AMPLITUDE ESTIMATION WITH APPLICATION TO SYSTEM IDENTIFICATION *

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

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

ABSTRACT

We investigate herein the problem of amplitude estimation of sinusoidal signals from observations corrupted by colored noise. A relatively large number of amplitude estimators are described which encompass Least Squares (LS) and Weighted Least Squares (WLS) methods. Additionally, filterbank approaches, which are widely used for spectral analysis, are extended to amplitude estimation. Specifically, we consider the recently introduced MAtched-FIlterbank (MAFI) approach and show that, by appropriately designing the prefilters, the MAFI approach includes the WLS approach. The amplitude estimation techniques discussed in this paper do not model the noise, and yet they are all asymptotically statistically efficient. It is their different finite-sample properties that are of particular interest to this study. Numerical examples are provided to illustrate the differences among the various estimators. Though amplitude estimation applications are numerous, we focus on system identification using sinusoidal probing signals.

1. INTRODUCTION

Consider the noise-corrupted observations of K complex-valued sinusoids

$$x(n) = \sum_{k=1}^{K} \alpha_k e^{j\omega_k n} + v(n), \; n = 0, 1, \dots N-1, \qquad (1)$$

where α_k denotes the complex amplitude of the kth sinusoid having frequency ω_k and $v(n)$ is the observation noise which is complex-valued and assumed to be stationary (and possibly colored) with zero-mean and finite unknown Power Spectral Density (PSD) $\phi(\omega)$. We assume that $\{\omega_k\}_{k=1}^{K}$ are known, with $\omega_k \neq \omega_l$, for $k \neq l$. The problem of interest is to estimate $\{\alpha_k\}_{k=1}^{K}$ from the observations $\{x(n)\}_{n=0}^{N-1}$.

We describe three general classes of amplitude estimators, namely the LS, WLS, and MAFI approaches. The amplitude estimators under discussion can be further categorized depending on whether they estimate one amplitude at a time or all amplitudes simultaneously. The various amplitude estimators are summarized in Section 5. We also discuss a system identification application using sinusoidal probing signals. We show that, by using proper amplitude estimators, we can avoid the iterative search required by the standard system identification routines, such as the Output Error Method (OEM), and achieve very good performance at a usually reduced computational load.

This paper uses the following notation to distinguish among the various amplitude estimators. For instance, LSE$(1,0,1)$ denotes the LS estimator that does not split the data, uses no prefiltering, and estimates one amplitude at a time. Likewise, MAFI(L,K,K) denotes the MAFI estimator that splits the data into L subvectors, utilizes K prefilters, and estimates K amplitudes simultaneously. The remaining amplitude estimators are similarly designated.

2. LS AMPLITUDE ESTIMATORS

2.1. LSE$(1,0,K)$

Let us write the data sequence in the following form

$$\mathbf{x} = \tilde{\mathbf{A}}\alpha + \mathbf{v}, \qquad (2)$$

where $\tilde{\mathbf{A}}$ is an $N \times K$ Vandermonde matrix, $\mathbf{x} = [x(0) \; \dots \; x(N-1)]^T$, $\alpha = [\alpha_1 \; \dots \; \alpha_K]^T$, $\mathbf{v} = [v(0) \; \dots \; v(N-1)]^T$, and where $(\cdot)^T$ denotes the transpose. The LS estimate of α is

$$\hat{\alpha} = (\tilde{\mathbf{A}}^H \tilde{\mathbf{A}})^{-1} \tilde{\mathbf{A}}^H \mathbf{x}, \qquad (3)$$

where $(\cdot)^H$ denotes the conjugate transpose. Note that $\mathbf{v}$ is not modeled. Even so, (3) is asymptotically efficient [1].

2.2. LSE$(1,0,1)$

Since the observation noise $v(n)$ is not modeled, one way to reduce the computational burden quite a bit is to include $K-1$ sinusoids in the noise term and estimate one amplitude at a time. The LSE$(1,0,1)$ is easily derived as

$$\hat{\alpha}_k = \frac{1}{N} \sum_{n=0}^{N-1} x(n) e^{-j\omega_k n}, \; k = 1, 2, \dots, K, \qquad (4)$$

which is the Discrete Fourier Transform (DFT) of $\{x(n)\}_{n=0}^{N-1}$. LSE$(1,0,1)$ is biased but asymptotically unbiased [1]. Moreover, (4) is asymptotically efficient [1]. In finite samples, (4) may be better or worse than (3) depending on the characteristics of the scenario under study [1].

3. WLS AMPLITUDE ESTIMATORS

3.1. WLSE$(L,0,K)$

We define the subvectors $\mathbf{y}(l) = [x(l) \; x(l+1) \; \dots \; x(l+M-1)]^T, \; l = 0, 1, \dots, L-1,$, where $L \triangleq N - M + 1$. We have

$$\mathbf{y}(l) = \mathbf{A}\mathbf{s}(l) + \epsilon(l), \qquad (5)$$

*THIS WORK WAS SUPPORTED IN PART BY THE SENIOR INDIVIDUAL GRANT PROGRAM OF THE SWEDISH FOUNDATION FOR STRATEGIC RESEARCH, THE NATIONAL SCIENCE FOUNDATION GRANT MIP-9457388, AND THE OFFICE OF NAVAL RESEARCH GRANT N00014-96-0817.

where $\mathbf{A}$ is an $M \times K$ Vandermonde matrix, $\mathbf{s}(l) = [\alpha_1 e^{j\omega_1 l} \ \ldots \ \alpha_K e^{j\omega_K l}]^T$, and $\epsilon(l) = [v(l) \ \ldots \ v(l+M-1)]^T$. Alternatively, we can rewrite (5) as

$$\mathbf{y}(l) = \mathbf{A}_l \alpha + \epsilon(l), \tag{6}$$

where $\mathbf{A}_l \triangleq \mathbf{A}\,\mathrm{diag}\{e^{j\omega_1 l} \ \ldots \ e^{j\omega_K l}\} \triangleq \mathbf{A}\mathbf{D}_l$. We will use (5) mostly for analysis and (6) for estimation.

The WLS (Markov-like) estimate of α in (6) is given by

$$\hat{\alpha} = \left[\sum_{l=0}^{L-1} \mathbf{A}_l^H \hat{\mathbf{Q}}^{-1} \mathbf{A}_l\right]^{-1} \left[\sum_{l=0}^{L-1} \mathbf{A}_l^H \hat{\mathbf{Q}}^{-1} \mathbf{y}(l)\right], \tag{7}$$

where $\hat{\mathbf{Q}}$ is an estimate of $\mathbf{Q} = E\{\epsilon(l)\epsilon^H(l)\}$. To estimate $\mathbf{Q}$, we may proceed as follows. Let $\hat{\mathbf{R}} = \frac{1}{L}\sum_{l=0}^{L-1} \mathbf{y}(l)\mathbf{y}^H(l)$. One can verify that as $L \to \infty$, $\hat{\mathbf{R}}$ goes to $\mathbf{R} = \mathbf{A}\mathbf{P}\mathbf{A}^H + \mathbf{Q}$, where $\mathbf{P} = \mathrm{diag}\{|\alpha_1|^2 \ \ldots \ |\alpha_K|^2\}$. Hence, one way to estimate $\mathbf{Q}$ is as

$$\hat{\mathbf{Q}} = \hat{\mathbf{R}} - \mathbf{A}\hat{\mathbf{P}}\mathbf{A}^H, \tag{8}$$

where $\hat{\mathbf{P}}$ is made from some initial estimates of $\{\alpha_k\}_{k=1}^K$. In the following we try to circumvent the need for initial amplitude estimates in two different ways.

First, we show a way to simplify the WLSE$(L, 0, K)$ that uses (7) with (8). Note that [1] $\hat{\mathbf{Q}}^{-1}\mathbf{A}_l = \hat{\mathbf{R}}^{-1}\mathbf{A}\boldsymbol{\Gamma}\mathbf{D}_l \approx \hat{\mathbf{R}}^{-1}\mathbf{A}\mathbf{D}_l\boldsymbol{\Gamma} = \hat{\mathbf{R}}^{-1}\mathbf{A}_l\boldsymbol{\Gamma}$, where $\boldsymbol{\Gamma} \triangleq \hat{\mathbf{P}}\mathbf{A}^H\hat{\mathbf{Q}}^{-1}\mathbf{A} + \mathbf{I}_K$ with $\mathbf{I}_K$ being the $K \times K$ identity matrix. Hence, (7) reduces to

$$\hat{\alpha} \approx \left[\sum_{l=0}^{L-1} \mathbf{A}_l^H \hat{\mathbf{R}}^{-1} \mathbf{A}_l\right]^{-1} \left[\sum_{l=0}^{L-1} \mathbf{A}_l^H \hat{\mathbf{R}}^{-1} \mathbf{y}(l)\right]. \tag{9}$$

The amplitude estimator in (9) can be interpreted as an extension of the Capon method in [2] to multiple sinusoids.

A different estimate of $\mathbf{Q}$ can be obtained as described next. Observe that $\mathbf{A}\mathbf{P}\mathbf{A}^H = \sum_{k=1}^K [\alpha_k \mathbf{a}(\omega_k)][\alpha_k \mathbf{a}(\omega_k)]^H \triangleq \sum_{k=1}^K \beta_k \beta_k^H$, where $\mathbf{a}(\omega) = [1 \ e^{j\omega} \ \ldots \ e^{j(M-1)\omega}]^T$. Thus,

$$\mathbf{y}(l) = \sum_{k=1}^K \beta_k e^{j\omega_k l} + \epsilon(l). \tag{10}$$

From (10), we can estimate β_k one at a time via LS as

$$\hat{\beta}_k = \frac{1}{L}\sum_{l=0}^{L-1} \mathbf{y}(l) e^{-j\omega_k l} \triangleq \mathbf{g}(\omega_k). \tag{11}$$

The use of (11) in (8) leads to

$$\hat{\mathbf{Q}} = \hat{\mathbf{R}} - \sum_{k=1}^K \mathbf{g}(\omega_k)\mathbf{g}^H(\omega_k). \tag{12}$$

The WLSE$(L, 0, K)$ that uses (7) with (12) is an extension of APES in [3] to multiple sinusoids with known frequencies.

Remark: We note that $\epsilon(k)$ and $\epsilon(l)$ in (6) are correlated (for $k \neq l$), which implies that (7) is suboptimal. Yet, the WLS methods are likely to outperform the LS methods because the latter completely ignore the correlation in $v(n)$.

3.2. WLSE$(L, 0, 1)$

It is straightforward to show that the WLSE$(L, 0, 1)$ that uses (7) with (8) is

$$\hat{\alpha}_k = \frac{\mathbf{a}^H(\omega_k)\hat{\mathbf{R}}^{-1}\mathbf{g}(\omega_k)}{\mathbf{a}^H(\omega_k)\hat{\mathbf{R}}^{-1}\mathbf{a}(\omega_k)}, \tag{13}$$

whereas the WLSE$(L, 0, 1)$ that uses (7) with (12) is

$$\hat{\alpha}_k = \frac{\mathbf{a}^H(\omega_k)[\hat{\mathbf{R}} - \mathbf{g}(\omega_k)\mathbf{g}^H(\omega_k)]^{-1}\mathbf{g}(\omega_k)}{\mathbf{a}^H(\omega_k)[\hat{\mathbf{R}} - \mathbf{g}(\omega_k)\mathbf{g}^H(\omega_k)]^{-1}\mathbf{a}(\omega_k)}. \tag{14}$$

Note that, unlike (9), the equation (13) is *exactly* equivalent to using (7) with (8). Equations (13) and (14) are recognized to have the same form as the Capon [2] and, respectively, the APES [3] spectral estimators. The two estimators were derived in [4] [5] by a different approach, namely the MAFI approach, which we will consider in a generalized form in the next section. It is interesting that the above two estimators, while both *asymptotically efficient*, have quite *different finite-sample properties*. Specifically, it was shown in [4] [5] that (13) is biased downward, whereas (14) is unbiased (within a second-order approximation) and in general has a better performance than the former.

4. MAFI AMPLITUDE ESTIMATORS

Let $\mathbf{H}^H \in \mathcal{C}^{\bar{K} \times M}$ be a matrix each row of which is a Finite Impulse Response (FIR) filter (for some $1 \leq \bar{K} \leq M$). The MAFI idea can be explained as follows: a) design $\mathbf{H}^H$ so that, when applied to $\{\mathbf{y}(l)\}$, it maximizes the SNR at the $\bar{K}$ filter outputs and b) estimate the amplitudes from the filtered data by, e.g., the LS or WLS technique. Mathematically, $\mathbf{H}$ can be obtained as follows:

$$\mathbf{H} = \arg\max_{\mathbf{H}} \underbrace{\mathrm{tr}\left[(\mathbf{H}^H\hat{\mathbf{Q}}\mathbf{H})^{-1}\mathbf{H}^H(\mathbf{A}\hat{\mathbf{P}}\mathbf{A}^H)\mathbf{H}\right]}_{\text{“Generalized SNR”}}, \tag{15}$$

where $\mathbf{H}$ is constrained in a way that is specified later (in particular, to guarantee that $\mathbf{H}$ is finite), and $\mathrm{tr}(\cdot)$ denotes the trace of a matrix.

It was shown in [1] that the solution $\mathbf{H}$ is not unique. A simple solution with $\bar{K} = K$ is given by [1]

$$\mathbf{H} = \hat{\mathbf{Q}}^{-1}\mathbf{A}(\mathbf{A}^H\hat{\mathbf{Q}}^{-1}\mathbf{A})^{-1}. \tag{16}$$

The above $\mathbf{H}$ also satisfies the constraint $\mathbf{H}^H\mathbf{A} = \mathbf{I}_K$, which says that each (row) filter in $\mathbf{H}^H$ passes one sinusoid undistorted, and completely annihilates the others.

From (6), the filtered data corresponding to (16) is

$$\mathbf{z}(l) \triangleq \mathbf{H}^H\mathbf{y}(l) = \mathbf{D}_l\alpha + \mathbf{H}^H\epsilon(l) \triangleq \mathbf{D}_l\alpha + \nu(l). \tag{17}$$

The covariance matrix of $\nu(l)$ can be estimated as $\mathbf{H}^H\hat{\mathbf{Q}}\mathbf{H} = (\mathbf{A}^H\hat{\mathbf{Q}}^{-1}\mathbf{A})^{-1}$. It follows that the WLS (Markov-like) estimate of α in (17) is

$$\hat{\alpha} = \left[\sum_{l=0}^{L-1} \mathbf{A}_l^H \hat{\mathbf{Q}}^{-1} \mathbf{A}_l\right]^{-1} \left[\sum_{l=0}^{L-1} \mathbf{A}_l^H \hat{\mathbf{Q}}^{-1} \mathbf{y}(l)\right], \tag{18}$$

which coincide with (7).

The MAFI interpretation of the WLS method makes an interesting connection between using the MAFI and WLS techniques for amplitude estimation. However, the MAFI approach is more general since the solution $\mathbf{H}$ is not unique [1]. We derive in the following a new MAFI amplitude estimator. Other interesting ones may exist and are yet to be discovered. Let $z_k(l)$ and $\nu_k(l)$ denote the k-th element of $\mathbf{z}(l)$ and, respectively, $\nu(l)$ in (17). Then

$$z_k(l) = \alpha_k e^{j\omega_k l} + \nu_k(l), \ k = 1, 2, \ldots, K. \tag{19}$$

By LS, the MAFI$(L, K, 1)$ estimate of the α_k is given by

$$\hat{\alpha}_k = \frac{1}{L}\sum_{l=0}^{L-1} z_k(l)e^{-j\omega_k l}. \quad (20)$$

Unlike (13) and (14) (which are also members of MAFI$(L, K, 1)$ [4]), (20) *does* require the knowledge of the number and frequencies of the sinusoids, which makes it behave more like a MAFI(L, K, K) estimator. In particular, it performs quite well when some sinusoids are closely spaced, as shown in Section 5.

5. NUMERICAL EXAMPLES

For notational simplicity, we adopt the following acronyms for the various amplitude estimators: **i**. LSE1: using (4); **ii**. LSEK: using (3); **iii**. Capon1: using (13); **iv**. APES1: using (14); **v**. CaponK: using (9); **vi**. APESK: using (7) along with (12); and **vii**. MAFI1: using (20) along with (12). The test data consist of three complex sinusoids corrupted by an AutoRegressive (AR) noise described by $v(n) = 0.99v(n-1) + e(n)$, where $e(n)$ is a complex white Gaussian noise with zero-mean and variance σ^2. The frequencies of the sinusoids are 0.1, 0.11, and 0.3 Hz, the amplitudes are $e^{j\pi/4}, e^{j\pi/3}$, and $e^{j\pi/4}$, and $N = 32$. We define the Signal-to-Noise Ratio (SNR) of the k-th sinusoid by $\text{SNR}_k = 10\log_{10} N|\alpha_k|^2/\phi(\omega_k)$. For those methods that depend on M, we choose $M = N/4 = 8$ [1]. Figure 1(a) shows the MSEs of the seven amplitude estimators for the sinusoid at 0.3 Hz, along with the corresponding Cramér-Rao Bound (CRB), as the SNR varies. Figure 1(b) shows the counterpart curves for the sinusoid at 0.1 Hz. A brief summary based on these results (also see [1]) is as follows. APES1 is recommended in applications where it is known a priori that no two sinusoids are closely spaced (see, e.g., the application discussed in the next section), or when the closely-spaced sinusoids are of no interest. The reason to prefer APES1 to APESK or MAFI1 in such cases is that the former is more flexible than the latter two since APES1 does not necessarily require the knowledge of the sinusoidal frequencies. In terms of computational cost, APES1 and MAFI1 are similar to one another and both are simpler than APESK. When it is desired to estimate closely spaced sinusoids in colored noise, however, MAFI1 may be preferred. In general, we do not recommend the use of Capon1 since it has a computational complexity similar to that of APES1 but is biased. Although we did notice that CaponK gives close-to-CRB performance at very low SNRs, in most cases of interest, other methods like APES1 or MAFI1 may be preferred. LSEK is statistically efficient and may be preferred when the observation noise is white; in cases where the white noise assumption is invalid, it is preferable to use APES1 or MAFI1. LSE1 gives comparatively rather poor estimation accuracy but is computationally quite simple. The performance differences stated so far occur only when N is relatively small. As N increases, all methods tend to the CRB, independent of the noise correlation. Hence, when N is sufficiently large, LSE1 should be preferred because of its computational simplicity.

6. SYSTEM IDENTIFICATION

Consider the linear discrete-time system described by [6]

$$x(n) = H(z^{-1})u(n) + v(n), \; n = 0, 1, \ldots, N-1, \quad (21)$$

where $u(n) = \sum_{k=1}^{K} \gamma_k e^{j\omega_k n}$, and

$$H(z^{-1}) = \frac{B(z^{-1})}{A(z^{-1})} = \frac{b_1 z^{-1} + \ldots + b_q z^{-q}}{1 + a_1 z^{-1} + \ldots + a_p z^{-p}}. \quad (22)$$

We assume that $K \geq p + q$. If p and q are unknown, K should be chosen sufficiently large. The problem is to estimate $\{a_i\}_{i=1}^{p}$ and $\{b_j\}_{j=1}^{q}$ from $\{x(n)\}_{n=0}^{N-1}$.

The commonly-used Output Error Method (OEM) minimizes the following criterion [6]

$$C_{\text{OEM}}(\mathbf{a}, \mathbf{b}) = \sum_{n=0}^{N-1} \left|x(n) - H(z^{-1})u(n)\right|^2, \quad (23)$$

where $\mathbf{a} = [a_1 \; \ldots \; a_p]^T$ and $\mathbf{b} = [b_1 \; \ldots \; b_q]^T$. Let $\alpha_k(\mathbf{a}, \mathbf{b}) = \gamma_k H(e^{j\omega_k})$. For sufficiently large N, the cost function $C_{\text{OEM}}(\mathbf{a}, \mathbf{b})$ is approximately equivalent to

$$C_1(\mathbf{a}, \mathbf{b}) = \sum_{n=0}^{N-1} \left|x(n) - \sum_{k=1}^{K} \alpha_k(\mathbf{a}, \mathbf{b})e^{j\omega_k n}\right|^2. \quad (24)$$

The method that we propose for estimating $\mathbf{a}$ and $\mathbf{b}$ is based on (24) and consists of two steps, as detailed next.

Step 1 Use an appropriate amplitude estimator, such as APES1, to obtain estimates $\{\hat{\alpha}_k\}_{k=1}^{K}$ of $\{\alpha_k\}_{k=1}^{K}$ from the measurements $\{x(n)\}_{n=0}^{N-1}$. The large-sample variance of the estimated amplitudes $\{\hat{\alpha}_k\}_{k=1}^{K}$ is proportional to $\{\phi(\omega_k)\}_{k=1}^{K}$ [1], which can be estimated from the residual $\hat{v}(n) = x(n) - \sum_{k=1}^{K} \hat{\alpha}_k e^{j\omega_k n}, \; n = 0, 1, \ldots, N-1$.

Step 2 Obtain estimates of $\{a_i, b_j\}$ by minimizing

$$C_2(\mathbf{a}, \mathbf{b}) = \sum_{k=1}^{K} \frac{1}{\hat{\phi}(\omega_k)} \left|\hat{\alpha}_k - \alpha_k(\mathbf{a}, \mathbf{b})\right|^2. \quad (25)$$

To do so we can use a host of methods provided that we have good initial estimates of $\mathbf{a}$ and $\mathbf{b}$. To that end, we assume that p and q are known. (Standard techniques for system order determination can be found in, e.g., [6].) We pick up the $p + q$ largest $\{\hat{\alpha}_k\}$ and define a criterion made from the corresponding terms of (25)

$$C_3(\mathbf{a}, \mathbf{b}) = \sum_{k=1}^{p+q} \frac{1}{\hat{\phi}(\omega_k)} \left|\hat{\alpha}_k - \alpha_k(\mathbf{a}, \mathbf{b})\right|^2, \quad (26)$$

where we have assumed that $\{\hat{\alpha}_k\}_{k=1}^{p+q}$ are the $p + q$ chosen amplitudes. The minimization of (26) is simple since one can set $\hat{\alpha}_k = \alpha_k(\mathbf{a}, \mathbf{b})$, which is equivalent to

$$\frac{\hat{\alpha}_k}{\gamma_k} A(e^{j\omega_k}) = B(e^{j\omega_k}), \; k = 1, 2, \ldots, p+q. \quad (27)$$

Equation (27) can be rewritten as a linear system of $p + q$ equations with $p + q$ unknowns $\{a_i, b_j\}$.

Remark: By the Extended Invariance Principle (EXIP) [7], the estimates of $\{a_i, b_j\}$ obtained by minimizing (25) achieve the CRB asymptotically, and hence they have a better asymptotic accuracy than the OEM estimates whenever $v(n)$ is colored. It also follows from this observation that in the case of $K = p + q$, the estimates obtained from (27) are asymptotically efficient. This latter result (of a somewhat limited interest since it requires $K = p + q$) was first proved in [8] in a much more complicated way.

Consider now an example. The system is defined by(22) with $A(z^{-1}) = 1 - 1.9109z^{-1} + 1.7251z^{-2} - 0.7033z^{-3} + 0.245z^{-4}$ and $B(z^{-1}) = z^{-1} + 1.0562z^{-2} + 0.61z^{-3} + 0.1912z^{-4} + 0.04z^{-5}$. The noise $v(n)$ is an

AR noise similar to the one used in Section 5 except that now $e(n)$ is a real-valued white Gaussian noise with zero-mean and $\sigma^2 = 0.01$. The probing signal is $u(n) = 2\cos(2\pi 0.05n) + 2\cos(2\pi 0.15n) + 2\cos(2\pi 0.25n) + 2\cos(2\pi 0.35n) + 2\cos(2\pi 0.45n)$. The probing signal and observation noise are real-valued since this is the usual case in practice. To reduce the number of graphs, we only show the averaged Root Mean Squared Error (RMSE) for the a-parameters $\text{RMSE}\{\hat{\mathbf{a}}\} = \frac{1}{p}\sum_{i=1}^{p} \text{RMSE}\{\hat{a}_i\}$, and similarly for the b-parameters. Figures 2(a) and 2(b) show the averaged RMSEs of the a-parameters and, respectively, the b-parameters obtained by using OEM and the proposed technique, as N increases. Figure 2(c) shows the required number of flops as N increases. For the proposed technique, we compute both the initial estimates given by solving (27), via LSEK, APES1 or MAFI1, and the minimizer of (25), obtained by a standard gradient-type nonlinear optimization routine. We use $M = 20$ for APES1 and MAFI1. As one can see, the initial system parameter estimates given by APES1 or MAFI1 are significantly better than those given by OEM, and yet the former two are computationally more efficient that the latter. The estimates obtained by minimizing (25) are only slightly better than the initial estimates obtained by APES1 or MAFI1, but at a significantly increased computational cost. Hence, minimizing (25) for refined estimation accuracy is not recommended.

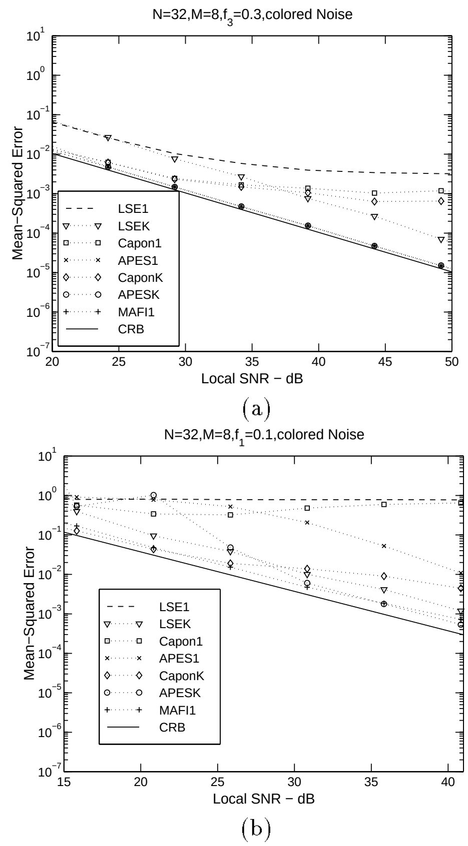

Figure 1. Empirical MSEs and the CRB versus local SNR. (a) α_3. (b) α_1.

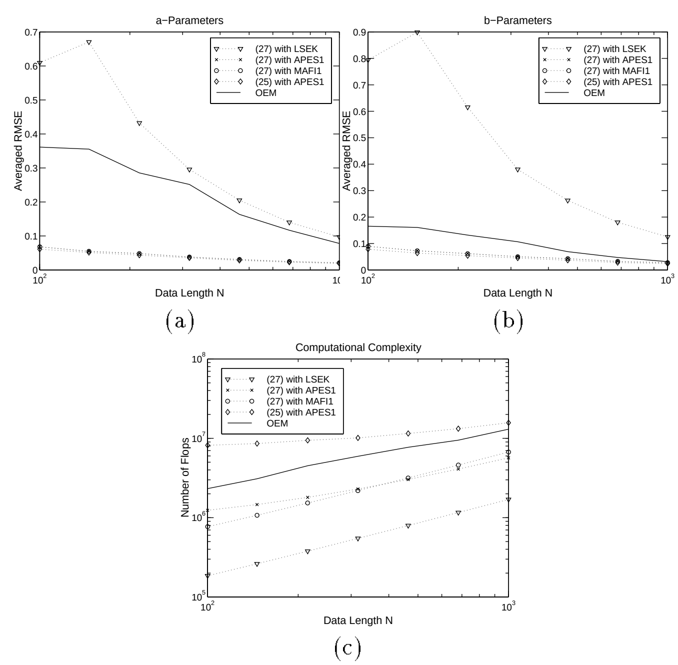

Figure 2. Averaged RMSEs and the number of flops versus N. (a) RMSE of a-parameters. (b) RMSE of b-parameters. (c) Number of flops.

REFERENCES

[1] P. Stoica, H. Li, and J. Li, "Amplitude estimation for sinusoidal signals with applications to system identification," submitted to *IEEE Transactions on Signal Processing*, August 1998.

[2] J. Capon, "High resolution frequency-wavenumber spectrum analysis," *Proceedings of the IEEE*, vol. 57, pp. 1408–1418, August 1969.

[3] J. Li and P. Stoica, "An adaptive filtering approach to spectral estimation and SAR imaging," *IEEE Transactions on Signal Processing*, vol. 44, pp. 1469–1484, June 1996.

[4] P. Stoica, A. Jakobsson, and J. Li, "Matched-filter bank interpretation of some spectral estimators," *Signal Processing*, vol. 66, pp. 45–59, April 1998.

[5] H. Li, J. Li, and P. Stoica, "Performance analysis of forward-backward matched-filterbank spectral estimators," *IEEE Transactions on Signal Processing*, vol. 46, pp. 1954–1966, July 1998.

[6] T. Söderström and P. Stoica, *System Identification*. London: Prentice Hall, 1989.

[7] 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.

[8] P. V. Kabaila, "On output-error methods for system identification," *IEEE Transactions on Automatic Control*, vol. 28, pp. 12–23, January 1983.