quaternion potential with dipole axis opposite to a uniform field. This equilibrium is stable for the Meissner state of superconductors.

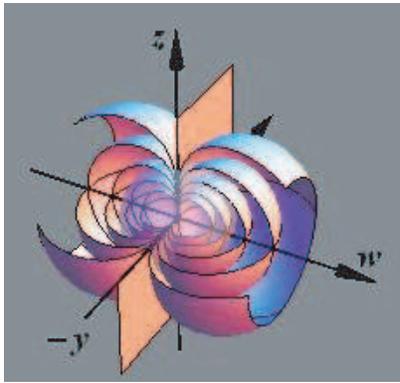


Figure 6: Contour plot shows tangent iso-surfaces for space charge of quaternion dipole.

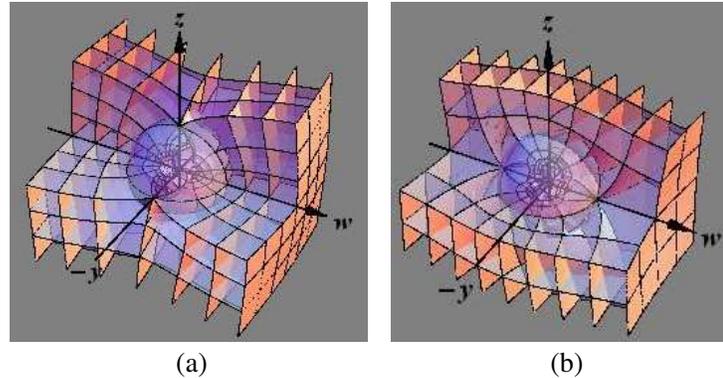


Figure 7: (a) Quaternion dipole aligned with uniform field draws field lines through the central sphere as for normal materials. (b) Field lines of quaternion dipole opposed to a uniform field curve around the central sphere as in the Meissner state of superconductors.

Expanding complex expressions into quaternion form as before we find the real quaternion potential Φ explicitly as:

$$\phi + I\psi \equiv w + Ir \pm \frac{1}{w + Ir}, \quad \phi = w \pm \frac{w}{w^2 + r^2}, \quad \begin{cases} + & \text{opposed sources} \\ - & \text{aligned sources} \end{cases} \quad (6a)$$

Our contour plots compute Cartesian co-ordinates (w, y, z) respectively for fixed Φ and for fixed charge density ρ from the Poisson Equation (4c). For aligned sources Fig. 5(a) shows the zero-potential surface as symmetry plane and unit sphere nesting the interior surfaces from the central dipole source. The opposite case in Fig. 5(b) shows unbounded surfaces separated from nests inside two envelopes that sharpen into cusps on the symmetry axis. Space charge for (6) only comes from the dipole term, so its iso-surfaces in Fig. 6 are valid for both cases, but with reversed polarity.

For corresponding parametric plots Table 1 lists potential and field components in $\Phi + I\Psi$ in (6) as independent variables controlling position co-ordinates w, x, y, z . As second-order equation in $w + Ir$, inversion of (6a) gives two solutions with opposite signs on the discriminant in explicit co-ordinate form:

$$w + Ir = \frac{1}{2} \left(\phi + I\psi \pm \sqrt{(\phi + I\psi)^2 \pm 4} \right), \quad \begin{cases} + & \text{normal state} \\ - & \text{Meissner state} \end{cases}, \quad \begin{cases} + & \text{exterior solution} \\ - & \text{interior solution} \end{cases} \quad (7)$$

Here further evaluation in terms of real functions gives cumbersome expressions not needed for our *Mathematica* plots. Parametric surfaces in Figs. 7(a) and (b) copy direct equi-potential plots in Figs. 5(a) (b). Field lines traced by surface pairs highlight different topologies for aligned vs. opposed sources in Figs. 7(a) and (b). Fig. 7(a) presents equilibrium states as for dielectric or paramagnetic spheres polarized by uniform fields. Fig. 7(b) presents equilibrium states as from perfect diamagnetism in the superconducting Meissner state. In both cases our quaternion method offers an analytic tool for modeling atmospheric space charge and plasmas including thundercloud electrification and ball lightning.

5. CONCLUSIONS

For static 3D potential calculation including space charge we conclude that our quaternion method

- regularizes differentiation by twelve conditions on derivatives, moments and eigenvalues,
- validates regularity of polynomials and power series for standard functions by inductive proofs.

Using quaternion potentials from standard functions we conclude that our solutions

- plot iso-surfaces for field lines copying parametric plots of their inverse functions,
- plot space charge for thunderclouds as nested tori converging on a dipolar ring singularity,

- plot space charge for spheres in normal fields as for dielectric or paramagnetic media, or by tangential fields seen in superconducting Meissner states and proposed for ball lightning.

Appendix A: Smooth Quaternion Closure

For field calculation in Maxwell theory quaternion functions need unique and smooth derivatives. Ref. 14 reviews definitions and presents theorems on Fueter- and Cullen-regularity for quaternion polynomials and power series with ditto coefficients. For such regularity only four conditions connect the sixteen partial derivatives of quaternion space. Also they ensure smooth functions only on concentric spheres centred at the origin [13, 14].

In our quaternion polynomials and power series all terms $a_n(w + ix + jy + kz)^n$ have real coefficients a_n . Their derivative follows as $na_n(w + ix + jy + kz)^{n-1}$ by standard rules for differentiation w.r.t. w . Expanding the brackets returns each term to explicit quaternion form $W_n + iX_n + jY_n + kZ_n$ with real functions $W_n(w, x, y, z)$ etc. of real co-ordinates w, x, y, z as components. Our quaternion derivative follows as $W_{n,w} + iX_{n,w} + jY_{n,w} + kZ_{n,w}$ with index notation for the differentiations. Additionally we need $16 - 4 = 12$ relations between partial derivatives securing completeness and closure for satisfactory potential theory [15]. We regularize quaternion functions by standard 3D vector relations and matrix calculus on their real components W, X, Y, Z through:

$$\begin{aligned} \text{Gradient: } & X_w = -W_x, Y_w = -W_y, Z_w = -W_z \\ \text{Rotation: } & Y_z = Z_y, Z_x = X_z, X_y = Y_x \\ \text{Moment: } & yZ = zY, zX = xZ, xY = yX \end{aligned} \quad \text{Eigenvalue: } \begin{pmatrix} X_x & X_y & X_z \\ Y_x & Y_y & Y_z \\ Z_x & Z_y & Z_z \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = W_w \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad (\text{A1})$$

Column-wise our gradient and rotation rules add up to three of Fueter’s regularity conditions, but our moment and eigenvalue rules only meet his fourth condition $W_w = X_x + Y_y + Z_z$ trivially at $W_w = X = Y = Z = 0$. Our equi-potentials $W = \text{const.}$ prove normal to non-orthogonal field surfaces $X = \text{const.}$ etc., as shown in Fig. A1.

Inductive Proof

First we must validate Eq. (A1) for the exponent $n = 1$, and then for $n + 1$, given their validity for exponent n .

Inspection readily confirms (A1) for first-order terms with $q = w + ix + jy + kz$. Upon multiplication by $q^n = W_n + iX_n + jY_n + kZ_n$ the quaternion product increases order parameter n one step to $n + 1$ according to:

$$q^{n+1} = q^n \cdot q \Rightarrow \begin{cases} W_{n+1} = W_n w - X_n x - Y_n y - Z_n z, & X_{n+1} = W_n x + X_n w + Y_n z - Z_n y \\ Y_{n+1} = W_n y - X_n z + Y_n w + Z_n x, & Z_{n+1} = W_n z + X_n y - Y_n x + Z_n w \end{cases} \quad (\text{A2a,b})$$

where the moment conditions in (A1) already cancel six terms of order n . Partial derivatives with respect to suitable variables validate induction for our leading *curl* and *grad* components through:

$$\text{Grad: } W_{n+1,x} + X_{n+1,w} = W_{n,x} w - X_{n,x} x - Y_{n,x} y - Z_{n,x} z + W_{n,w} x + X_{n,w} w + Y_{n,w} y + Z_{n,w} z = 0 \quad (\text{A3a})$$

$$\text{Rot: } Y_{n+1,z} - Z_{n+1,y} = W_{n,z} y + Y_{n,z} w - W_{n,y} z - Z_{n,y} w = -Z_{n,w} y + Y_{n,w} z = (-Z_{n,y} + Y_{n,z})_{,w} = 0 \quad (\text{A3b})$$

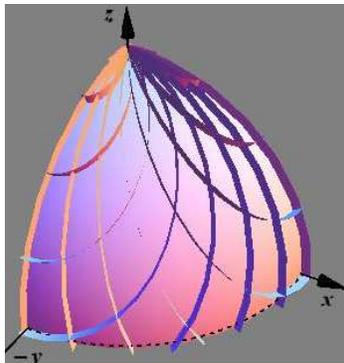


Figure A1: $\text{Log}(q)$ plots field surfaces normal to equi-potential sphere, but not to each other.

and likewise for their two other vector components. This total cancellation of order- n terms also validates induction for our moment rules by algebra, and for our eigenvalue equation by differentiation:

$$\text{Moment: } Y_{n+1}z - Z_{n+1}y = W_n \not{j}z + Y_n wz - W_n \not{z}y - Z_n wy = (Y_n \not{z} - Z_n \not{y})w = 0 \quad (\text{A4a})$$

$$\text{Eigen- value: } \begin{cases} W_{n+1,w}x - X_{n+1,x}x - Y_{n+1,y}y - Z_{n+1,z}z = W_{n,w}wx + W_n \not{x} - X_{n,w}x^2 - Y_{n,w}xy \\ -Z_{n,w}xz - W_n \not{x}x^2 - W_n \not{x} - X_{n,x}wx - W_n \not{y}xy - X_{n,y}wy - W_n \not{z}xz - X_{n,z}wz = 0 \end{cases} \quad (\text{A4b})$$

and likewise for two other vector components. Our inductive proof expands quaternion powers just as Newton's binomial formula does for real or complex numbers. For Fourier series the goniometric addition formulas enable equivalent expansions. Such addition formulas will also qualify many other standard functions for 3D potential and field calculation from quaternion space [16].

ACKNOWLEDGMENT

The author wishes to thank A. M. Cohen at Eindhoven University of Technology for inspiring comments, and Oliver Francks in Bad Hindelang, Germany, for permission to reproduce the photo's in Fig. 1.

REFERENCES

1. Singer, S., *The Nature of Ball Lightning*, Plenum Press, New York, 1978.
2. Smirnov, B. M., *Physics of Ball Lightning*, Physics Reports, Vol. 224, 151–236, 1993.
3. Kikuchi, H., *Electrohydrodynamics in Dusty and Dirty Plasmas*, ASSL, Kluwer, The Netherlands, 2001.
4. Callebaut, D. K. and G. K. Karugila, "Powerful nonlinear waves from moderate first order perturbations," *PIERS Proceedings*, 732–735, Hangzhou, China, August 22–26, 2005.
5. Dijkhuis, G. C., "Ball lightning as thermodynamic limit of the periodic system," in: *The Atmosphere and Ionosphere: Dynamics, Processes and Monitoring*, V. Bychkov, G. Golubkov, A. Nikitin, Eds., Springer Verlag, in press.
6. Francks, O., private communication, November 2009.
7. Dwyer, J. R., et al., "Energetic radiation produced during rocket-triggered lightning," *Science*, Vol. 299, 4, 2003.
8. Tsuchiya, H., et al., "Detection of high-energy gamma rays from winter thunderclouds," *Phys. Rev. Lett.*, Vol. 99, 2007.
9. Dijkhuis, G. C., "Scaling law for fusion power from boson vortex," *International Wroclaw Symposium on Electromagnetic Compatibility*, 5, 1988.
10. Kuipers, J. B., *Quaternions and Rotation Sequences*, Princeton University Press, 2002.
11. Dijkhuis, G. C., "On potential flow solutions from the division algebras," *Proceedings of the 6th European Symposium on Aerothermodynamics for Space Vehicles*, ESA SP-659, 2009.
12. Dijkhuis, G. C., "Boundary layer model and calculation for horizontal thundercloud electrification," *Environmental and Space Electromagnetics*, H. Kikuchi, Ed., Springer-Verlag, Tokyo, 1991.
13. Hamilton, W. R., "On a new species of imaginary quantities connected with the theory of quaternions," *Proceedings of the Royal Irish Academy*, Vol. 2, 424–434, 1844.
14. Gentili, G., et al., "A new approach to Cullen-regular functions of a quaternionic variable," *C.R. Acad. Sci. Paris*, Ser. I 342, 741–744, 2006.
15. Landau, L. D. and E. W. Lifshitz, *Electrodynamics of Continuous Media*, Pergamon Press, 1963.
16. Abramowitz, M. and I. A. Stegun, *Handbook of Mathematical Functions*, Dover, 1968.

Spectral Theory of Beam Scatterings for Object Imaging Using Scanning Millimeter Wave Radar Sensor

Yasumitsu Miyazaki

Department of Media Informatics, Aichi University of Technology
50-2 Manori, Nishihasama-cho, Gamagori 443-0047, Japan

Abstract— Object imaging system using electromagnetic millimeter waves may be very excellent system for object recognition in case of large optical attenuation media. Physical characteristics of millimeter waves are different scattering and transparent properties, comparing with optical waves. In this paper, beam scattering theory of object imaging using scanning millimeter wave radar is studied for object shape recognition and millimeter wave computer tomography. Object imaging system of electromagnetic millimeter waves may contribute to important medical applications and industrial fields. Millimeter wave radar of short wavelengths and high frequency using short pulses of millimeter wave carriers yield precise distance measurement and image processing. Distance measurement using short pulse millimeter carriers are derived by time differences between transmitted pulse and received pulse, and image information are given by reflected waves by scanning millimeter beam waves, like optical vision information. Temporal and spatial characteristics of electromagnetic scattering and reflection by objects are studied. Fundamental characteristics and application of scanning millimeter wave radars of synthetic aperture type and multi-wavelengths for image recognition of objects are shown. Using spatial spectral functions, asymptotic expressions concerned with Hermite-Gaussian eigenfunctions are discussed for the incident Gaussian beam, and reflected and scattered waves, using boundary relations for each spectral function.

1. INTRODUCTION

Shape image recognition system of objects using millimeter electromagnetic waves has excellent characteristics for object sensing and detection, comparing with optical shape recognition system, because physical properties of transparent and reflecting material consisting objects for millimeter electromagnetic waves have very different properties compared with optical characteristics [1–4]. In case that target objects are buried in optical lossy media with large attenuation of optical waves and, invisible and optical large attenuation structures are located between optical incident generator and sensing target, optical sensing and detection system for object shape recognition can not be applied [5, 6]. Object imaging system using millimeter waves may be very excellent system for object recognition in case of large optical attenuation media. Optical and X-ray computer tomographies have been rapidly developed and recently, ultra-sonic CT has been also well accomplished [7]. Millimeter wave CT has not been so much discussed. In this paper, beam scattering theory of object imaging using scanning millimeter wave radar is studied for object shape recognition and millimeter wave computer tomography.

Gaussian beam waves with temporal Gaussian impulse radiated from millimeter wave antenna incident to complicated targets are considered. Spatial and temporal spectral functions are introduced using transverse wave number spectrum for the spatial space and angular frequencies for time domain. Spatial spectral functions are expressed for incident and scattering electromagnetic fields. Asymptotic expressions concerned with Hermite-Gaussian eigenfunctions are derived for the incident Gaussian beam, and reflected and scattered waves, using beam waist and beam spot size. Radiating millimeter beam packets from scanning millimeter radars and reflected beam packets of temporal Gaussian pulse form and spatial Gaussian beam form are expressed as Fourier components for time coordinates and as spectral functions expanded by Hermite-Gaussian functions for space coordinates. Incident, reflected and scattered fields are studied using beam mode expansions derived by Hermite-Gaussian spectral functions.

Scattered and transmitted waves are discussed using eigenfunction orthogonalities, satisfying boundary conditions on complicated target shape for spectral functions of incident, reflected and transmitted waves. Boundary conditions of objects are applied to derive reflected and scattered field for spectral components. Based on temporal and spatial characteristics of millimeter wave scattering, image recognition of object targets using scanning millimeter wave radars of scanning millimeter wave radars of synthetic aperture type are shown with computer image processing. Object imaging is analyzed by the Synthetic Aperture Radar Method. Correlation functions by

transmitted and received fields yield image processing information for transverse image patterns at each time. Velocities of objects are given by frequency variations of received fields. Scanning millimeter wave radar in the front of measurement equipment radiating short Gaussian pulse and forward Gaussian beam can construct spatial image at each time analyzing reflected and scattered waves by objects. Based on this fundamental theory, image processing radar sensor systems may be constructed.

2. OBJECT SHAPE RECOGNITION BY SCANNING MILLIMETER WAVE RADAR

Fundamental architectures of scanning millimeter wave radars for object image processing are shown in Figs. 1(a) and (b). Distances between the antenna and object are given by pulse time difference measurement of transmitting millimeter wave pulse and receiving pulse. Measurement situation of antenna and object, and the millimeter wave radar system for image processing are shown in Fig. 2.

Figure 1(b) shows object shape recognition systems using scanning millimeter wave radar. S_{ij} are radiation apertures of scanning millimeter pulse radar and f_{ij} are multiple millimeter wave frequencies.

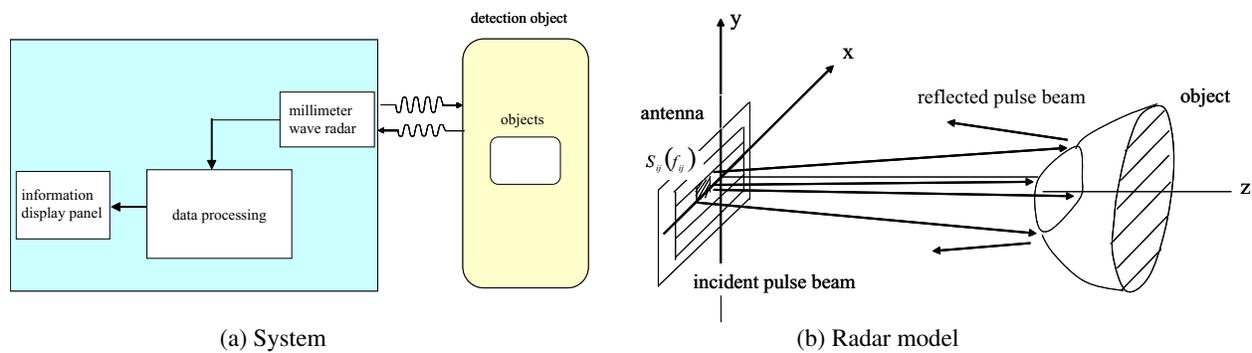


Figure 1: Fundamental architecture of scanning millimeter wave radar.

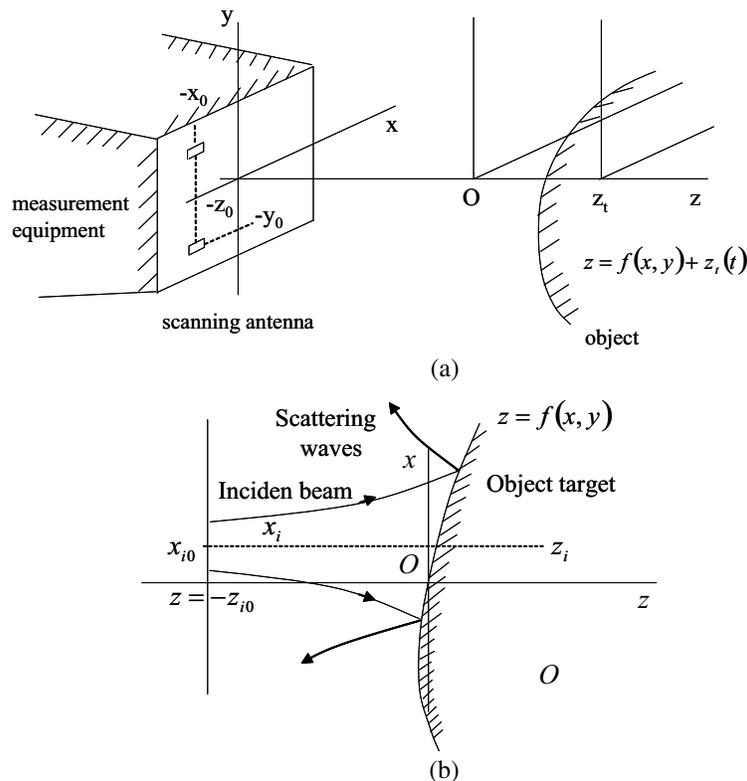


Figure 2: Image processing antenna and measured object. (a) Reflection and scattering of millimeter wave pulse beam. (b) Object imaging and beam scattering.

3. REFLECTION AND SCATTERING BY MILLIMETER WAVE PULSES

Beam waves with beam center $(-x_0, -y_0)$ of millimeter wave pulses are radiated from scanning millimeter wave antenna at position $z = -z_0$ as shown in Figs. 2 and 3. Body surfaces of object at time t are given by $z = f(x, y) + z_t(t)$ of position center z_t .

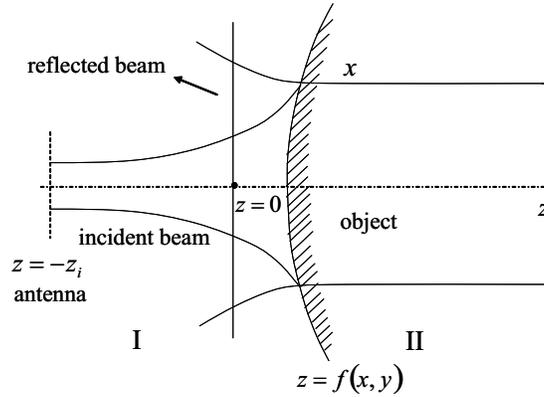


Figure 3: Incident beam and object.

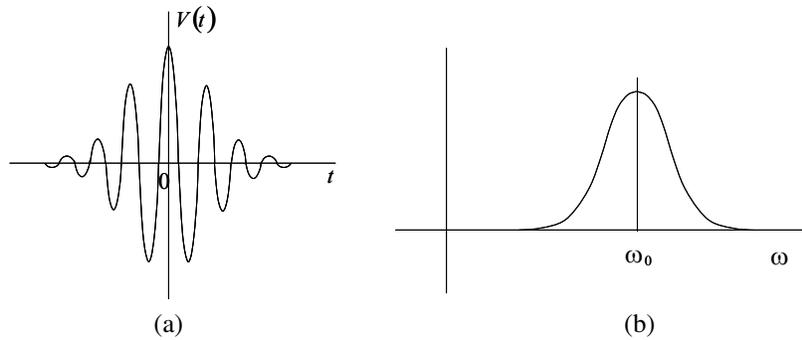


Figure 4: Pulse characteristics of incident beam. (a) Temporal characteristics of millimeter wave radar. (b) Frequency characteristics of millimeter wave pulse.

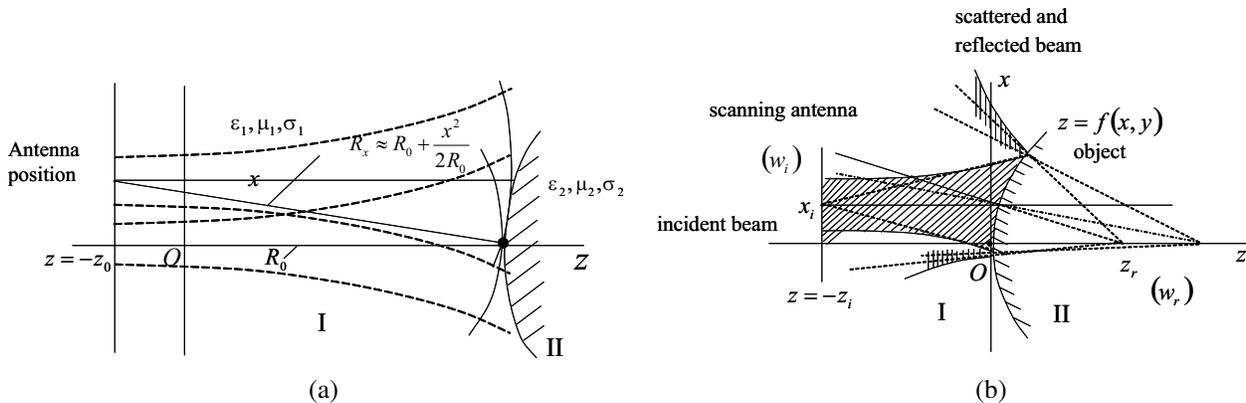


Figure 5: Scanning antenna and scattering object. (a) Incident beam and scattering object. (b) Incident and scattered beams.

Incident pulses as shown in Fig. 4 are

$$\mathbf{E}^{(i)}(\mathbf{r}, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{V}(\omega) \tilde{\mathbf{E}}^{(i)}(\mathbf{r}, \omega) e^{j\omega t} d\omega \quad (1)$$

where temporal function $V(t)$ and spectral function $\tilde{V}(\omega)$ are

$$V(t) = e^{-\left(\frac{2t}{T}\right)^2} e^{j\omega_0 t} = \frac{1}{2\pi} \int \tilde{V}(\omega) e^{j\omega t} d\omega \tag{2}$$

$$\tilde{V}(\omega) = \frac{\sqrt{\pi}T}{2} e^{-\left(\frac{\omega-\omega_0}{4}\right)^2}$$

When plane waves with polarization of direction $\boldsymbol{\tau}$ are incident, $\tilde{\mathbf{E}}^{(i)}(\mathbf{r}, \omega) = E_0 e^{-j\beta z} \mathbf{i}_\tau$ and using $\beta_0 = \omega_0/c$

$$\mathbf{E}^{(i)}(\mathbf{r}, t) = E_0 e^{j\omega_0 t - j\beta_0 z} e^{-\left(\frac{2}{T}\right)^2 \left(t - \frac{z}{c}\right)^2} \mathbf{i}_\tau \tag{3}$$

and when the fundamental beam mode with y position is incident, in two dimensional case,

$$\tilde{\mathbf{E}}^{(i)} = E_0 \mathbf{i}_y e^{-j\beta(z+z_0)} \frac{1}{\sqrt{1-j\zeta}} e^{-\frac{a^2 \omega^2}{2(1-j\zeta)}} \tag{4}$$

where $\zeta = \frac{(z+z_0)}{\beta} a^2$.

Object surface $z = f(x)$ with object center $x = 0$ separates the region of free space I and region of object II as shown in Fig. 5(a). Material constants of region I and II are $\varepsilon_1, \mu_1, \sigma_1$ and $\varepsilon_2, \mu_2, \sigma_2$. Incident wave, reflected wave and transmitted wave are $\tilde{\mathbf{E}}_i, \tilde{\mathbf{E}}_r, \tilde{\mathbf{E}}_t$. In the region I, $\tilde{\mathbf{E}}^{(I)} = \tilde{\mathbf{E}}_i + \tilde{\mathbf{E}}_r$, in the region II, $\tilde{\mathbf{E}}^{(II)} = \tilde{\mathbf{E}}_t$.

$$\mathbf{E}_n(\mathbf{r}_j, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \tilde{\mathbf{E}}_n(\mathbf{r}_j, \omega) e^{j\omega t} d\omega \tag{5}$$

$$\tilde{\mathbf{E}}_n(\mathbf{r}_n, \omega) = \int_{-\infty}^{\infty} \hat{\mathbf{E}}_n(\beta_{tn}) e^{-j\beta_{tn} x_n - j\sqrt{\beta_n^2 - \beta_{tn}^2} z} d\beta_{tn}$$

Wave numbers β_n are $\beta_i = \beta_r = \beta_1 = \omega \sqrt{\varepsilon_1^* \mu_1}$, $\beta_t = \beta_2 = \omega \sqrt{\varepsilon_2^* \mu_2}$, $\varepsilon_i^* = \varepsilon_i - j\sigma_i/\omega$. β_{tn} are β_{ti} , β_{tr} and β_{tt} .

If β_{tn} are small, we can use asymptotic approximation of $\sqrt{\beta_n^2 - \beta_{tn}^2} \cong \beta_{tn} - \frac{1}{2} \frac{\beta_{tn}^2}{\beta_n} - \frac{1}{8} \frac{\beta_{tn}^4}{\beta_n^3}$, and $\hat{\mathbf{E}}_n(\beta_{tn})$ can be expanded as series of Hermite-Gaussian functions

$$\hat{\mathbf{E}}_n(\beta_{tn}) = \sum_{m=0}^{\infty} a_{mn} \phi_m(\beta_{tn}) \boldsymbol{\tau}_n \tag{6}$$

where

$$\phi_m(\beta_{tn}) = \frac{1}{\sqrt{\alpha_n (2^m m! \sqrt{\pi})^{\frac{1}{2}}}} e^{-\frac{1}{2} \beta_n'^2} H_m(\beta_n') \tag{7}$$

$$\beta_n' = \beta_{tn} / \alpha_n$$

and coefficients a_{nm} are

$$a_{nm} \cong \int_{-\infty}^{\infty} \hat{\mathbf{E}}_n(\beta_{tn}) \phi_m(\beta_{tn}) \boldsymbol{\tau}_n d\beta_{tn} \tag{8}$$

When incident beam wave is beam wave with the polarization of y -direction, using incident coordinates (x_i, z_i) , incident beam waves are

$$\tilde{E}_{iy} = \int_{-\infty}^{\infty} \hat{E}_i(\beta_{ti}) e^{-j\beta_{ti} x_i - j\sqrt{\beta_1^2 - \beta_{ti}^2} (z_i + z_{oi})} d\beta_{ti},$$

$$\tilde{H}_{ix} = \int_{-\infty}^{\infty} \hat{E}_i(\beta_{ti}) \frac{(j\omega\varepsilon_1 + \sigma_1) \left(-j\sqrt{\beta_1^2 - \beta_{ti}^2}\right)}{-\beta_1^2} e^{-j\beta_{ti} x_i - j\sqrt{\beta_1^2 - \beta_{ti}^2} (z_i + z_{oi})} d\beta_{ti} \tag{9}$$

$$\tilde{H}_{iz} = \int_{-\infty}^{\infty} \hat{E}_i(\beta_{ti}) \frac{(j\omega\varepsilon_1 + \sigma_1) (-j\beta_{ti})}{-\beta_1^2} e^{-j\beta_{ti} x_i - j\sqrt{\beta_1^2 - \beta_{ti}^2} (z_i + z_{oi})} d\beta_{ti}$$

Transmitted waves are obtained as like Eq. (9), using transmission coordinates (x_t, z_t) . Reflected waves are using

$$\tilde{E}_{ry} = \int_{-\infty}^{\infty} \hat{E}_r(\beta_{tr}) e^{-j\beta_{tr}x_r + j\sqrt{\beta_1^2 - \beta_{tr}^2}(z - z_{or})} d\beta_{tr}, \quad \tilde{E}_{ty} = \int_{-\infty}^{\infty} \hat{E}_t(\beta_{tt}) e^{-j\beta_{tt}x_t + j\sqrt{\beta_2^2 - \beta_{tt}^2}(z - z_{ot})} d\beta_{tt} \quad (10)$$

Boundary conditions on object surface $z = f(x)$ are

$$\mathbf{n} \times (\mathbf{E}^{(1)} - \mathbf{E}^{(2)}) = \mathbf{0}, \quad \mathbf{n} \times (\mathbf{H}^{(1)} - \mathbf{H}^{(2)}) = \mathbf{0} \quad (11)$$

Here, $\mathbf{E}^{(1)} = \mathbf{E}_i + \mathbf{E}_r$ and $\mathbf{E}^{(2)} = \mathbf{E}_t$, $\eta_i = \frac{\omega\varepsilon_i + \sigma/j}{k_i^2}$, the following reflected and transmitted waves are obtained

$$\begin{aligned} \hat{E}_r(\beta_t) &= \frac{\eta_1 \sqrt{k_1^2 - \beta_t^2} - \eta_2 \sqrt{k_2^2 - \beta_t^2}}{\eta_1 \sqrt{k_1^2 - \beta_t^2} + \eta_2 \sqrt{k_2^2 - \beta_t^2}} \frac{B_i(\beta_t)}{B_t(\beta_t)} e^{-j\sqrt{k_1^2 - \beta_t^2}z_o} \hat{E}_i(\beta_t), \\ \hat{E}_t(\beta_t) &= \frac{2\eta_1 \sqrt{k_1^2 - \beta_t^2}}{\eta_1 \sqrt{k_1^2 - \beta_t^2} + \eta_2 \sqrt{k_2^2 - \beta_t^2}} \frac{B_i(\beta_t)}{B_t(\beta_t)} e^{-j\sqrt{k_1^2 - \beta_t^2}z_o} \hat{E}_z(\beta_t) \end{aligned} \quad (12)$$

In case of rectangular target of object as

$$f(x) = \begin{cases} 0 & (|x| > W/2) \\ h & (|x| < W/2) \end{cases} \quad (13)$$

and defining $\varphi^{(0)} = |\tilde{E}_i(x, 0)|$, we obtained asymptotically

$$B_r(\beta_t) = \int_{-\infty}^{\infty} e^{\mp j\sqrt{k_1^2 - \beta_t^2}f(x)} e^{-j\beta_t x} \varphi^{(0)}(x) dx, \quad B_t(\beta_t) = \int_{-\infty}^{\infty} e^{-j\sqrt{k_2^2 - \beta_t^2}f(x)} e^{-j\beta_t x} \varphi^{(0)}(x) dx \quad (14)$$

From Eq. (12), back scattered and reflected waves are obtained. Pulse responses due to back scattering by object are derived by Eq. (12).

When incident fundamental beam expressed in Eq. (4) is reflected by object surface of convex curvature with the radius R at $z = 0$, reflected beam is

$$\tilde{\mathbf{E}}^{(r)} = E_0 \mathbf{i}_y e^{+j\beta(z+z_0-z_r)} \frac{1}{\sqrt{1-j\zeta_r}} e^{-\frac{a_r^2 x^2}{2(1-j\zeta_r)}} \quad (15)$$

where $\zeta_r = \frac{2(z-z_r)}{\beta} a_r^2$, $z_r = \frac{R}{2+R/z_0}$, $a_r = \frac{R}{2z_0+R} a_i$.

Received pulse responses $S_r(x, t)$ are, when reflection coefficient of object is R_s

$$\mathbf{E}^{(r)}(\mathbf{r}, t) = E_0 R_s e^{j\omega_0 t - j\beta_0(z+z_0-z_r)} e^{-\frac{a_r^2}{2} x^2} e^{-\left(\frac{2}{T}\right)^2 \left(t - \frac{\ell}{c}\right)^2} \mathbf{i}_r \quad (16)$$

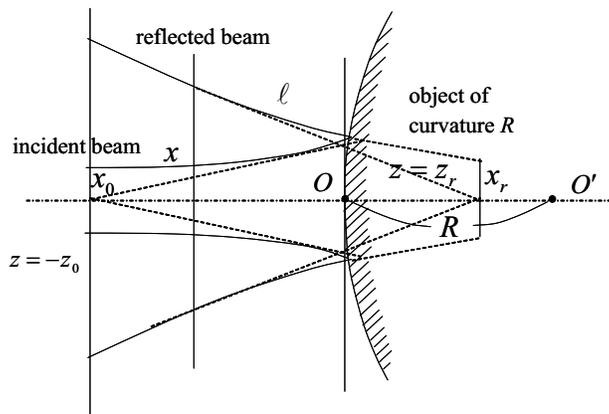


Figure 6: Scanning millimeter radar and reflected beam.

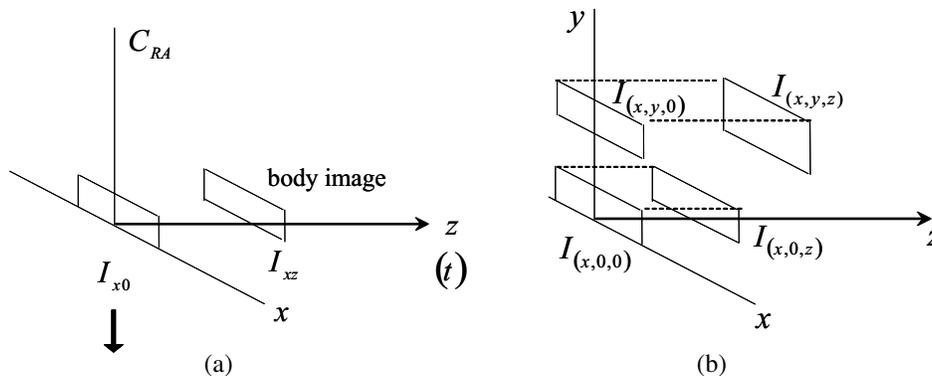


Figure 7: Scanning millimeter wave radar and synthetic aperture processing. (a) x coordinate one dimensional Data on each z plane and C_{RA} . (b) x, y coordinates two dimensional images on each z plane.

where $\ell(x, z_0, z_r, R)$ is given by ray path of incident beam and reflected beam.

Body shape recognition is performed by image processing of received signals $S_r(x, t)$ given by multiple scanning radar beams, as shown in Figs. 7(a) and (b) where I_{xz} and $I(x, y, z)$ are image characteristics given by azimuth compression $C_{RA}(x, t)$.

4. SUMMARIES

In the automatic image systems, image processing systems using millimeter wave scanning radar are very useful. Particularly, object shape detection and recognition systems are extremely important. Object shape detection systems by using scanning millimeter wave radar are discussed by electromagnetic field theory with spectral functions of Gaussian beams. Pulse responses of very short millimeter waves are shown for reflected and scattered beam waves by conducting and dielectric object. Synthetic aperture radar system can be accomplished, based on this fundamental theory for object shape detection in automatic image systems

REFERENCES

1. Miyazaki, Y. and Y. Akao, "Reflection and mode conversion of beam wave through a curved interface between dielectrics and thick dielectric lens," *Trans. of IECE, Japan*, Vol. 51-B, No. 1, 25–31, 1968.
2. Miyazaki, Y., "Space-time impulse responses of electromagnetic backscatterings of beam waves for sensing," *Proc. of OFSET*, 361–364, 2000.
3. Miyazaki, Y., "Electromagnetic scattering and filtering in random media for image sensing," *PIERS Proceedings*, 689–697, Cambridge, USA, July 2–6, 2008.
4. Mizuno, K., "Imaging arrays in the millimeter wave region: Millimeter wave camera," *J. of IEICE Japan*, Vol. 79, No. 2, 177–181, 1996.
5. Miyazaki, Y., "Electromagnetic scattering theory of car body imaging using scanning millimeter wave radar," *PIERS Proceedings*, 1423–1428, Beijing, China, March 23–27, 2009.
6. Takahashi, K. and Y. Miyazaki, "FDTD computer simulation of millimeter wave driving radar for ITS," *11th World Congress on OTS Nagoya*, Vol. 4088, 6, 2004.
7. Miyazaki, Y. and K. Kouno, "FDTD analysis of spatial filtering of scattered waves for optical CT of medical diagnosis," *IEEJ Trans. FM*, Vol. 129, No. 1, 693698, 2009.

FDTD Parallel Computing of Electromagnetic Wave Scattering by Clouds for Microwave Remote Sensing of Weather Satellite

Yasumitsu Miyazaki¹, Nobuo Goto², and Koichi Takahashi¹

¹Department of Media Informatics, Aichi University of Technology
50-2 Manori, Nishihassama-cho, Gamagori 443-0047, Japan

²Institute of Technology and Science, The University of Tokushima
2-1 Minamijosanjima-cho, Tokushima 770-8506, Japan

Abstract— The spaceborne remote sensing is very useful for direct measurement of weather environment in wide area and disaster prevention system in case of strong rainfalls. In this paper, we analyze the electromagnetic scattering by statistically distributed clouds in rain region in satellite remote sensing. The satellite antenna is very high over earth surface and moves on two dimensional plane. Clouds of rain region can be evaluated by SAR system. Characteristics of electromagnetic wave scattering from clouds are analyzed for improvement of the space resolution in the microwave remote sensing by weather satellite. The characteristics of electromagnetic backscattering for microwave remote sensing are analyzed by FDTD and analytical method.

1. INTRODUCTION

The spaceborne remote sensing is very important for direct measurement of weather environment in wide area and is effective for safety system of ITS and disaster prevention system to disasters caused by strong rainfalls. In this paper, we consider cloud observation system using satellite remote sensing and analyze the electromagnetic scattering by statistically distributed clouds in rain region [1]. The satellite antenna is very high over earth surface and moves on two dimensional plane. Clouds of rain region can be evaluated by SAR system [2].

In this study, characteristics of electromagnetic wave scattering from clouds are analyzed, as a purpose to improve the space resolution of the microwave remote sensing using weather satellite. We have analyzed the electromagnetic scattering from the object to evaluate the resolution of microwave remote sensing using satellite. Radiation field from satellite antenna is calculated for incident field to FDTD numerical method using Green's function [3–6]. Electromagnetic scattering by object is calculated by FDTD method. Received field can be derived by analytic expression using equivalent electric and magnetic currents on the boundary given by FDTD method.

We analyzed electromagnetic backscattering from clouds at weather satellite by analytic expression with FDTD. Received field of signals at satellite position is calculated using the surface equivalent electric and magnetic currents derived by FDTD method. Gaussian pulse is transmitted from the satellite antenna. SAR images of clouds have been reconstructed from the received electric field.

For the analysis of microwave scattering by clouds and rains, the analysis region including clouds and raindrops, parallel computation of FDTD using grid computer is applied to evaluate scattering field, where the total analysis space is divided into subdomains [7].

2. INCIDENT FIELD RADIATED BY SATELLITE ANTENNA

Fields in large region I between antenna and area of clouds and rain in Figs. 1–2 are analytically studied and fields in small region II around raindrops is numerically studied by FDTD. The electromagnetic fields radiated from current density $\mathbf{J}(\mathbf{r}, t)$ are

$$\begin{aligned}\mathbf{E}(\mathbf{r}, t) &= \int_{S_a} d\mathbf{r}' \int_{-\infty}^{\infty} dt' G(\mathbf{r}, \mathbf{r}'; t, t') \left\{ -\mu_0 \frac{\partial \mathbf{J}(\mathbf{r}', t')}{\partial t'} \right\} \\ \mathbf{H}(\mathbf{r}, t) &= \int_{S_a} d\mathbf{r}' \int_{-\infty}^{\infty} dt' G(\mathbf{r}, \mathbf{r}'; t, t') \left\{ -\nabla' \times \mathbf{J}(\mathbf{r}', t') \right\}\end{aligned}\quad (1)$$

where, S_a is the antenna surface and Green's function $G(\mathbf{r}, \mathbf{r}'; t, t')$ is

$$G(\mathbf{r}, \mathbf{r}'; t, t') = \frac{1}{4\pi |\mathbf{r} - \mathbf{r}'|} \delta\left(t - t' - \frac{|\mathbf{r} - \mathbf{r}'|}{c}\right)$$

When antenna with $2a \times 2b$ at $z = h_1$ in Fig. 3 radiates Gaussian beam with beam waist beam spot size $r_{0x} = r_{0y}$ and wave number k , radiated field $E_y(x, y, z)$ in free space region I is given by Eq. (2).

If, $\zeta_x = \frac{2z}{kr_{0x}^2}$, $\zeta_y = \frac{2z}{kr_{0y}^2}$

$$E_y(x, y, z) = E_0 e^{-jkz} \frac{1}{\sqrt{1-j\zeta_x}} \frac{1}{\sqrt{1-j\zeta_y}} e^{-\frac{x^2}{r_{0x}^2} \left(\frac{1}{1-j\zeta_x} \right) - \frac{y^2}{r_{0y}^2} \left(\frac{1}{1-j\zeta_y} \right)} \quad (2)$$

3. SCATTERED FIELD ANALYSIS BY FDTD METHOD FOR RAIN CLOUD REGION II

Raindrops are distributed randomly in rain region. The incident wave for FDTD calculation is assumed the Gaussian beam of frequency 20 GHz radiated from the surface S_0 , $z = h_4$ ($z' = h_4 - h_2$) in region II of Fig. 4. In FDTD method, the analysis region is discretized as $x' = i\Delta s$ ($0 \leq i \leq Nx$), $y' = j\Delta s$ ($0 \leq j \leq Ny$), $z' = k\Delta s$ ($-N_0 \leq k \leq Nz$) and $t = n\Delta t$ ($0 \leq n \leq T$). When radiation field in the antenna is expressed by equivalent current $J_y(x, y, z, t)$, the electromagnetic fields at point (i, j, k) at time $n\Delta t$ are calculated by difference equation from Maxwell's equation,

$$E_y^n(i, j, k) = c_1 E_y^{n-1}(i, j, k) - c_2 J_y^{n-1}(i, j, k) + c_3 \{ H_x^{n-1}(i, j, k) - H_x^{n-1}(i, j, k-1) - H_z^{n-1}(i, j, k) + H_z^{n-1}(i, j, k) \} \quad (3)$$

$$H_y^n(i, j, k) = H_y^{n-1}(i, j, k) - c_4 \{ E_x^n(i, j, k+1) - E_x^n(i, j, k) - E_z^n(i+1, j, k) + E_z^n(i, j, k) \}$$

where,

$$c_1 = \frac{1 - \sigma\Delta t / (2\varepsilon)}{1 + \sigma\Delta t / (2\varepsilon)}, \quad c_2 = \frac{\Delta t / \varepsilon}{1 + \sigma\Delta t / (2\varepsilon)}, \quad c_3 = \frac{\Delta t / \varepsilon}{1 + \sigma\Delta t / (2\varepsilon)} \frac{1}{\Delta s}, \quad c_4 = \frac{\Delta t}{\mu\Delta s}$$

Here, Δs is cell size and Δt is time step. ε , μ and σ are the dielectric constant, conductivity, and permeability of material. Here, equivalent current corresponding to incident field radiated from

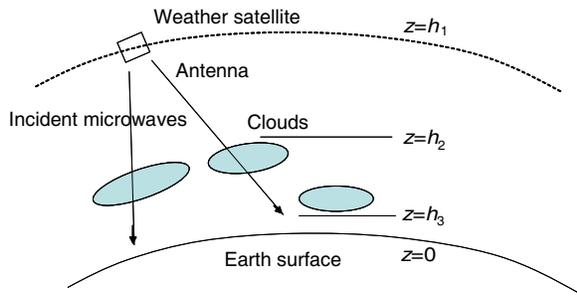


Figure 1: Weather satellite and microwave scattering.

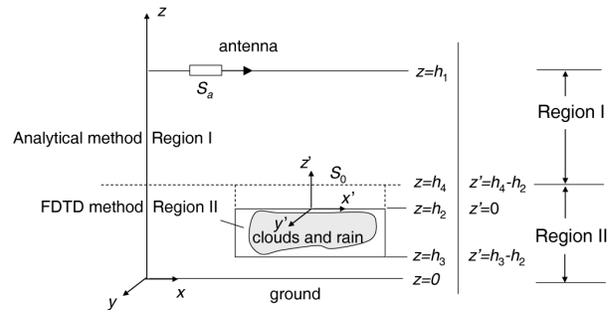


Figure 2: Calculation model for microwave remote sensing of clouds and rain.

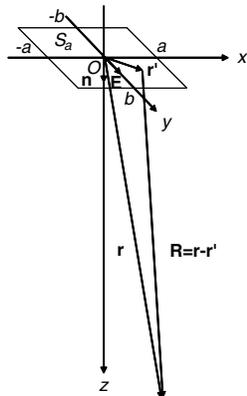


Figure 3: Radiation field by rectangular aperture.

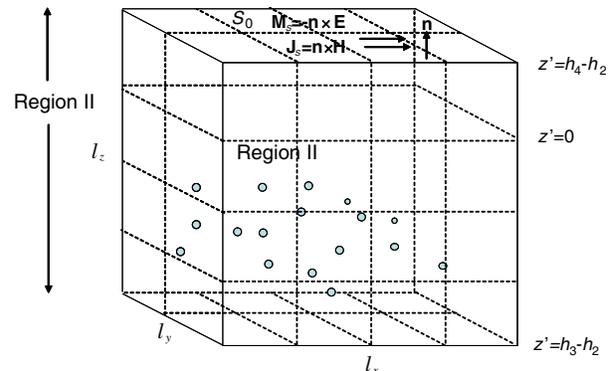


Figure 4: Three dimensional analysis model of FDTD method.

antenna at top boundary surface around surface S_0 of region II for FDTD calculation of scattering and reflection in region II of clouds and rains is considered. Current distribution is assumed to be Gaussian on $z'' = 0$ plane

$$J_y(x'', y'', z'' = 0, t) = J_0 e^{-\frac{(x''-x_0)^2}{r_0^2}} e^{-\frac{(y''-y_0)^2}{r_0^2}} e^{-\frac{(t-t_0)^2}{t_w^2}} \cos 2\pi ft \quad (4)$$

In case of normal and oblique incidence for incident angle $x'' = x' \cos \theta - z' \sin \theta$, $z'' = x' \sin \theta + z' \cos \theta$. Mur's boundary condition is applied to obtain the electromagnetic fields at the boundary of the analysis space. In raindrop region, complex dielectric constant ε^* and complex refractive index n_r^* are given by $\varepsilon^* = \varepsilon + \sigma / (j\omega) = n_r^{*2} \varepsilon_0$, $n_r^* = n_1 - jn_2$ where, $\varepsilon = \varepsilon_0 (n_1^2 - n_2^2)$, $\sigma = 2\varepsilon_0 \omega n_1 n_2$. In this analysis, the shape of a raindrop is assumed to be cubic and generated by giving the length of a side length of squares a_i , positions of one apex (x_i, y_i, z_i) and relative dielectric constant $\varepsilon_r(i, j, k) = \varepsilon(i, j, k) / \varepsilon_0$.

By dividing the total FDTD region II into subdomains, FDTD equations at point (i, j, k) except the boundary surface is expressed as

$$\mathbf{x}^{n+1}(i, j, k) = \mathbf{a}\mathbf{x}^n(i, j, k) + \boldsymbol{\varphi}^n(i, j, k) \quad (5)$$

Here, we define vectors $\mathbf{x}^n(i, j, k)$ and $\boldsymbol{\varphi}^n(i, j, k)$, using coefficient matrix \mathbf{a} . By definition of vectors \mathbf{X}^n and $\boldsymbol{\Phi}^n$ as

$$\begin{aligned} \mathbf{X}^n &= {}^t [\mathbf{x}^n(0, 0, 0), \mathbf{x}^n(0, 0, 1), \dots, \mathbf{x}^n(N_x, N_y, N_z)], \\ \boldsymbol{\Phi}^n &= {}^t [\boldsymbol{\varphi}^n(0, 0, 0), \boldsymbol{\varphi}^n(0, 0, 1), \dots, \boldsymbol{\varphi}^n(N_x, N_y, N_z)], \end{aligned}$$

we have $\mathbf{X}^{n+1} = \mathbf{A}\mathbf{X}^n + \boldsymbol{\Phi}^n$ for FDTD equation where matrix \mathbf{A} consists of diagonal small matrix \mathbf{a} . Here, n is the time step and when $n = 0, 1, 2, \dots, T$, we have

$$\begin{bmatrix} I & 0 & 0 & 0 & \dots & \dots & \dots & 0 & 0 \\ -\mathbf{A} & I & 0 & 0 & \dots & \dots & \dots & 0 & 0 \\ 0 & -\mathbf{A} & I & 0 & \dots & \dots & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & & & -\mathbf{A} & I & 0 \\ 0 & 0 & 0 & 0 & \dots & \dots & -\mathbf{A} & I \end{bmatrix} \begin{bmatrix} \mathbf{X}^0 \\ \mathbf{X}^1 \\ \vdots \\ \mathbf{X}^{T-1} \\ \mathbf{X}^T \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \boldsymbol{\Phi}^0 \\ \vdots \\ \boldsymbol{\Phi}^{T-2} \\ \boldsymbol{\Phi}^{T-1} \end{bmatrix} \quad (6)$$

Eq. (6) shows that the computation of FDTD is replaced to the solution of huge system of linear equation [7].

For fundamental calculation, we study simple case of $f = 20$ GHz, $x_0 = y_0 = 0$ m, $r_0 = 0.1$ m, $t_0 = 0.15$ ns, $t_w = 0.06$ ns, $J_0 = 1$ A/m². Here, FDTD area is $-\ell_x/2 \leq x' \leq \ell_x/2$, $-\ell_y/2 \leq y' \leq \ell_y/2$, $z'_1 \leq z' \leq z'_2$ where $\ell_x = N_x \Delta s = 0.07$ (m), $\ell_y = N_y \Delta s = 0.07$ (m), rain area is $z'_1 \leq z' \leq 0$, $z'_1 = h_3 - h_2 = -N_0 \Delta s = -0.05$ (m) and $z'_2 = h_4 - h_2 = N_z \Delta s = 0.025$ (m). Here, $N_x = N_y = 150$, $N_0 = 100$ and $N_z = 50$.

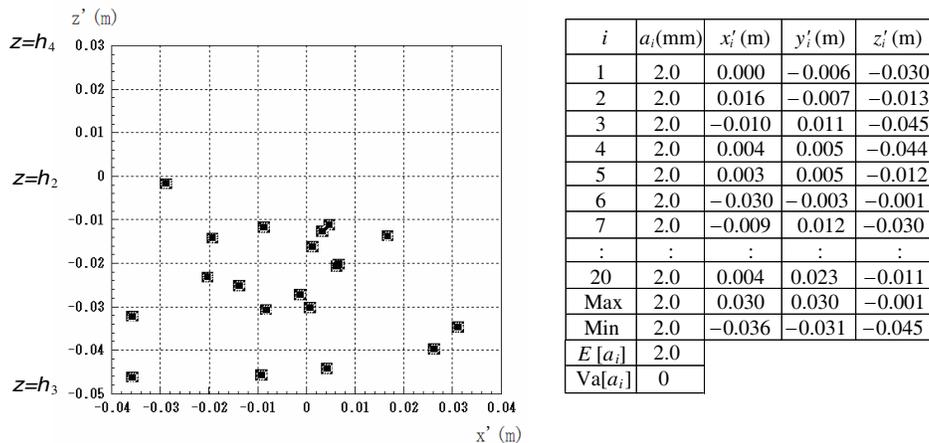


Figure 5: Cross sectional view of raindrop distribution, $N = 20$, $a_i = 2$ (mm), $n_r^* = 6.46 - j2.81$.

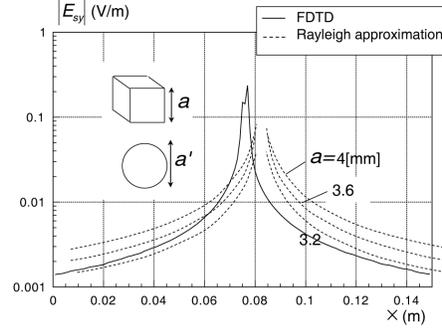


Figure 6: Comparison of scattered field amplitudes of cubic and spherical raindrop ($y = z = 0.075$ m), the length of a side length of cubic model $a = 2$ mm.

Table 1: Simulation parameters.

Parameters	Values
f : Frequency of incident wave	20 GHz
λ : Wavelength of incident wave	0.015 m (c/f)
ℓ_x : Length of analysis space (x')	0.075 m (5λ)
ℓ_y : Length of analysis space (y')	0.075 m (5λ)
ℓ_z : Length of analysis space (z')	0.075 m (5λ)
V : Volume of rain region ($\ell_x \times \ell_y \times (h_2 - h_3)$)	$2.81 \times 10^{-4} \text{ m}^3$
x_0 : Center point of beam (x')	0 m
y_0 : Center point of beam (y')	0 m
r_0 : Beam waist beam spot size	0.1 m (6.7λ)
Δ_s : Length of a cell	5×10^{-4} ($\lambda/30$)
Δ_t : Time increment	0.75 ps ($T/66.4$)
N : Number of raindrops	20
a_i : Length of a side of a raindrop	2 mm
n_r : Complex refractive index of raindrops	$6.46 - j2.81$

Rayleigh scattering approximation and numerical analysis for scattered fields concerned with a single raindrop are compared to test the accuracy of FDTD analysis as in Fig. 6.

These figures show the fundamental characteristics for microwave remote sensing of clouds and rain region.

Figures 7 and 8 show the reflection field from raindrops in cloud and rain region II using cloud and rain model shown in Fig. 5. Fig. 7 shows the case of normal incidence ($\theta = 0^\circ$) and Fig. 8 shows the case of oblique incidence ($\theta = 5.7^\circ$). Corresponding satellite position is $x' = 0$ and $x' = -5 \times 10^3$ (m) when satellite height is $z = h_1 = 5 \times 10^5$ (m).

Using the reflection field on virtual surface ($z' = h_4 - h_2$), surface equivalent current is calculated and the received field at satellite antenna is obtained.

4. REFLECTION FIELD RECEIVED BY SATELLITE ANTENNA

The received electric field at receiving point is derived by surface equivalent current on virtual surface $z' = h_4 - h_2$. Electric field components in spherical coordinate are

$$E_\theta(\mathbf{r}, t) \cong -\eta_0 W_\theta(\mathbf{r}, t) - U_\phi(\mathbf{r}, t), \quad E_\phi(\mathbf{r}, t) \cong -\eta_0 W_\phi(\mathbf{r}, t) + U_\theta(\mathbf{r}, t) \quad (7)$$

where vector potentials generated by surface equivalent current are given by

$$\begin{aligned} \mathbf{W}(\mathbf{r}, t) &\cong \frac{1}{4\pi cr} \frac{\partial}{\partial t} \left[\iint_{S_0} \mathbf{J}_s \left(\mathbf{r}, t - \frac{r - \mathbf{r}' \cdot \hat{\mathbf{r}}}{c} \right) dS' \right], \\ \mathbf{U}(\mathbf{r}, t) &\cong \frac{1}{4\pi cr} \frac{\partial}{\partial t} \left[\iint_{S_0} \mathbf{M}_s \left(\mathbf{r}, t - \frac{r - \mathbf{r}' \cdot \hat{\mathbf{r}}}{c} \right) dS' \right] \end{aligned} \quad (8)$$

where $\eta = \sqrt{\mu_0/\epsilon_0}$.

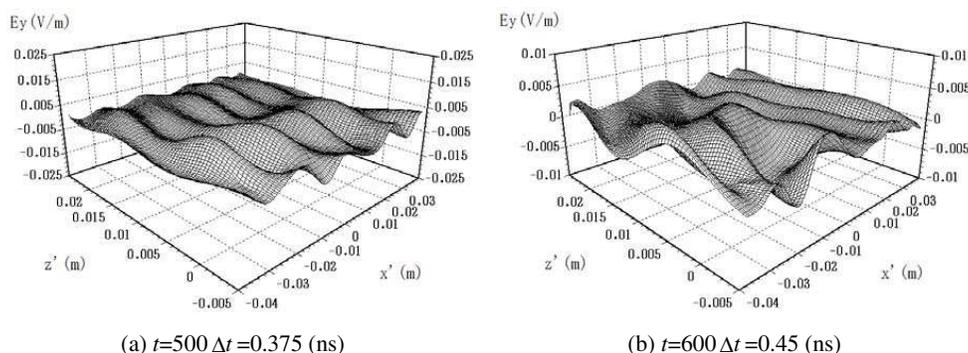
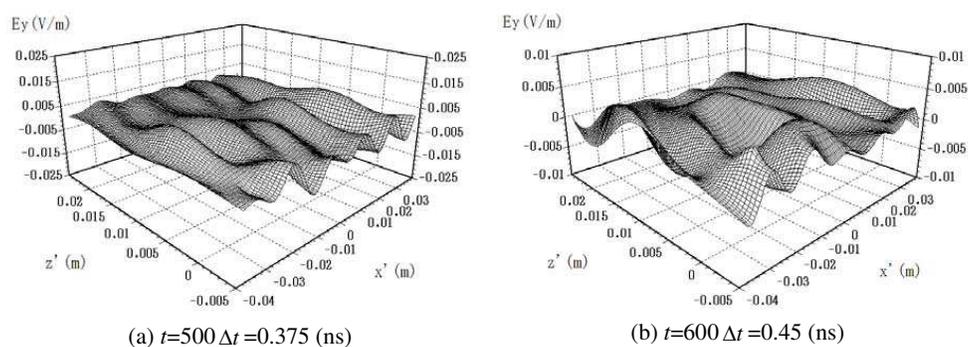
Figure 7: Electric field E_y at $y' = 0.0375$ (m), $\theta = 0^\circ$.Figure 8: Electric field E_y at $y' = 0.0375$ (m), $\theta = 5.7^\circ$.

Image information processing of cloud and rain distribution by satellite antenna system is investigated by auto-correlation function for received electric fields along scanning antenna route, as

$$R(\tau_x) = \int_{-T/2}^{T/2} E_i(t) E_j^*(t + \tau_x) dt \quad (9)$$

when measuring time is T and received fields at different position x_i and x_j of scanning antenna are E_i , E_j .

5. CONCLUSION

In this paper, electromagnetic backscattering by single raindrop and randomly distributed raindrops in clouds and rain region are analyzed by FDTD method. These results show the fundamental characteristics because the scattered wave which reaches the receiving antenna contributes to SAR image of clouds and rain region. To obtain the electromagnetic scattering in large area which contains hundreds of raindrops, parallel and successive computation of FDTD is studied.

REFERENCES

1. Oguchi, T., "Electromagnetic wave propagation and scattering in rain and other hydrometers," *IEEE Proc.*, Vol. 71, 1029–1078, 1983.
2. Oliver, C. J., "Information from SAR images," *Applied Physics*, Vol. 24, 1493–1514, 1991.
3. Miyazaki, Y. and T. Kawakami, "Electromagnetic backscattering of microwaves by objects on random ground surfaces in satellite remote sensing," *Proc. of CJMW 2000*, 217–220, 2000.
4. Sonoda, J. and Y. Miyazaki, "Analysis of electromagnetic scattering using FDTD method and near to far field transformation for satellite remote sensor," *Technical Paper of IEE Japan*, Vol. EMT-00-102, 19–24, 2000.
5. Azuma, H. and Y. Miyazaki, "FDTD analysis and simulation experiment of TV signal disturbance characteristics due to airplanes," *Technical Paper of IEE Japan*, Vol.120-C, No. 1, 61–67, 2000.
6. Takahashi, K. and Y. Miyazaki, "FDTD computer simulation of millimeter wave driving radar for ITS," *11th World Congress on ITS Nagoya*, Vol. 4083, 6, 2004.

7. Rodriguez, G., Y. Miyazaki, and N. Goto, “Matrix-based FDTD parallel algorithm for big areas and its applications to high-speed wireless communications,” *IEEE Trans. Antennas & Propagat.*, Vol. 54, No. 3, 785–796, 2006.

THz Applications for the Engineering Approach to Modelling Frequency Dispersion within Normal Metals at Room Temperature

Stepan Lucyszyn and Yun Zhou
Imperial College London, UK

Abstract— The use of the accurate classical relaxation-effect model for frequency dispersion in normal metals at room temperature with THz structures can be mathematically cumbersome and not insightful. Recent work has demonstrated that it is possible to dramatically simplify otherwise complex analysis and allows for a much deeper insight to be gained into the classical relaxation-effect model. This paper gives applications for this engineering approach. Here, using the concept of Q -factor for a metal, the synthesized equivalent transmission line model is validated. Then, using the concept of complex skin depth, analysis is performed on hollow metal-pipe rectangular waveguides and their associated cavity resonators. This work proves that the mathematical modelling of THz structures can be greatly simplified by taking an electrical engineering approach to these electromagnetic problems.

1. INTRODUCTION

The classical relaxation-effect model accurately describes the frequency dispersive nature of normal metals at room temperature. Indeed, the robustness of this model has been tested in recent years for normal metals at room temperature from dc to terahertz frequencies [1–4]. For example, it has been used to validate measurements [1], alternative frequency dispersion models [2, 3] and commercial electromagnetic simulation software [4]. However, when compared to the over-simplified classical skin-effect model (with associated variables indicated by the suffix “ o ”), the classical relaxation-effect modelling approach (with associated variables indicated by the suffix “ R ”) for THz structures at room temperature can be mathematically cumbersome and not insightful.

Given an angular frequency $\omega = 2\pi f$, where f is the frequency of the driving electromagnetic fields, and the phenomenological temperature-dependent scattering relaxation time for the free electrons (i.e., mean time between collisions) τ in normal metals at room temperature, a recent paper addressed analysis for all values of $\omega\tau$ and even allowed accurate simplifying assumptions to be made in the region of $0 \leq \omega\tau \leq 2$. In this paper, the equivalent transmission line model is tested and metal-pipe rectangular waveguide structures are modelled [4] using the concepts of Q -factors and complex skin depth for normal metals at room temperature [5].

2. EQUIVALENT TRANSMISSION LINE MODEL VALIDATION

The modelling of frequency dispersion in a metal using a synthesized equivalent transmission line model has been recently reported but not extensively validated [5]. Therefore, the modelling of an electromagnetic wave travelling to an arbitrary depth of one wavelength will be modelled here using $ABCD$ parameter matrices and its results converted into S -parameters for comparison.

For a uniform passive transmission line, the $ABCD$ parameter matrix is given by the following textbook representation:

$$[ABCD] = \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} \cosh(\gamma l) & Z_o \sinh(\gamma l) \\ Y_o \sinh(\gamma l) & \cosh(\gamma l) \end{bmatrix} \quad Y_o = \frac{1}{Z_o} [S] \quad (1)$$

where, for any arbitrary uniform transmission line, γ is the complex propagation constant, l is the physical length and Z_o is the complex characteristic impedance. For the equivalent transmission line model, the characteristic impedance equates to the surface impedance of the metal, i.e., $Z_o = Z_S$, when the electromagnetic waves decay exponentially with distance into the metal (as found with normal metals at room temperature) [5]. The corresponding S -parameters are given by the following textbook representation:

$$S_{11} = S_{22} = \frac{(Z_T - Z_T^*) \cosh(\gamma l) + (Z_o - Z_T^* Z_T Y_o) \sinh(\gamma l)}{D} \quad (2)$$

$$\begin{aligned}
 S_{21} &= S_{12} = \frac{2R_T}{D} \\
 D &= 2Z_T \cosh(\gamma l) + (Z_o + Z_T^2 Y_o) \sinh(\gamma l) [\Omega] \\
 Z_T &= R_T + jX_T [\Omega] \quad \text{and} \quad Z_T^* = R_T - jX_T [\Omega]
 \end{aligned}$$

where Z_T is the complex terminating impedance at both ends of the transmission line and $j = \sqrt{-1}$ is the complex operator. When the terminating impedance is equal to the complex conjugate of the characteristic impedance of the transmission line (i.e., $Z_T = Z_o^*$):

$$\begin{aligned}
 S_{11} &= S_{22} = \frac{-je^{-2\gamma l} Q_c (1 + jQ_c)}{1 + (e^{-\gamma l} Q_c)^2} \cong -je^{-2\gamma l} Q_c (1 + jQ_c) \quad \text{with} \quad \Re\{\gamma\}l > 2.3 \\
 S_{21} &= S_{12} = \frac{e^{-\gamma l} (1 + jQ_c)}{1 + (e^{-\gamma l} Q_c)^2} \cong e^{-\gamma l} (1 + jQ_c) \quad \text{with} \quad \Re\{\gamma\}l > 2.3 \\
 \text{Insertion Loss} &= 20 \log|s_{21}| \text{ [dB]} \Rightarrow G_{\text{MAX}} \text{ [dB]} \\
 \text{Attenuation} &= 20 \log(\Re\{s_{21}\}) \text{ [dB]} \Rightarrow \text{Absorption Loss} = 20 \log(e^{-\Re\{\gamma\}l}) \text{ [dB]}
 \end{aligned} \tag{3}$$

where Q_c is the component Q -factor for a normal metal [5]. For a depth of one wavelength (i.e., $l = \lambda$), with the classical relaxation-effect model, the propagation constant per unit wavelength and voltage-wave transmission coefficients $S_{21} = S_{12}$ are given by the following simple expressions:

$$\begin{aligned}
 \gamma_R \lambda_R &= 2\pi \frac{\gamma_R}{\Im\{\gamma_R\}} = 2\pi(Q_{cR} + j) [\lambda_R^{-1}] \\
 S_{11} &= S_{22} \cong 0 \quad \text{and} \quad S_{21} = S_{12} \cong e^{-2\pi Q_{cR}} (1 + jQ_{cR})
 \end{aligned} \tag{4}$$

$Q_{cR} = \frac{\Re\{\gamma_R\}}{\Im\{\gamma_R\}} = (1 + \xi Q_{mR})^2$ where $\xi = \sqrt{\sqrt{u^{-4} + u^{-2}} + u^{-1}} - u^{-1}$ and $Q_{mR} = \frac{\Re\{\gamma_R^2\}}{\Im\{\gamma_R^2\}} = \omega\tau = u$ where Q_{cR} and Q_{mR} are the component and material Q -factors, respectively, for a normal metal using the classical relaxation-effect model [5].

The equivalent transmission line model for the classical relaxation-effect model is shown in Fig. 1. The corresponding $ABCD$ parameter matrix for the elementary lumped-element circuit is given in (5):

$$\begin{aligned}
 [ABCD_1] &= \begin{bmatrix} A_1 & B_1 \\ C_1 & D_1 \end{bmatrix} = \begin{bmatrix} 1 + j\omega\mu\sigma_R\Delta z^2 & j\omega\mu\Delta z \\ \sigma_R\Delta z & 1 \end{bmatrix} \\
 \Delta z &= \frac{\lambda_R}{N} = \frac{\pi\delta_{S_o}}{N} Q_{cR}^{3/2} \left[1 + \frac{1}{Q_{cR}^2} \right]
 \end{aligned} \tag{5}$$

where $\mu = \mu_o\mu_r$ is the magnetic permeability, μ_o is the magnetic permeability of free space, μ_r is the relative magnetic permeability, $\sigma_R = \sigma_o/(1 + j\omega\tau)$ is the intrinsic bulk conductivity due to free charge carriers, $\Delta z = f(\omega\tau)$ represents the propagation distance for one elementary lumped-element circuit (this must be very much shorter than the wavelength λ of the electromagnetic wave within the metal), δ_{S_o} is the normal skin depth calculated using the classical skin-effect model and N is the number of cascaded elementary lumped-element circuit sections.

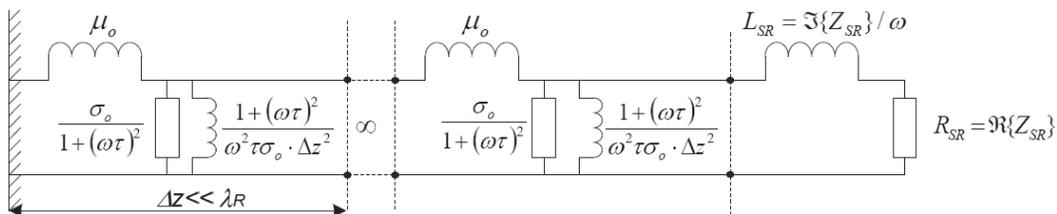


Figure 1: Equivalent transmission line model for the classical relaxation-effect model, showing expressions for the distributed-element parameters [5].

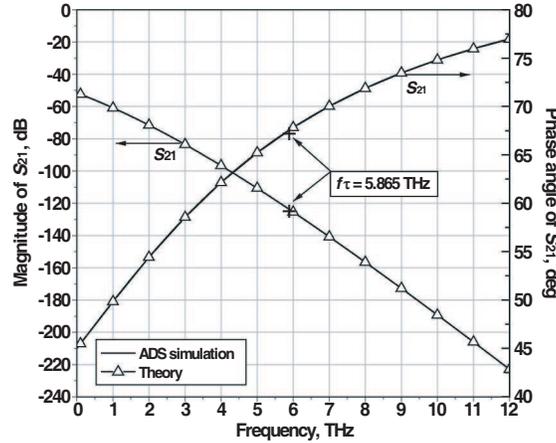


Figure 2: Frequency responses for the forward voltage-wave transmission coefficient for gold at room temperature, for $N = 800$ sections.

For the synthesized equivalent transmission line model, the overall $ABCD$ parameter matrix can be calculated from:

$$[ABCD_N] = \prod_{i=1}^N [ABCD_1(i)] \quad (6)$$

A good test for the validity of the equivalent transmission line model, for both the reflection and transmission characteristics, is to extract the surface impedance $Z_{SR}|_N$ and propagation constant per unit wavelength $\gamma_R \lambda_R|_N$, respectively, from the overall $ABCD$ parameter matrix $[ABCD_N]$:

$$Z_{SR}|_N = \sqrt{\frac{B_N}{C_N}} \quad \text{and} \quad \gamma_R \lambda_R|_N = \cosh^{-1}(A_N) = \sinh^{-1}\left(\frac{B_N}{Z_{SR}|_N}\right) \quad (7)$$

The results from theory, using (1) to (4), can be compared with those from the corresponding synthesized equivalent transmission line model, extracted using commercial circuit simulation software (Agilent Technologies EEsof ADS), for gold at room temperature (having $\sigma_o = 4.517 \times 10^7$ [S/m], $\tau = 27.135$ [fs], $\mu_r = 0.99996 \cong 1$ [1]). Fig. 2 shows the frequency responses for the forward voltage-wave transmission coefficient. It can be clearly seen that there is an excellent fit between theory and that from the equivalent transmission line model having $N = 800$ sections.

It is interesting to note that the phase angle for S_{21} increases with frequency, unlike a normal transmission line. As a result, the differential-phase group delay will be negative. However, this does not break the laws of physics, since the transit time propagation delay is still positive.

A more detailed comparison between theory, using (1) to (4), results from the $ABCD$ parameter matrix calculation, using (5) to (7), and commercial circuit simulation software (Microwave Office®) is given in Table 1, for an arbitrary frequency of $\omega\tau = 1$, i.e., $f = 1/(2\pi\tau) = 5.865$ THz.

The reason for choosing this frequency is for mathematical simplicity, since $Q_{cR}(\omega\tau = 1) = (1 + \sqrt{2})$ [5]; hence, the propagation constant per unit wavelength simplifies to:

$$\gamma_R \lambda_R(\omega\tau = 1) = 2\pi \left[(1 + \sqrt{2}) + j \right] [\lambda_R^{-1}] \quad (8)$$

The results of the test in (7) can be seen in Table 1, for $N = 400$ and 800. It can be seen that the reflection characteristic is identical to the theoretical value, to 8 decimal places, with 800 sections. However, the transmission characteristic is only determined to 3 decimal places, with 800 sections. Note that an extracted value of $\Im\{\gamma_R\} \lambda_R = 0$ gives the same answer as the theoretical value of 2π . With regard to the S -parameters and their associated power characteristics, it can be seen in Table 1 that there is an excellent agreement between the results from the commercial circuit simulation software and the $ABCD$ parameter matrix calculations. Also, the results are closer to the corresponding theoretical values as the number of sections increase from 400 to 800.

Table 1: Comparison of modelled parameters for gold at room temperature, at 5.865 THz, determined from theory, direct $ABCD$ parameter matrix calculations and synthesized equivalent transmission line models using commercial circuit simulation software.

Parameters for $l = \lambda_R, \omega\tau = 1, Z_T = Z_{SR}^*$	Theory	Synthesized Transmission Line			
		ABCD Parameter Matrix Calculations		Microwave Office®	
No. of Sections, N	---	400	800	400	800
$Z_{SR} [\Omega]$	0.46079043 +j1.11244650	0.46079964 +j1.11246875	0.46079043 +j1.11244650	---	---
$\gamma_R \lambda_R [\lambda_R^{-1}]$	15.1689 +j6.2832	15.1681 -j0.0013	15.1685 -j0.0004	---	---
S_{11}	0	0.0006 +j0.0271	0.0002 +j0.0135	0.0006 +j0.0270	0.0001 +j0.0133
Return Loss [dB]	-247.51	-31.35	-37.41	-31.37	-36.9
$S_{21} = e^{-2\pi(1+\sqrt{2})}$ $\cdot [1 + j(1 + \sqrt{2})] \times 10^7$	2.583 +j6.237	2.579 +j6.242	2.582 +j6.240	2.564 +j6.238	2.567 +j6.239
G_{MAX} [dB]	-123.4126	-123.4093	-123.4099	-123.41	-123.4
Attenuation Loss [dB]	-131.756	-131.748	-131.752	-131.9	-131.9
$\forall S_{21} = \tan^{-1}(1 + \sqrt{2}) [^\circ]$	67.50	67.55	67.52	67.58	67.56

3. MODELLING OF WAVEGUIDE STRUCTURES AT ROOM TEMPERATURE

Equations for the surface impedance of a metal, the attenuation constant of uniform metal-pipe rectangular waveguides (MPRWGs) and the Q -factor of resonant waveguide cavities were previously given for accurate calculations at THz frequencies [4].

3.1. Metal-pipe Rectangular Waveguides

With hollow metal-pipe rectangular waveguides operating in the dominant TE_{10} mode of propagation, above its cut-off angular frequency $\omega_c|_{TE_{10}}$, the attenuation constant attributed to the finite conductivity of the metal walls is given by the well-known textbook expression:

$$\alpha_{WG}|_{TE_{10}} = \frac{R_S(\omega)}{\eta_o} G_{10}(\omega) \text{ [Np/m]} \quad (9)$$

$$Z_S(\omega) = \sqrt{\frac{j\omega\mu_o\mu_r}{\sigma + j\omega\varepsilon_o}} \text{ [\Omega/square]}; \quad \eta_o = \sqrt{\frac{\mu_o}{\varepsilon_o}} \text{ [\Omega]}; \quad G_{10}(\omega) = \frac{1 + 2\frac{b}{a} \left(\frac{\omega_c|_{TE_{10}}}{\omega}\right)^2}{b\sqrt{1 - \left(\frac{\omega_c|_{TE_{10}}}{\omega}\right)^2}} \text{ [m}^{-1}] \quad (10)$$

where $Z_S(\omega) = R_S(\omega) + jX_S(\omega)$ is the surface impedance of the metal, $R_S = \Re\{Z_S\}$ is the surface resistance, $X_S = \Im\{Z_S\}$ is the surface reactance, η_o is the intrinsic impedance for a plane wave in free space, ε_o is the electric permittivity of free space, $\sigma \equiv \sigma' - j\sigma''$ is the intrinsic bulk conductivity of the metal due to free charge carriers, and $G_{10}(\omega)$ is the geometric factor for the TE_{10} mode with this particular structure. With reference to Fig. 3, the internal width and height dimensions are a and b , respectively.

Now, the attenuation constant for this uniform guided-wave structure can be calculated using the over-simplified classical skin-effect model. Using this approach the resulting errors, relative to the values calculated using the classical relaxation-effect model, were previously determined by the following complicated expression [4].

$$E_{\alpha_o}(\omega) = \left(\frac{\alpha_{WG_o} - \alpha_{WGR}}{\alpha_{WGR}} \right) \cdot 100\% \cong \left[\sqrt{\sqrt{1 + (\omega\tau)^2} + \omega\tau} - 1 \right] \cdot 100\% \quad [4] \quad (11)$$

However, in contrast, a relatively simple expression for this error can now be derived using the

concept of Q -factor for normal metals at room temperature [5]:

$$E_{\alpha_o}(\omega) \cong \left(\frac{R_{S_o} - R_{SR}}{R_{SR}} \right) \cdot 100\% \quad \text{with} \quad Z_{SR} = R_{S_o} \left[\frac{1}{(1 + \xi\omega\tau)} + j(1 + \xi\omega\tau) \right] \quad (12)$$

$$\therefore E_{\alpha_o}(\omega) = \left(\sqrt{Q_{cR}} - 1 \right) \cdot 100\% = \xi Q_{mR} \cdot 100\% \approx 0.539 Q_{mR} \cdot 100\% \quad \text{for} \quad 0 \leq \omega\tau \leq 2 \quad [5] \quad (13)$$

For example, using (13), the calculated error using this simple approximation is 110% at 12 THz (i.e., at $\omega\tau = 2.046$) and this can be compared with the calculated error, using (11), of 108% at the same frequency.

3.2. Cavity Resonators

The unloaded Q -factor for a hollow metal-pipe rectangular waveguide cavity resonator in the mnl mode is given as:

$$Q_U(\omega'_o)|_{\text{TE}_{mnl}} = \frac{\omega_{I,mnl}}{R_S(\omega'_o)} \Gamma_{mnl} \quad (14)$$

where $\omega_{I,mnl}$ is the ideal angular resonant frequency for the TE_{mnl} modes in an ideal (lossless) hollow cavity and $R_S(\omega'_o)$ is the surface resistance of the metal walls at the angular resonant frequency ω'_o . Here, overall frequency detuning of ω'_o from $\omega_{I,mnl}$ is because of both perturbation and frequency detuning due to ohmic losses [4]. The corresponding geometric factor Γ_{mnl} for this particular resonant cavity structure is given by:

$$\Gamma_{mnl} = \mu_o \psi_{mnl} [H] \quad \text{where} \quad \psi_{mnl} = \frac{\lambda_{I,mnl}}{8} \left[\frac{2b(a^2 + d^2)^{3/2}}{2b(a^3 + d^3) + ad(a^2 + d^2)} \right] \quad [\text{m}] \quad (15)$$

where d is the internal length dimension for the cavity resonator. The corresponding textbook expression for the wavelength in free space $\lambda_{I,mnl}$ at the associated resonant frequencies for the TE_{mnl} modes in an ideal hollow cavity $f_{I,mnl}$ is give by:

$$\lambda_{I,mnl} = \frac{c}{f_{I,mnl}} \quad \text{where} \quad f_{I,mnl} = \frac{c}{2} \sqrt{\left(\frac{m}{a}\right)^2 + \left(\frac{n}{b}\right)^2 + \left(\frac{l}{d}\right)^2} \equiv \frac{\omega_{I,mnl}}{2\pi} \quad (16)$$

where c is the speed of light in free space. For example, with the dominant TE_{101} mode, it can be seen that (15) and (16) reduce to:

$$\lambda_{I,101} = \frac{2ad}{\sqrt{a^2 + d^2}} = \begin{cases} \sqrt{2}a & \text{for } b = a/2 \text{ and } d = a \quad [5] \\ \sqrt{8/3}a & \text{for } b = a/2 \text{ and } d = \sqrt{2}a \quad [4] \end{cases} \quad (17)$$

$$\psi_{101} = \frac{abd(a^2 + d^2)}{2[2b(a^2 + d^2) + ad(a^2 + d^2)]} = \begin{cases} a/8 \equiv \text{Volume/Area} & \text{for } b = a/2 \text{ and } d = a \quad [5] \\ 3a/[2(\sqrt{2} + 10)] & \text{for } b = a/2 \text{ and } d = \sqrt{2}a \quad [4] \end{cases} \quad (18)$$

With the classical skin-effect model, it will be seen that the surface resistance is given by the textbook expression:

$$R_S(\omega'_o) \rightarrow R_{S_o}(\omega'_{oo}) = \frac{\omega'_{oo} \mu_o \mu_r \delta_{S_o}(\omega'_{oo})}{2} \quad (19)$$

The resulting unloaded Q -factor is, therefore, given as:

$$Q_{U_o}(\omega'_{oo})|_{\text{TE}_{mnl}} = \frac{2}{\mu_r \delta_{S_o}(\omega'_{oo})} \left(\frac{\omega_{I,mnl}}{\omega'_{oo}} \right) \psi_{mnl} \quad (20)$$

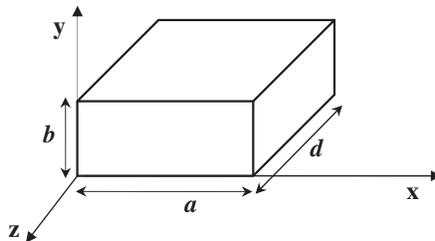


Figure 3: Internal spatial variable definitions for a uniform hollow MPRWG.

With the classical relaxation-effect model, it has recently been shown in [5] that the surface resistance becomes:

$$R_S(\omega'_o) \rightarrow R_{SR}(\omega'_{oR}) = \omega'_{oR} \mu_o \mu_r \Im \{ \delta_{cR}(\omega'_{oR}) \} \quad (21)$$

where complex skin depth is given by $\delta_{cR} \equiv \delta'_{cR} - j\delta''_{cR}$ [5]. The resulting unloaded Q -factor is, therefore, now given by the following:

$$Q_{UR}(\omega'_{oR})|_{TE_{mnl}} = \frac{1}{\mu_r \Im \{ \delta_{cR}(\omega'_{oR}) \}} \left(\frac{\omega_{I-mnl}}{\omega'_{oR}} \right) \psi_{mnl} \quad (22)$$

Note that (22) represents the formal derivation of $Q_U(\omega'_o)$ in [4], for a hollow rectangular waveguide cavity having metal walls with $\mu_r \neq 1$ and $\omega\tau > 0$; this is also the more general form of $Q_{UR}(\omega_{oR})|_{TE_{101}}$ in [5].

It is interesting to see how the unloaded Q -factor for this component can be represented using the exact expression from the classical skin-effect model (20) at $\omega\tau = 1$, by including both component and material Q -factors for the normal metal walls:

$$Q_{Uo}(\omega'_{oo})|_{TE_{mnl}} = \frac{Q_{Uo}(\omega\tau=1)|_{TE_{mnl}}}{\sqrt{Q_{mR}(\omega'_{oo})}} \quad \text{and} \quad Q_{UR}(\omega'_{oR})|_{TE_{mnl}} = Q_{Uo}(\omega\tau=1)|_{TE_{mnl}} \sqrt{\frac{Q_{cR}(\omega'_{oR})}{Q_{mR}(\omega'_{oR})}} \quad (23)$$

With the unloaded Q -factor calculated using the over-simplified classical skin-effect model the resulting error, relative to the values calculated using the classical relaxation-effect model, was previously determined after undergoing a relatively lengthy process [4]. However, in contrast, a relatively simple expression for this error can now be derived using the concept of Q -factor for normal metals at room temperature:

$$E_{Qo}(\omega'_{oR}) = \left| \frac{Q_{Uo}(\omega'_{oo}) - Q_{UR}(\omega'_{oR})}{Q_{UR}(\omega'_{oR})} \right| \cdot 100\% = \left| \sqrt{\frac{Q_{mR}(\omega'_{oR})}{Q_{mR}(\omega'_{oo}) Q_{cR}(\omega'_{oR})}} - 1 \right| \cdot 100\% \quad (24)$$

$$E_{Qo}(\omega'_{oR}) \cong \left| \frac{1}{\sqrt{Q_{cR}(\omega'_{oR})}} - 1 \right| \cdot 100\% \quad \text{since} \quad \omega'_{oR} \cong \omega'_{oo} \quad (25)$$

$$\therefore E_{Qo}(\omega'_{oR}) \approx \left(\frac{1}{1 + 1.8553 Q_{mR}(\omega'_{oR})^{-1}} \right) \cdot 100\% \quad \text{for} \quad 0 \leq \omega\tau \leq 2 \quad [5] \quad (26)$$

For example, using (26), the calculated error using this simple approximation is 40% at 7.3 THz (i.e., at $\omega\tau = 1.245$) and this can be compared with the error determined using the lengthy technique, described in [4], of 41% at the same frequency.

4. CONCLUSION

This paper has validated the recently introduced engineering approach to modelling frequency dispersion in normal metals at room temperature. While relatively simple examples were previously given, to show how algebraic expressions can be dramatically simplified [5], similar benefits have been demonstrated here with more complicated analytical problems. Using the concept of Q -factor for a metal, the equivalent transmission line model was validated. Then, using the concept of complex skin depth, analysis was performed on hollow metal-pipe rectangular waveguides and their associated cavity resonators. This work proves that the mathematical modelling of THz structures can be greatly simplified by taking an electrical engineering approach to these electromagnetic problems.

ACKNOWLEDGMENT

This work was supported by the UK's Engineering and Physical Sciences Research Council (EPSRC) under Platform Grant EP/E063500/1.

REFERENCES

1. Lucyszyn, S., "Investigation of anomalous room temperature conduction losses in normal metals at terahertz frequencies," *IEE Proceedings — Microwaves, Antennas and Propagation*, Vol. 151, No. 4, 321–329, 2004.

2. Lucyszyn, S., “Investigation of Wang’s model for room temperature conduction losses in normal metals at terahertz frequencies,” *IEEE Trans. on Microwave Theory Tech.*, Vol. 53, 1398–1403, 2005.
3. Lucyszyn, S., “Evaluating surface impedance models for terahertz frequencies at room temperature,” *PIERS Online*, Vol. 3, No. 4, 554–559, 2007.
4. Zhou, Y. and S. Lucyszyn, “HFSSTM modelling anomalies with THz metal-pipe rectangular waveguide structures at room temperature,” *PIERS Online*, Vol. 5, No. 3, 201–211, 2009.
5. Lucyszyn, S. and Y. Zhou, “Engineering approach to modelling frequency dispersion within normal metals at room temperature for THz applications,” *Progress In Electromagnetics Research*, PIER 101, 257–275, 2010.