COMPUTATIONALLY EFFICIENT MAXIMUM-LIKELIHOOD ESTIMATION OF STRUCTURED COVARIANCE MATRICES

Hongbin Li[1] *Petre Stoica*[2] *Jian Li*[1]

[1]Department of Electrical & Computer Engineering, University of Florida, Gainesville, FL 32611-2044, USA.
[2]Department of Technology, Uppsala University, P. O. Box 27, S-751 03 Uppsala, Sweden.

ABSTRACT

A computationally efficient method for structured covariance matrix estimation is presented. The proposed method provides an Asymptotic (for large samples) Maximum Likelihood estimate of a structured covariance matrix and is referred to as AML. A closed-form formula for estimating Hermitian Toeplitz covariance matrices is derived which makes AML computationally much simpler than most existing Hermitian Toeplitz matrix estimation algorithms. The AML covariance matrix estimator can be used in a variety of applications. We focus on array processing herein and show that AML enhances the performance of angle estimation algorithms, such as MUSIC, by making them attain the corresponding Cramér-Rao bound (CRB) for uncorrelated signals.

1. INTRODUCTION

The covariance matrix of a stationary signal is Hermitian and Toeplitz. However, the conventional sample covariance matrix obtained from a finite number of observations seldom has this structure. Estimating structured covariance matrices is of particular interest in a variety of applications including array processing and time series analysis. An important technique for structured covariance matrix estimation is the maximum likelihood (ML) approach [1, 2, 3] (and the references therein). Since, for Hermitian Toeplitz matrices, a closed-form solution to the exact ML estimation problem does not exist [2], the ML methods presented in the previous studies are iterative and computationally involved, and they are not guaranteed to yield the global optimal solution. To avoid this difficulty, suboptimal methods have been considered, a notable example being the *iterated Toeplitz approximation method* (ITAM) [4]. However, in spite of the fact that the ITAM estimator produces a covariance matrix estimate that is in general closer to the true matrix than the sample covariance matrix in the Frobenius norm sense, there is no guarantee that better application-related performances, such as angle estimation in array processing, will result. In fact, using the ITAM covariance matrix estimate with MUSIC [5] (referred to as ITAM-MUSIC) provides inconsistent (for high SNR) angle estimates (see Section 4 for details).

In this paper we present a computationally efficient method for structured covariance matrix estimation. The method provides an Asymptotic (for large samples) Maximum Likelihood estimate of a structured covariance matrix and will be referred to as the AML algorithm. AML makes use of the *extended invariance principle* (EXIP) for parameter estimation which was introduced in [6]. A closed-form formula is obtained for Hermitian Toeplitz matrix estimation by the AML approach. To assess the performance of the proposed technique, we investigate the impact of using the AML covariance matrix estimate on angle estimation. In particular, we obtain the angle estimates by using MUSIC with the AML covariance matrix estimate and the approach is referred to as AML-MUSIC. By exploiting the Toeplitz structure of the covariance matrix in angle estimation, we implicitly assume the *a priori* knowledge that the incident signals are uncorrelated. With this additional knowledge, the corresponding Cramér-Rao bound (CRB), referred to as the structured CRB or S-CRB, should be lower than the CRB without this knowledge, which is referred to as the unstructured CRB or U-CRB. AML-MUSIC is shown to (asymptotically) attain the S-CRB whereas, as is well-known, using MUSIC with the unstructured sample covariance matrix (referred to as the standard MUSIC) can at best approach the U-CRB.

2. PROBLEM FORMULATION

Assume that $\mathbf{y}(n) \in \mathcal{C}^{M\times 1}, n = 1, 2, \ldots, N$, denote N independent samples of a circularly symmetric complex Gaussian stationary random process with zero-mean and Hermitian Toeplitz covariance matrix $\mathbf{R}(\phi)$ that is a known function of an unknown parameter vector ϕ, where $\phi \in \mathcal{R}^{(2M-1)\times 1}$ consists of the real and imaginary parts of the first column or row of $\mathbf{R}$. The problem of interest herein is to determine a Hermitian Toeplitz matrix estimate $\mathbf{R}(\hat{\phi})$ of $\mathbf{R}(\phi)$ from $\{\mathbf{y}(n)\}$.

The previous situation occurs in many applications including array processing, in which $\{\mathbf{y}(n)\}_{n=1}^{N}$ denote the array output vectors when (i) the incoming signals are uncorrelated and (ii) a uniform linear array (ULA) is employed [5]. Let K uncorrelated signals impinge on a ULA of M sensors, and assume that the additive noise is spatially white and independent of the signals. Then the spatial covariance matrix has the form [5]:

$$\mathbf{R}(\phi) = \mathbf{A}(\theta)\mathbf{S}\mathbf{A}^H(\theta) + \sigma^2\mathbf{I}_M, \qquad (1)$$

where $\mathbf{A} \in \mathcal{C}^{M\times K}$ is the Vandermonde array manifold matrix, $\theta \in \mathcal{R}^{K\times 1}$ denotes the vector consisting of the arrival angles, $\mathbf{S} \in \mathcal{R}^{K\times K}$ denotes the diagonal signal covariance matrix, $(\cdot)^H$ denotes the conjugate transpose, σ^2 denotes the noise variance, and $\mathbf{I}_M$ denotes the $M \times M$ identity matrix. The exact ML estimate $\hat{\phi}$ of ϕ is obtained by maximizing the likelihood function, which is equivalent to

$$\hat{\phi} = \arg \min_{\phi \in D_{\phi}} L_{\phi}(\phi), \tag{2}$$

where $D_{\phi} = \mathcal{R}^{(2M-1)\times 1}$, and

$$L_{\phi}(\phi) = \ln |\mathbf{R}(\phi)| + \operatorname{tr}\left[\mathbf{R}^{-1}(\phi)\tilde{\mathbf{R}}\right], \tag{3}$$

with $|\cdot|$ denoting the determinant, $\operatorname{tr}(\cdot)$ denoting the trace and $\tilde{\mathbf{R}}$ being the sample covariance matrix:

$$\tilde{\mathbf{R}} = \frac{1}{N}\sum_{n=1}^{N} \mathbf{y}(n)\mathbf{y}^H(n). \tag{4}$$

If we impose on $\mathbf{R}$ no structure except for Hermitian symmetry, then it is known that the ML estimate of $\mathbf{R}$ is given by $\tilde{\mathbf{R}}$ [1]; whereas if we observe the structure of $\mathbf{R}$ implied by the parameterization of $\mathbf{R}(\phi)$, the ML estimate of $\mathbf{R}$ is given by $\mathbf{R}(\hat{\phi})$. However, solving for the ML solution from (2) turns out to be very complicated because of the nonlinearity of the cost function. This limits the interest in using the exact ML structured covariance matrix estimate in practical applications.

3. DERIVATION OF THE AML ESTIMATOR

Let $\mathbf{r} = \operatorname{vec}(\mathbf{R}) \in \mathcal{C}^{M^2\times 1}$, where $\operatorname{vec}(\cdot)$ denotes the operation of stacking the columns of a matrix on top of one another, and let $\gamma \in \mathcal{R}^{M^2\times 1}$ denote the vector which is made from the real and imaginary parts of the elements of $\mathbf{R}$ above and on the main diagonal. Evidently there is an $M^2 \times M^2$ matrix $\mathbf{F}$ such that $\gamma = \mathbf{F}\mathbf{r}$. Furthermore, since the mapping from $\mathbf{r}$ to γ is one-to-one, $\mathbf{F}$ must be nonsingular. By invoking EXIP, we have the following result.

Theorem 1 *Let $\tilde{\gamma} = \mathbf{F}\operatorname{vec}(\tilde{\mathbf{R}})$. Then*

$$\tilde{\phi} = \arg \min_{\phi \in D_{\phi}} [\tilde{\gamma} - \gamma(\phi)]^T \, \Gamma^{-1} [\tilde{\gamma} - \gamma(\phi)] \tag{5}$$

is an asymptotically (in N) valid approximation of the ML estimate $\hat{\phi}$, where $\Gamma = \operatorname{cov}(\tilde{\gamma})$ or equivalently a consistent (in N) estimate of the covariance matrix of $\tilde{\gamma}$.

Proof: See [6].
Let $\tilde{\mathbf{r}} = \operatorname{vec}(\tilde{\mathbf{R}})$. Using the facts that $\tilde{\gamma} = \mathbf{F}\tilde{\mathbf{r}}$ and $\Gamma = \mathbf{F}\mathbf{C}\mathbf{F}^H$, where $\mathbf{C} = \operatorname{cov}(\tilde{\mathbf{r}})$, we can readily check that (5) is equivalent to

$$\tilde{\phi} = \arg \min_{\phi \in D_{\phi}} [\tilde{\mathbf{r}} - \mathbf{r}(\phi)]^H \, \mathbf{C}^{-1} [\tilde{\mathbf{r}} - \mathbf{r}(\phi)]. \tag{6}$$

It turns out to be more convenient to work with (6) than with (5), since we thus avoid the transformation from $\tilde{\mathbf{r}}$ to $\tilde{\gamma}$. An expression for $\mathbf{C}$, needed in (6), is obtained as follows. Let $\mathbf{r}_m$ denote the m-th column of $\mathbf{R}$ and let $\tilde{\mathbf{r}}_m$ denote the m-th column of $\tilde{\mathbf{R}}$. We have $E\{\tilde{\mathbf{r}}_m\} = \mathbf{r}_m$, and

$$E\left\{\tilde{\mathbf{r}}_{m_1}\tilde{\mathbf{r}}_{m_2}^H\right\} = \mathbf{r}_{m_1}\mathbf{r}_{m_2}^H + R_{m_2 m_1}\mathbf{R}/N, \tag{7}$$

where $(\cdot)^*$ denotes the complex conjugate and $R_{m_2 m_1}$ denotes the $m_2 m_1$-th element of $\mathbf{R}$. Hence

$$\mathbf{C} \triangleq E\left\{(\tilde{\mathbf{r}} - \mathbf{r})(\tilde{\mathbf{r}} - \mathbf{r})^H\right\} = (\mathbf{R}^T \otimes \mathbf{R})/N, \tag{8}$$

where $\otimes$ denotes the matrix Kronecker product. Using the natural and consistent (in N) estimate $\tilde{\mathbf{C}} = (\tilde{\mathbf{R}}^T \otimes \tilde{\mathbf{R}})/N$ for $\mathbf{C}$ in (6) leads to

$$\tilde{\phi} = \arg \min_{\phi \in D_{\phi}} [\tilde{\mathbf{r}} - \mathbf{r}(\phi)]^H (\tilde{\mathbf{R}}^{-T} \otimes \tilde{\mathbf{R}}^{-1}) [\tilde{\mathbf{r}} - \mathbf{r}(\phi)]. \tag{9}$$

Consider next the function $\mathbf{r}(\phi)$ for the case that $\mathbf{R}(\phi)$ is Hermitian Toeplitz. Let

$$\mathbf{Q}_m = \begin{bmatrix} \mathbf{0}_{M-m,m} & \mathbf{I}_{M-m} \\ \mathbf{0}_{m,M-m} & \mathbf{0}_{m,m} \end{bmatrix}, \quad m = 1, 2, \ldots, M-1, \tag{10}$$

where $\mathbf{0}_{r,s}$ denotes the $r \times s$ matrix with zero elements. Then

$$\begin{aligned} \mathbf{R} &\triangleq \begin{bmatrix} \rho_0 & \rho_1 & \cdots & \rho_{M-1} \\ \rho_1^* & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho_1 \\ \rho_{M-1}^* & \cdots & \rho_1^* & \rho_0 \end{bmatrix} \\ &= \rho_0 \mathbf{I}_M + \sum_{m=1}^{M-1} (\rho_m \mathbf{Q}_m + \rho_m^* \mathbf{Q}_m^T). \end{aligned} \tag{11}$$

Let

$$\begin{aligned} \boldsymbol{\Sigma} = \big[& \operatorname{vec}(\mathbf{I}_M) \quad \operatorname{vec}(\mathbf{Q}_1) \quad \operatorname{vec}(\mathbf{Q}_1^T) \quad \ldots \\ & \operatorname{vec}(\mathbf{Q}_{M-1}) \quad \operatorname{vec}(\mathbf{Q}_{M-1}^T) \big], \end{aligned} \tag{12}$$

and

$$\xi = \begin{bmatrix} \rho_0 & \rho_1 & \rho_1^* & \ldots & \rho_{M-1} & \rho_{M-1}^* \end{bmatrix}^T. \tag{13}$$

It follows from (11) that

$$\mathbf{r}(\phi) = \boldsymbol{\Sigma}\xi = \boldsymbol{\Sigma} \underbrace{\begin{bmatrix} 1 & & & & & \\ & 1 & i & & \mathbf{0} & \\ & 1 & -i & & & \\ & & & \ddots & & \\ & \mathbf{0} & & & 1 & i \\ & & & & 1 & -i \end{bmatrix}}_{\boldsymbol{\Omega}} \phi \triangleq \boldsymbol{\Psi}\phi, \tag{14}$$

where $\boldsymbol{\Psi} = \boldsymbol{\Sigma}\boldsymbol{\Omega}$, and

$$\phi = [\, \rho_0 \;\; \operatorname{Re}(\rho_1) \;\; Im(\rho_1) \;\; \ldots \;\; \operatorname{Re}(\rho_{M-1}) \;\; \operatorname{Im}(\rho_{M-1}) \,]^T. \tag{15}$$

Using (14) in (9) and minimizing the so-obtained quadratic function yields the following (asymptotic) ML estimate of ϕ:

$$\tilde{\phi} = \left[\operatorname{Re}\left(\boldsymbol{\Psi}^H \tilde{\mathbf{C}}^{-1} \boldsymbol{\Psi}\right)\right]^{-1} \left[\operatorname{Re}\left(\boldsymbol{\Psi}^H \tilde{\mathbf{C}}^{-1} \tilde{\mathbf{r}}\right)\right]. \tag{16}$$

To see this, let $\bar{\mathbf{r}} = \tilde{\mathbf{C}}^{-1/2}\tilde{\mathbf{r}}$, and $\bar{\mathbf{\Psi}} = \tilde{\mathbf{C}}^{-1/2}\tilde{\mathbf{\Psi}}$, where $\tilde{\mathbf{C}}^{-1/2}$ denotes the Hermitian square root of $\tilde{\mathbf{C}}$. Then

$$\begin{aligned} & [\tilde{\mathbf{r}} - \mathbf{r}(\phi)]^H \tilde{\mathbf{C}}^{-1}[\tilde{\mathbf{r}} - \mathbf{r}(\phi)] = \|\mathbf{r} - \bar{\mathbf{\Psi}}\phi\|^2 \\ = \ & \phi^T \bar{\mathbf{\Psi}}^H \bar{\mathbf{\Psi}}\phi - \phi^T \bar{\mathbf{\Psi}}^H \bar{\mathbf{r}} - \bar{\mathbf{r}}^H \bar{\mathbf{\Psi}}\phi + \bar{\mathbf{r}}^H \bar{\mathbf{r}} \\ = \ & \phi^T [\mathrm{Re}(\bar{\mathbf{\Psi}}^H \bar{\mathbf{\Psi}})]\phi - 2\phi^T \mathrm{Re}(\bar{\mathbf{\Psi}}^H \bar{\mathbf{r}}) + \bar{\mathbf{r}}^H \bar{\mathbf{r}}. \quad (17) \end{aligned}$$

Equation (16) immediately follows from (17). Next we note that the $\mathrm{Re}(\cdot)$ in (16) can be dropped since both $\mathbf{\Psi}^H \tilde{\mathbf{C}}^{-1}\mathbf{\Psi}$ and $\mathbf{\Psi}^H \tilde{\mathbf{C}}^{-1}\tilde{\mathbf{r}}$ can be shown to be real-valued due to the special structure of $\mathbf{\Psi}$. With this observation, we obtain the final formula of AML for estimating a Hermitian Toeplitz covariance matrix:

$$\tilde{\phi} = \left(\mathbf{\Psi}^H \tilde{\mathbf{C}}^{-1}\mathbf{\Psi}\right)^{-1} \left(\mathbf{\Psi}^H \tilde{\mathbf{C}}^{-1}\tilde{\mathbf{r}}\right). \quad (18)$$

The sparse structures of $\mathbf{\Sigma}$ and $\mathbf{\Omega}$ should of course be exploited for the AML implementation. For example, we can first compute $\tilde{\mathbf{C}}^{-1}\tilde{\mathbf{r}}$ and $\tilde{\mathbf{C}}^{-1}\mathbf{\Sigma}$ as follows:

$$\tilde{\mathbf{C}}^{-1}\tilde{\mathbf{r}} = (\tilde{\mathbf{R}}^{-T} \otimes \tilde{\mathbf{R}}^{-1})\mathrm{vec}(\tilde{\mathbf{R}}) = \mathrm{vec}(\tilde{\mathbf{R}}^{-1}), \quad (19)$$

and

$$\begin{aligned} \tilde{\mathbf{C}}^{-1}\mathbf{\Sigma} = \big[\ & \mathrm{vec}(\tilde{\mathbf{R}}^{-1}\mathbf{I}_M \tilde{\mathbf{R}}^{-1}) \quad \mathrm{vec}(\tilde{\mathbf{R}}^{-1}\mathbf{Q}_1 \tilde{\mathbf{R}}^{-1}) \\ & \mathrm{vec}[(\tilde{\mathbf{R}}^{-1}\mathbf{Q}_1 \tilde{\mathbf{R}}^{-1})^H] \quad \ldots \quad \mathrm{vec}(\tilde{\mathbf{R}}^{-1}\mathbf{Q}_{M-1} \tilde{\mathbf{R}}^{-1}) \\ & \mathrm{vec}[(\tilde{\mathbf{R}}^{-1}\mathbf{Q}_{M-1} \tilde{\mathbf{R}}^{-1})^H] \ \big]. \quad (20) \end{aligned}$$

Then calculating either $\mathbf{\Psi}^H \tilde{\mathbf{C}}^{-1}\mathbf{\Psi} = \mathbf{\Omega}^H \mathbf{\Sigma}^T \tilde{\mathbf{C}}^{-1}\mathbf{\Sigma}\mathbf{\Omega}$ or $\mathbf{\Psi}^H \tilde{\mathbf{C}}^{-1}\tilde{\mathbf{r}} = \mathbf{\Omega}^H \mathbf{\Sigma}^T \tilde{\mathbf{C}}^{-1}\tilde{\mathbf{r}}$ requires only a few additions.

Remark 1: If we relax the Gaussian assumption, then the estimate given in (2) is no longer the ML estimate. In such a case, it seems that the use of the "covariance matching" criterion (6) (or (5)) makes more sense than using (2). *Remark 2:* $\mathbf{R}(\tilde{\phi})$ as given by (18) is not guaranteed to be positive semidefinite. However, this may occur only if $\mathbf{R}$ is close to singular and N is relatively small. If $N \gg 1$, then by the consistency of $\tilde{\phi}$, the matrix $\mathbf{R}(\tilde{\phi})$ must be positive semidefinite. Our experimental experience suggests that for a number of data samples as small as, for example, $N = 15$, the estimated covariance matrix is always observed positive semidefinite, even when $\mathbf{R}$ is nearly singular.

4. NUMERICAL EXAMPLES

To illustrate the impact of using structured covariance matrix estimates on angle estimation, consider the problem of estimating the arrival angles $\theta_1 = 0°$ and $\theta_2 = 5°$ of two uncorrelated signals with the same power impinging on a ULA of $M = 5$ sensors separated by half wavelength. We compare the performances of the standard MUSIC, ITAM-MUSIC, AML-MUSIC as well as a WSF technique recently proposed in [7]. Specifically, we use the root-MUSIC algorithm for the first three MUSIC-related methods. The WSF algorithm was shown in [7] to be a large sample efficient method for estimating the arrival angles of uncorrelated signals. Note that WSF can be viewed as a structured covariance matrix estimator by plugging in the signal and noise parameter estimates. Define the SNR for the k-th incoming signal as $\mathrm{SNR}_k = 10\log_{10}(s_k/\sigma^2)$, where s_k denotes the variance of the k-th signal. Figure 1 shows the

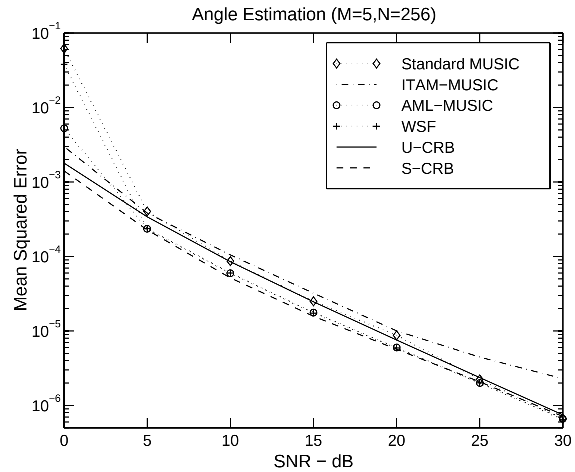

Figure 1. MSE's of the estimates of θ_1 and the corresponding U-CRB and S-CRB versus SNR when $N = 256$ and $M = 5$.

mean-squared errors (MSE's) of the corresponding angle estimates of θ_1, and the corresponding U-CRB and S-CRB, versus SNR. The MSE's are based on 200 independent trials. We note that:

- The standard MUSIC asymptotically (in SNR) achieves the U-CRB, which is a well-known fact;
- AML-MUSIC and WSF asymptotically achieve the S-CRB, with the former having a lower threshold SNR than the latter;
- The U-CRB asymptotically approaches the S-CRB;
- ITAM-MUSIC never attains the S-CRB and performs worse than the standard MUSIC when the SNR increases.

The ITAM estimator was originally proposed as an algorithm which can be used to enhance the performance of such algorithms as MUSIC and ESPRIT when the SNR is relatively low [4]. As indicated by Figure 1(b), ITAM's performance is indeed quite good at low SNR. However, ITAM is not an optimal method and there is no surprise that ITAM-MUSIC never achieves the S-CRB. On the other hand, the inconsistency (in SNR) of ITAM appears surprising at first sight. To explain it briefly, note that as the SNR goes to infinity, we have $\lim_{\sigma^2 \to 0} \tilde{\mathbf{R}} = \mathbf{A}(\theta)\tilde{\mathbf{S}}\mathbf{A}^H(\theta)$, where $\tilde{\mathbf{S}}$ is the sample signal covariance matrix which is *not* diagonal for finite N. In spite of the fact that $\tilde{\mathbf{R}}$ is not Toeplitz in this case, the *signal* and *noise subspaces* can be obtained *exactly* from $\tilde{\mathbf{R}}$ when the SNR goes to infinity. For a subspace-based algorithm like root-MUSIC, perfect angle estimates can be obtained if the exact subspace is available. However, ITAM attempts to find a Toeplitz matrix that is as close to $\tilde{\mathbf{R}}$ as possible and no efforts are made to ensure appropriate subspace approximation. As a result, the subspaces of $\tilde{\mathbf{R}}$ are distorted by the sequences of approximation introduced by ITAM, and ITAM-MUSIC is hence inconsistent in SNR. It is interesting to note that, even though AML assumes a Toeplitz structure as ITAM, it does not suffer from the inconsistency problem suffered by ITAM. In the appendix, we show that using the AML criterion (9) in array processing when the SNR is high is equivalent to seeking a Toeplitz matrix that is closest to the range space of $\mathbf{A}(\theta)$. Consequently, the AML covariance matrix estimate provides consistent (in SNR) subspace estimates and AML-MUSIC in turn yields consistent angle estimates at high SNR.

AML-MUSIC and WSF in general perform quite similarly for most cases except, as we have found, that in some difficult scenarios, such as when the SNR is relatively low or when the signals are close to each other, the former tends to perform better than the latter. In addition, AML is a general covariance matrix estimator and should of course not be limited to the application of angle estimation; whereas WSF is essentially an angle estimator. The following example indicates that using AML in angle estimation involves very modest additional computations. Define η as the ratio of flops needed by ITAM-MUSIC, AML-MUSIC or WSF to that by MUSIC. Figure 2 shows the curves of η versus M when SNR=10 dB and $N = 256$. Clearly AML-MUSIC needs the least computations, which is due to the computational efficiency of AML, while WSF becomes computationally much more involved than the others when M increases.

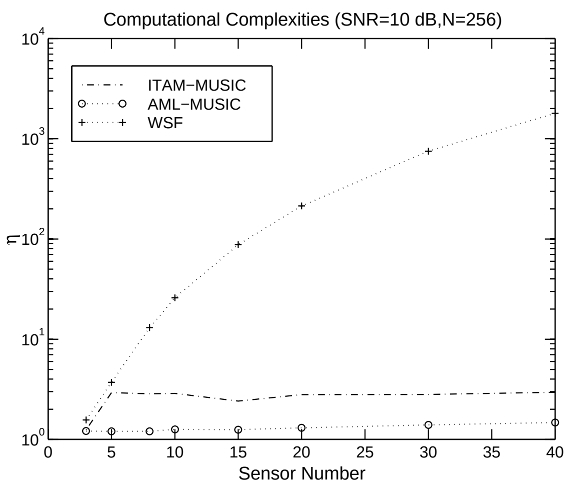

Figure 2. The flops ratio, η, versus M when SNR=10 dB and $N = 256$.

5. CONCLUSIONS

We have presented an asymptotic ML method, referred to as AML, for structured covariance matrix estimation. A closed-form formula for Toeplitz covariance matrix estimation has been derived. We have shown that using the AML covariance matrix estimate improves the angle estimation accuracy of MUSIC by achieving the relevant S-CRB. Finally, we remark on the fact that even though we have concentrated on Hermitian Toeplitz matrix estimation in this paper, it is straightforward to extend the proposed technique to estimate any other matrices that have a linear structure.

APPENDIX ANALYSIS OF AML AT HIGH SNR

In this appendix we show that the AML estimate of the spatial covariance matrix $\mathbf{R}$ provides accurate estimate of the signal subspace of $\mathbf{R}$ at high SNR. Let the eigendecomposition of $\tilde{\mathbf{R}}$ be

$$\tilde{\mathbf{R}} = \tilde{\mathbf{E}}_s\tilde{\mathbf{\Lambda}}_s\tilde{\mathbf{E}}_s^H + \tilde{\mathbf{E}}_n\tilde{\mathbf{\Lambda}}_n\tilde{\mathbf{E}}_n^H, \tag{21}$$

where $\tilde{\mathbf{\Lambda}}_s$ is the diagonal matrix containing the K largest eigenvalues with the columns of $\tilde{\mathbf{E}}_s$ being the associated eigenvectors, and $\tilde{\mathbf{\Lambda}}_n$ is the diagonal matrix containing the remaining eigenvalues with the columns of $\mathbf{E}_n$ being the corresponding eigenvectors. Since, for sufficiently small σ^2, we have $\tilde{\mathbf{\Lambda}}_n = O(\sigma^2)$, it follows that

$$\tilde{\mathbf{R}}^{-1} = \tilde{\mathbf{E}}_s\tilde{\mathbf{\Lambda}}_s^{-1}\tilde{\mathbf{E}}_s^H + \tilde{\mathbf{E}}_n\tilde{\mathbf{\Lambda}}_n^{-1}\tilde{\mathbf{E}}_n^H \approx \tilde{\mathbf{E}}_n\tilde{\mathbf{\Lambda}}_n^{-1}\tilde{\mathbf{E}}_n^H. \tag{22}$$

Hence we can rewrite the cost function in (9) at high SNR as

$$\begin{aligned} & \left[\mathrm{vec}(\tilde{\mathbf{R}}^T - \mathbf{R}^T(\phi))\right]^T (\tilde{\mathbf{R}}^{-T} \otimes \tilde{\mathbf{R}}^{-1}) \left[\mathrm{vec}(\tilde{\mathbf{R}} - \mathbf{R}(\phi))\right] \\ = \ & \mathrm{tr}\left[(\tilde{\mathbf{R}} - \mathbf{R}(\phi))\tilde{\mathbf{R}}^{-1}(\tilde{\mathbf{R}} - \mathbf{R}(\phi))\tilde{\mathbf{R}}^{-1}\right] \\ \approx \ & \mathrm{tr}\left[\left(\mathbf{I}_M - \mathbf{R}(\phi)\tilde{\mathbf{E}}_n\tilde{\mathbf{\Lambda}}_n^{-1}\tilde{\mathbf{E}}_n^H\right)^2\right]. \end{aligned} \tag{23}$$

The second term in (23) is of the order $O(\sigma^{-2})$, and hence is the dominant one. To minimize the criterion function in (23), the AML estimation $\mathbf{R}(\tilde{\phi})$ of $\mathbf{R}$ must minimize this term (as $\sigma^2 \to 0$). However this is only possible if the range space of $\mathbf{R}(\tilde{\phi})$ is close to that of $\tilde{\mathbf{E}}_s$, which, in turn, approaches the range space of $\mathbf{A}(\theta)$ (the so-called signal subspace) as $\sigma^2 \to 0$.

REFERENCES

[1] J. P. Burg, D. G. Luenberger, and D. L. Wenger, "Estimation of structured covariance matrices," *Proceedings of the IEEE*, vol. 70, pp. 963–974, September 1982.

[2] D. R. Fuhrmann, "Application of Toeplitz covariance estimation to adaptive beamforming and detection," *IEEE Transactions on Signal Processing*, vol. 39, pp. 2194–2198, October 1991.

[3] D. B. Williams and D. H. Johnson, "Robust estimation of structured covariance matrices," *IEEE Transactions on Signal Processing*, vol. 41, pp. 2891–2906, September 1993.

[4] J. A. Cadzow, "Signal enhancement – a composite property mapping algorithm," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, pp. 49–62, January 1988.

[5] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, pp. 276–280, March 1985.

[6] P. Stoica and T. Söderström, "On reparameterization of loss functions used in estimation and the invariance principle," *Signal Processing*, vol. 17, pp. 383–387, August 1989.

[7] M. Jansson, B. Göransson, and B. Ottersten, "Analysis of a subspace-based spatial frequency estimator," *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, Germany, April 1997.