Amplitude Estimation of Multichannel Signal in Spatially and Temporally Correlated Noise

K. J. Sohn and H. Li
Stevens Institute of Technology, USA
B. Himed
Air Force Research Laboratory/SNRT, USA

Abstract—This paper examines the problem of complex amplitude estimation of a multichannel signal in the presence of colored noise with unknown spatial and temporal correlation. A number of amplitude estimators are developed, including the optimum maximum likelihood (ML) estimator, which involves nonlinear optimization, and several suboptimal but computationally more efficient estimators based on least-squares (LS) or weighted LS (WLS) estimation. The Cramér-Rao bound (CRB) for the estimation problem is presented. Numerical results are presented to illustrate the performance of these estimators with or without training data.

1. Introduction

Amplitude estimation occurs in numerous signal processing applications. A survey of amplitude estimation techniques for sinusoidal signals with known frequencies in colored noise is found in [1]. While [1] is primarily concerned with *single-channel sinusoidal* signals, we consider amplitude estimation of an *arbitrary multichannel* signal observed in space and time using a sensor array. The observed data is contaminated by *a spatially and temporally correlated* disturbance signal with *unknown* correlation. Among other applications, this problem is encountered in an airborne radar system equipped with multiple antennas (e. g., [2]), where the multichannel signal refers to the space-time steering vector of the antenna array, the amplitude refers to the radar cross section (RCS) of a target, and the disturbance lumps together the thermal noise, radar clutter, and other interferences. Amplitude estimation within such a context would be useful for estimating the spatial and temporal correlation of the disturbance, developing effective target detectors, and finding solutions to several other relevant problems.

To account for its temporal and spatial correlation, our approach is to model the disturbance as a multichannel autoregressive (AR) process. Using extensive real radar data, [2] has shown that multichannel AR models are appropriate and offer efficient representation of the disturbance signal in airborne radars. Our *parametric* approach to the modeling of the disturbance is another major distinction compared to the *non-parametric* approach of [1]. Based on the parametric approach, our problem of interest is to find estimates of the signal amplitude, the AR coefficient matrices, and the spatial covariance matrix of the multichannel signal that drives the AR model. In the sequel, we first examine the optimum ML detector, and show that it involves nonlinear optimization. We then introduce several suboptimal but computationally more efficient LS and WLS amplitude estimators, which can be used to initialize the nonlinear searching involved in the ML estimator. The CRB for the estimation problem is presented as a performance baseline. In our numerical comparison of the different estimators, we focus on the case with *no or very limited training* data, which is of particular interest for applications in non-stationary or dense-target environments (e. g., [3]).

2. Data Model and Problem Statement

The observed noisy multichannel signal $\mathbf{x}_0(n)$ can be written as

$$\mathbf{x}_0(n) = \alpha\mathbf{s}(n) + \mathbf{d}(n), \quad n = 0, 1, \dots, N-1, \tag{1}$$

where all vectors are $J\times 1$ vectors, J is the number of spatial channels, N is the number of temporal observations, $\mathbf{s}(n)$ denotes the signal vector that is assumed known but with unknown complex amplitude α, and $\mathbf{d}(n)$ denotes the disturbance that is correlated in space and time. In addition, there are a set of *disturbance-only training* (i. e., $\alpha = 0$) data $\mathbf{x}_k(n)$, $k = 1, 2, \dots, K$ and $n = 0, 1, \dots, N-1$, available to assist amplitude estimation. In radar systems, training data may be obtained from range cells adjacent to the test cell. However, training is generally limited or may even be unavailable, especially in non-stationary or dense-target environments [3]. We consider amplitude estimation with and without training; in the later case, we have $K = 0$.

Let $\mathbf{x}_k \triangleq [\mathbf{x}_k^T(0),\, \mathbf{x}_k^T(1),\, \dots,\, \mathbf{x}_k^T(N-1)]^T$, and $\mathbf{d}$ and $\mathbf{s}$ are formed similarly from $\mathbf{d}(n)$ and $\mathbf{s}(n)$, respectively. It is assumed that the training data $\{\mathbf{x}_k\}_{k=1}^K$ and $\mathbf{d}$ are independent and identically distributed (i.i.d.) with complex Gaussian distribution $\mathcal{CN}(\mathbf{0},\mathbf{R})$, where $\mathbf{R}$ denotes the unknown space-time covariance matrix. A J-channel AR process is used to model the disturbance:

$$\mathbf{x}_k(n)-\alpha\mathbf{s}(n)=-\sum_{p=1}^{P}\mathbf{A}^H(p)\{\mathbf{x}_k(n-p)-\alpha\mathbf{s}(n-p)\}+\boldsymbol{\varepsilon}_k(n),\quad k=0,1,\ldots,K, \tag{2}$$

where $\{\mathbf{A}^H(p)\}_{p=1}^P$ denote the unknown $J\times J$ AR coefficient matrices and $\boldsymbol{\varepsilon}_k(n)$ denotes the driving spatial noise with distribution $\mathcal{CN}(\mathbf{0},\mathbf{Q})$, where $\mathbf{Q}$ denotes the unknown $J\times J$ spatial covariance matrix. With some notational abuse, we have $\alpha=0$ (i.e., disturbance-only) for $k\neq 0$ in (2). To focus on the amplitude estimation problem, we assume the model order P is known. In practice when P is unknown, it can be estimated by using a variety of model selection techniques [4].

The problem is to estimate the amplitude α, which is the *signal parameter* of primary interest, as well as *nuisance parameters* $\{\mathbf{A}^H(n)\}$ and $\mathbf{Q}$, from observations $\{\mathbf{x}_k(n)\}$.

3. Amplitude Estimators

For compact presentation, let $\mathbf{A}^H\triangleq[\mathbf{A}^H(1),\ldots,\mathbf{A}^H(P)]\in\mathbb{C}^{J\times JP}$ which contains all the coefficient matrices involved in the P-th order AR model, $\mathbf{y}_k(n)\triangleq[\mathbf{x}_k^T(n-1),\ldots,\mathbf{x}_k^T(n-P)]^T$ which contains the regression subvectors formed from the observed signal $\mathbf{x}_0$ or the k-th training signal $\mathbf{x}_k$, and $\mathbf{t}(n)\triangleq[\mathbf{s}^T(n-1),\ldots,\mathbf{s}^T(n-P)]^T$, which contains the regression subvectors formed from the steering vector s. In the following, we first consider the optimal ML estimator, followed by the suboptimal LS and WLS estimators.

3.1. Optimal ML Amplitude Estimator

In Appendix 1, we show that the ML estimator of α is given by

$$\hat{\alpha}_{ML}=\min_{\alpha}\left|\hat{\mathbf{R}}_{xx}(\alpha)-\hat{\mathbf{R}}_{yx}^H(\alpha)\hat{\mathbf{R}}_{yy}^{-1}(\alpha)\hat{\mathbf{R}}_{yx}(\alpha)\right|, \tag{3}$$

where the correlation matrices are given by

$$\hat{\mathbf{R}}_{xx}(\alpha)=\sum_{n=P}^{N-1}[\mathbf{x}_0(n)-\alpha\mathbf{s}(n)][\mathbf{x}_0(n)-\alpha\mathbf{s}(n)]^H+\sum_{n=P}^{N-1}\sum_{k=1}^{K}\mathbf{x}_k(n)\mathbf{x}_k^H(n), \tag{4}$$

$$\hat{\mathbf{R}}_{yy}(\alpha)=\sum_{n=P}^{N-1}[\mathbf{y}_0(n)-\alpha\mathbf{t}(n)][\mathbf{y}_0(n)-\alpha\mathbf{t}(n)]^H+\sum_{n=P}^{N-1}\sum_{k=1}^{K}\mathbf{y}_k(n)\mathbf{y}_k^H(n), \tag{5}$$

$$\hat{\mathbf{R}}_{yx}(\alpha)=\sum_{n=P}^{N-1}[\mathbf{y}_0(n)-\alpha\mathbf{t}(n)][\mathbf{x}_0(n)-\alpha\mathbf{s}(n)]^H+\sum_{n=P}^{N-1}\sum_{k=1}^{K}\mathbf{y}_k(n)\mathbf{x}_k^H(n). \tag{6}$$

Although statistically optimal, there is no closed-form expression for the above ML estimate. The cost function (3) is a highly nonlinear bivariate function (α is complex-valued). A brute-force exhaustive search over the two-dimensional parameter space is generally impractical. Alternatively, we can resort to Newton-like iterative nonlinear searches, providing an initial estimate of α is available. Next, we discuss suboptimal estimators that can be used for initialization.

3.2. LS Estimator

A linear LS amplitude estimator based on $\mathbf{x}_0$ only is given by

$$\hat{\alpha}_{\mathrm{LS}}=\frac{\mathbf{s}^H\mathbf{x}_0}{\mathbf{s}^H\mathbf{s}}, \tag{7}$$

which ignores the coloredness of the disturbance signal. Albeit simple, the LS estimator is useful when training is unavailable. In addition, it can be used in combination with the WLS amplitude estimator presented next for improved estimation accuracy.

3.3. WLS Estimator

Suppose we have some initial estimates of $\mathbf{A}$ and $\mathbf{Q}$, denoted by $\hat{\mathbf{A}}$ and $\hat{\mathbf{Q}}$, respectively. Then, as shown in Appendix 2, a WLS amplitude estimator is given by

$$\hat{\alpha}_{\mathrm{WLS}}=\frac{\sum_{n=P}^{N-1}\{\mathbf{s}(n)+\sum_{p=1}^{P}\hat{\mathbf{A}}^H(p)\mathbf{s}(n-p)\}^H\hat{Q}^{-1}\{\mathbf{x}_0(n)+\sum_{p=1}^{P}\hat{A}^H(p)\mathbf{x}_0(n-p)\}}{\sum_{n=P}^{N-1}\{\mathbf{s}(n)+\sum_{p=1}^{P}\hat{\mathbf{A}}^H(p)\mathbf{s}(n-p)\}^H\hat{Q}^{-1}\{\mathbf{s}(n)+\sum_{p=1}^{P}\hat{A}^H(p)\mathbf{s}(n-p)\}}. \tag{8}$$

To find initial estimates $\hat{\mathbf{A}}$ and $\hat{\mathbf{Q}}$, we consider two cases with and without training. First, if training is available (i.e., $K\geq 1$), an ML estimator based on only the training data can be used to estimate $\mathbf{A}$ and $\mathbf{Q}$. Following similar steps in Appendix 1, we can show that the *training-only* ML *estimates* are given by

$$\hat{\mathbf{A}}^H = -\hat{\mathbf{R}}_{yx,t}^H \hat{\mathbf{R}}_{yy,t}^{-1}, \tag{9}$$

$$\hat{\mathbf{Q}} = \frac{1}{K(N-P)}\Big(\hat{\mathbf{R}}_{xx,t} - \hat{\mathbf{R}}_{yx,t}^H \hat{\mathbf{R}}_{yy,t}^{-1} \hat{\mathbf{R}}_{yx,t}\Big), \tag{10}$$

$\hat{\mathbf{R}}_{xx,t} = \sum_{n=P}^{N-1} \sum_{k=1}^{K} \mathbf{x}_k(n)\mathbf{x}_k^H(n)$, and $\hat{\mathbf{R}}_{yy,t}$ and $\hat{\mathbf{R}}_{yx,t}$ are correlation matrices formed similarly as in (5) and (6), however, using only the training signals.

On the other hand, if no training data are available ($K = 0$), we can create artificially one "training signal" by by subtracting. $\hat{\alpha}_{LS}\mathbf{s}(n)$ from the observed signal, i. e.,

$$\bar{\mathbf{x}}_0 \triangleq \mathbf{x}_0 - \hat{\alpha}_{LS}\mathbf{s}$$

where α_{LS} is given by (7). Then, the training-only ML estimator (9) and (10) can be used to estimate $\mathbf{A}$ and $\mathbf{Q}$ as if $K = 1$. Finally, it is noted that once the WLS estimate $\hat{\alpha}_{WLS}$ is obtained, it can be used to update estimates of $\mathbf{A}$ and $\mathbf{Q}$. We can iterate the above procedure a few times.

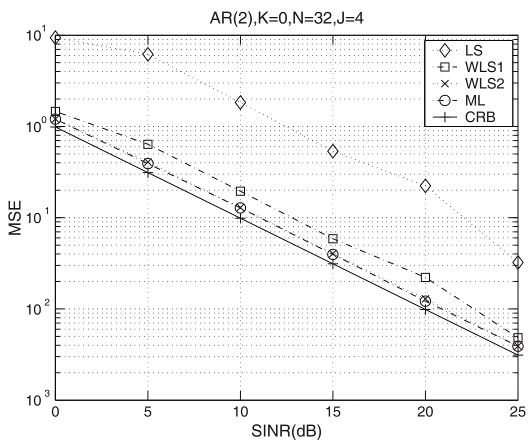

Figure 1: MSE of the signal amplitude estimate $\hat{\alpha}$ versus the input SINR when $J = 4$, $N = 32$, and $K = 0$.

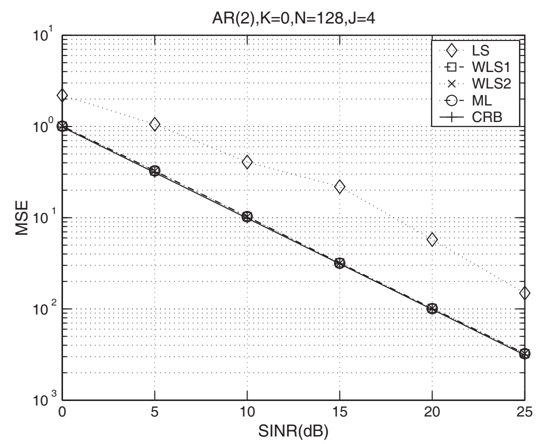

Figure 2: MSE of the signal amplitude estimate $\hat{\alpha}$ versus the input SINR when $J = 4$, $N = 128$, and $K = 0$.

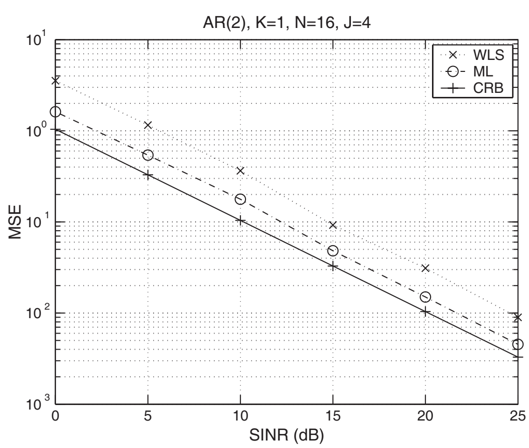

Figure 3: MSE of the signal amplitude estimate $\hat{\alpha}$ versus the input SINR when $J = 4$, $N = 16$, and $K = 1$.

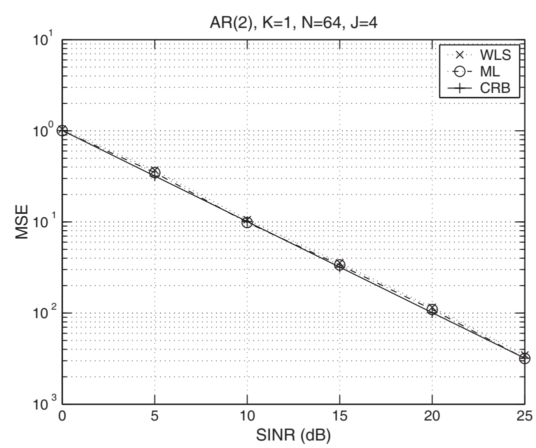

Figure 4: MSE of the signal amplitude estimate $\hat{\alpha}$ versus the input SINR when $J = 4$, $N = 64$, and $K = 1$.

4. Cramér-Rao Bound

The CRB provides a lower bound on the variance of the parameter estimates obtained by any unbiased estimators, and it can be used to access the accuracy of various amplitude estimation schemes. It can be shown that CRB for the signal amplitude estimation is given by

$$\mathrm{CRB}(\alpha)=\left[\sum_{n=P}^{N-1}\left\{\mathbf{s}(n)+\sum_{p=1}^{P}\mathbf{A}^H(p)\mathbf{s}(n-p)\right\}^H\mathbf{Q}^{-1}\left\{\mathbf{s}(n)+\sum_{p=1}^{P}\mathbf{A}^H(p)\mathbf{s}(n-p)\right\}\right]^{-1}. \quad (11)$$

5. Numerical Results

We present numerical results to compare the proposed amplitude estimation schemes. In the following, the SINR is defined as SINR$=|\alpha|^2\mathbf{s}^H\mathbf{R}^{-1}\mathbf{s}$ where $\mathbf{R}$ is the $JN\times JN$ joint space-time covariance matrix of the disturbance $\mathbf{d}$. For the no training case ($K=0$), we consider 1) LS amplitude estimator given by (7); 2) WLS1 amplitude estimator given by (8) with estimates $\hat{\mathbf{A}}$ and $\hat{\mathbf{Q}}$ obtained by using the artificially created training signal; 3) WLS2 estimator which extends WLS1 with another iteration; 4) ML amplitude estimator given by (3). For the case when training is available ($K>0$), we consider 1) WLS amplitude estimator given by (8) along with (9) and (10); 2) ML amplitude estimator given by (3). In both cases, the CRB (11) is included.

Figures 1 and 2 depict the mean-squared error (MSEs) of the various amplitude estimates versus the input SINR. We can see that 1) the MSE of WLS1 estimator is slightly larger than the CRB when $N=32$, but is close to the CRB when $N=128$; 2) as N increases, the MSEs of the WLS1, WLS2, and ML estimators are getting close to the CRB; 3) the MSE of the LS estimator is away from the CRB even at $N=128$.

Figures 3 and 4 depict the MSEs of the various amplitude estimates versus the input SINR when very limited training is available ($K=1$). It is seen that as N increases, the WLS estimates are close to the ML estimates and the CRB.

6. Conclusion

We have examined the problem of amplitude estimation of a known multichannel signal in the presence of a temporally and spatially correlated disturbance signal. To deal with temporal and spatial coloredness, the disturbance signal is modeled as a multichannel AR process with unknown AR coefficient matrices and spatial covariance matrix. We have derived the ML estimate of the signal amplitude which involves two-dimensional nonlinear searches. We have also introduced several suboptimal LS and WLS estimators that can be utilized to initialize the searching.

Appendix 1: Derivation of ML Estimators

The exact maximization of the joint PDF or likelihood function with respect to the unknown parameters produces a set of highly nonlinear equations that are difficult to solve. For large data records, the likelihood function can be well approximated by a joint conditional PDF (12) conditioned on $\{\mathbf{x}_k(n)\}_{n=0}^{P-1}\}, k=0,1,\ldots,K$ [5]. For brevity, the conditional PDF is referred to as the likelihood function henceforth. The loglikelihood function is proportional to (within an additive constant) [6]

$$\begin{aligned}
&-L\ln|\mathbf{Q}|-\sum_{k=1}^{K}\sum_{n=P}^{N-1}\left[\mathbf{x}_k(n)+\sum_{p=1}^{P}\mathbf{A}^H(p)\mathbf{x}_k(n-p)\right]^H\mathbf{Q}^{-1}\left[\mathbf{x}_k(n)+\sum_{p=1}^{P}\mathbf{A}^H(p)\mathbf{x}_k(n-p)\right]\\
&\quad-\sum_{n=P}^{N-1}\left[\{\mathbf{x}_0(n)-\alpha\mathbf{s}(n)\}+\sum_{p=1}^{P}\mathbf{A}^H(p)\{\mathbf{x}_0(n-p)-\alpha\mathbf{s}(n-p)\}\right]^H\mathbf{Q}^{-1}\\
&\quad\times\left[\{\mathbf{x}_0(n)-\alpha\mathbf{s}(n)\}+\sum_{p=1}^{P}\mathbf{A}^H(p)\{\mathbf{x}_0(n-p)-\alpha\mathbf{s}(n-p)\}\right] \qquad (12)
\end{aligned}$$

where $L=(K+1)(N-P)$. Taking the derivative of the likelihood function with respect to $\mathbf{Q}$ and equating the result to zero produce the ML estimates of $\mathbf{Q}$ given α and $\mathbf{A}$:

$$\begin{aligned}
\hat{\mathbf{Q}}(\alpha,\mathbf{A})\triangleq\frac{1}{L}\Bigg\{&\sum_{k=1}^{K}\sum_{n=P}^{N-1}\left[\mathbf{x}_k(n)+\sum_{p=1}^{P}\mathbf{A}^H(p)\mathbf{x}_k(n-p)\right]\left[\mathbf{x}_k(n)+\sum_{p=1}^{P}\mathbf{A}^H(p)\mathbf{x}_k(n-p)\right]^H\\
&+\sum_{n=P}^{N-1}\left[\{\mathbf{x}_0(n)-\alpha\mathbf{s}(n)\}+\sum_{p=1}^{P}\mathbf{A}^H(p)\{\mathbf{x}_0(n-p)-\alpha\mathbf{s}(n-p)\}\right]\\
&\times\left[\{\mathbf{x}_0(n)-\alpha\mathbf{s}(n)\}+\sum_{p=1}^{P}\mathbf{A}^H(p)\{\mathbf{x}_0(n-p)-\alpha\mathbf{s}(n-p)\}\right]^H\Bigg\}. \qquad (13)
\end{aligned}$$

Substituting the above $\hat{\mathbf{Q}}$ back in the likelihood function, we find that maximizing the loglikelihood reduces to minimizing $|\hat{\mathbf{Q}}(\alpha,\mathbf{A})|$. Therefore, the ML estimates of α and $\mathbf{A}$ can be obtained by minimizing $|\hat{\mathbf{Q}}(\alpha,\mathbf{A})|$ with respect to α and $\mathbf{A}$. In turn, we can get the ML estimate of $\mathbf{Q}$ by replacing α and $\mathbf{A}$ with their ML estimates in (13). Next, observe that

$$
\begin{aligned}
L\hat{\mathbf{Q}}(\alpha,\mathbf{A}) &= \hat{\mathbf{R}}_{xx}(\alpha)+\mathbf{A}^H\hat{\mathbf{R}}_{yx}(\alpha)+\hat{\mathbf{R}}_{yx}^H(\alpha)\mathbf{A}+\mathbf{A}^H\hat{\mathbf{R}}_{yy}(\alpha)\mathbf{A}\\
&= \left(\mathbf{A}^H+\hat{\mathbf{R}}_{yx}^H(\alpha)\hat{\mathbf{R}}_{yy}^{-1}(\alpha)\right)\hat{\mathbf{R}}_{yy}(\alpha)\left(\mathbf{A}^H+\hat{\mathbf{R}}_{yx}^H(\alpha)\hat{\mathbf{R}}_{yy}^{-1}(\alpha)\right)^H+\hat{\mathbf{R}}_{xx}(\alpha)-\hat{\mathbf{R}}_{yx}^H(\alpha)\hat{\mathbf{R}}_{yy}^{-1}(\alpha)\hat{\mathbf{R}}_{yx}(\alpha)
\end{aligned}
\tag{14}
$$

where the correlation matrices are given by (4), (5), and (6). Since $\hat{\mathbf{R}}_{yy}(\alpha)$ is non-negative definite and the remaining terms in (14) do not depend on $\mathbf{A}$, it follows that $\hat{\mathbf{Q}}(\alpha,\mathbf{A}) \geq \hat{\mathbf{Q}}(\alpha,\mathbf{A})|_{\mathbf{A}=\hat{\mathbf{A}}(\alpha)}$, where

$$
\hat{\mathbf{A}}^H(\alpha) = -\hat{\mathbf{R}}_{yx}^H(\alpha)\hat{\mathbf{R}}_{yy}^{-1}(\alpha). \tag{15}
$$

When $\hat{\mathbf{Q}}(\alpha,\ \mathbf{A})$ is minimized, the estimate $\hat{\mathbf{A}}^H(\alpha)$ of $\mathbf{A}^H$ will minimize any non-decreasing function including the determinant of $\hat{\mathbf{Q}}(\alpha,\ \mathbf{A})$) [7]. Hence, $\hat{\mathbf{A}}^H(\alpha)$ is the ML estimate of $\mathbf{A}^H$ given α. Replacing $\mathbf{A}^H$ in (14) by $\hat{\mathbf{A}}^H(\alpha)$ yields the ML amplitude estimator (3).

Appendix 2: Derivation of WLS Estimator

Suppose that $\mathbf{Q}$ and $\mathbf{A}$ are known. Then, taking the derivative of the loglikelihood function (12) and setting the result to zero yield

$$
\begin{aligned}
&\alpha\sum_{n=P}^{N-1}\left\{\mathbf{s}(n)+\sum_{p=1}^{P}\mathbf{A}^H(p)\mathbf{s}(n-p)\right\}^H\mathbf{Q}^{-1}\left\{\mathbf{s}(n)+\sum_{p=1}^{P}\mathbf{A}^H(p)\mathbf{s}(n-p)\right\}\\
&-\sum_{n=P}^{N-1}\left\{\mathbf{s}(n)+\sum_{p=1}^{P}\mathbf{A}^H(p)\mathbf{s}(n-p)\right\}^H\mathbf{Q}^{-1}\left\{\mathbf{x}_0(n)+\sum_{p=1}^{P}\mathbf{A}^H(p)\mathbf{x}_0(n-p)\right\}=0.
\end{aligned}
\tag{16}
$$

By solving (16), we have the ML estimate of α. It is different from the ML estimate (3) which assumes $\mathbf{Q}$ and $\mathbf{A}$ are unknown. In practice, $\mathbf{Q}$ and $\mathbf{A}$ are unknown. If these matrices are replaced by their estimates $\hat{\mathbf{Q}}$ and $\hat{\mathbf{A}}$, the resulting WLS estimator is given by (8).

Acknowledgement

This work was supported by the Air Force Research Laboratory under contract number FA8750-05-2-0001.

REFERENCES

1. Stoica, P., H. Li, and J. Li,, "Amplitude estimation of sinusoidal signals: survey, new results, and an application," *IEEE Transactions on Signal Processing*, Vol. 48, 338–352, February 2000.
2. Román, J. R., M. Rangaswamy, D. W. Davis, Q. Zhang, B. Himed, and J. H. Michels, "Parametric adaptive matched filter for airborne radar applications," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 36, No. 2, 677–692, April 2000.
3. Michels, J. H., B. Himed, and M. Rangaswamy, "Robust STAP detection in a dense signal airborne radar environment," *Signal Processing*, special issue on New Trends and Findings in Antenna Array Processing for Radar, Vol. 84, 1625–1636, 2004.
4. Söderström, T. and P. Stoica, *System Identification*, Prentice Hall International, London, UK, 1989.
5. Stoica, P., and R. L. Moses, *Spectral Analysis of Signals*, Upper Saddle River, Pearson Prentice Hall, NJ, 2005.
6. Sohn, K. J., H. Li, and B. Himed, "Parametric rao test for multichannel adaptive signal detection," submitted to *IEEE Transactions on Aerospace and Electronic Systems*, Prentice Hall International, London, UK, 2005.
7. Li, J., B. Halder, P. Stoica, and M. Viberg, "Computationally efficient angle estimation for signals with known waveforms," *IEEE Transactions on Signal Processing*, Vol. 43, No. 9, 2154–2163, September 1995.