Two-Dimensional System Identification Using Amplitude Estimation

Hongbin Li, *Member, IEEE*, Wei Sun, Petre Stoica, *Fellow, IEEE*, and Jian Li, *Senior Member, IEEE*

***Abstract*—In [1], we introduced an amplitude estimation based scheme for one-dimensional (1-D) system identification that overcomes several drawbacks (e.g., computational complexity, local convergence, and statistical inefficiency when spectrally colored noise is present) suffered by the conventional output error method (OEM). Along the same line, we herein propose a two-dimensional (2-D) system identification scheme that makes use of 2-D amplitude estimation. In particular, we consider the recently introduced 2-D Amplitude and Phase EStimation (APES) amplitude estimator, which has been shown to yield superior performance over its competitors. To benchmark the proposed scheme, we also derive the Cramér–Rao bound (CRB) for the 2-D system identification problem.**

***Index Terms*—Cramér–Rao bound, 2-D amplitude estimation, 2-D system identification.**

I. Introduction

CONSIDER the following two-dimensional (2-D) linear discrete-time system

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

where $u(n, \overline{n})$ denotes the probing signal and $v(n, \overline{n})$ the (possibly *spectrally colored*) measurement noise. The rational system transfer function is given by [2], [3]

$$H(z^{-1}, \overline{z}^{-1}) := \frac{B(z^{-1}, \overline{z}^{-1})}{A(z^{-1}, \overline{z}^{-1})} := \frac{\sum_{i=0}^{r-1} \sum_{j=0}^{s-1} b_{i,j} z^{-i} \overline{z}^{-j}}{\sum_{i=0}^{p-1} \sum_{j=0}^{q-1} a_{i,j} z^{-i} \overline{z}^{-j}} \tag{2}$$

where, without loss of generality, $a_{0,0} = 1$, $b_{0,0} = 0$, and $(z^{-1}, \overline{z}^{-1})$ are the unit-delay operators. Let $\boldsymbol{a} := [a_{0,1} \; \cdots \; a_{0,q-1} \; \cdots \; a_{p-1,0} \; \cdots \; a_{p-1,q-1}]^T$ and $\boldsymbol{b} := [b_{0,1} \; \cdots \; b_{0,s-1} \; \cdots \; b_{r-1,0} \; \cdots \; b_{r-1,s-1}]^T$. Assume the system orders r, s, p, and q are known (determined using some model selection methods in, e.g., [3], [4]). The problem of interest is to estimate $\boldsymbol{a}$ and $\boldsymbol{b}$ from the system outputs $\{x(n, \overline{n})\}$.

Manuscript received July 31, 2001; revised December 28, 2001. This work was supported by the New Jersey Commission on Science and Technology, the Senior Individual Grant Program of the Swedish Foundation for Strategic Research, and the National Science Foundation under Grant MIP-9457388. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Dimitras A. Pados.

H. Li and W. Sun are with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ 07030 USA (e-mail: hli@stevens-tech.edu).

P. Stoica is with the Department of Systems and Control, Uppsala University, SE-751 03, Uppsala, Sweden.

J. Li is with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611 USA.

Publisher Item Identifier S 1070-9908(02)03405-3.

II. Two-Dimensional System Identification

A standard technique to solve the above problem is the output error method (OEM) [2], [3], [5] that minimizes

$$C(\boldsymbol{a}, \boldsymbol{b}) = \sum_{n=0}^{N-1} \sum_{\overline{n}=0}^{\overline{N}-1} \left| x(n, \overline{n}) - H\left(z^{-1}, \overline{z}^{-1}\right) u(n, \overline{n}) \right|^2. \tag{3}$$

The above minimization is usually performed by some *iterative nonlinear* optimization scheme, a process that is computationally *intensive* and *sensitive* to the choice of initial values for the unknown parameters. Moreover, the OEM is *statistically inefficient* when $v(n, \overline{n})$ is spectrally colored [2], [3], [5].

Next, we introduce a *closed-form* 2-D system identification scheme that relies on amplitude estimation and is computationally *simpler* and yet statistically more *accurate* than the OEM when $v(n, \overline{n})$ is colored. The proposed scheme utilizes an input signal composed of K probing 2-D sinusoids

$$u(n, \overline{n}) = \sum_{k=1}^{K} \gamma_k e^{j2\pi(f_k n + \overline{f}_k \overline{n})}$$

where we assume $K \geq J$ and $J := (pq - 1) + (rs - 1)$ denotes the number of unknown parameters in (2). Let $\alpha_k(\boldsymbol{a}, \boldsymbol{b}) := \gamma_k H(e^{-j2\pi f_k}, e^{-j2\pi \overline{f}_k})$. For sufficiently large N and $\overline{N}$ such that the transient response in the output may be ignored, (1) with the sinusoidal input can be approximated as

$$x(n, \overline{n}) = \sum_{k=1}^{K} \alpha_k(\boldsymbol{a}, \boldsymbol{b}) e^{j2\pi(f_k n + \overline{f}_k \overline{n})} + v(n, \overline{n}), \quad n = 0, \ldots, N-1; \overline{n} = 0, \ldots, \overline{N}-1 \tag{4}$$

from which we can estimate $\{\alpha_k(\boldsymbol{a}, \boldsymbol{b})\}_{k=1}^K$ in an *unstructured* manner by using the 2-D amplitude estimation technique discussed in Section III. Once the estimates of $\{\alpha_k(\boldsymbol{a}, \boldsymbol{b})\}_{k=1}^K$ are obtained, we select the J largest ones (in magnitude) out of the K amplitude estimates and denote them by $\hat{\alpha}_k$, $k = 1, \ldots, J$. Given these amplitude estimates, we can determine $\boldsymbol{a}$ and $\boldsymbol{b}$ by setting $\hat{\alpha}_k = \alpha_k(\boldsymbol{a}, \boldsymbol{b})$, or, equivalently,

$$\hat{\alpha}_k A\left(e^{-j2\pi f_k}, e^{-j2\pi \overline{f}_k}\right) = \gamma_k B\left(e^{-j2\pi f_k}, e^{-j2\pi \overline{f}_k}\right), \qquad k = 1, \ldots, J. \tag{5}$$

In contrast to the *nonlinear* OEM criterion (3), we have obtained in (5) a set of J *linear* equations with J unknowns, which can be solved *in closed-form*, yielding estimates of the system parameters $\boldsymbol{a}$ and $\boldsymbol{b}$.

As shown for the one-dimensional (1-D) case in [1], the proposed 2-D system identification scheme is *asymptotically (large-sample) statistically efficient*, irrespective of the color of $v(n, \overline{n})$, as long as the 2-D amplitude estimates $\{\hat{\alpha}_k\}$ are asymptotically efficient. This conclusion also follows precisely from the extended invariance principle (EXIP) in [6].

III. Two-Dimensional Amplitude Estimation

In [1], we have introduced/studied a variety of 1-D amplitude estimators, which are all asymptotically statistically efficient (thus *asymptotically equivalent*) but with quite *different finite-sample properties*. The 2-D extensions of the estimators can be shown to have similar behaviors to those of their 1-D counterparts. Among the various choices, we consider the 2-D Amplitude and Phase EStimation (APES) amplitude estimator [7], which surpasses their rivals in several aspects (see [1] for details) and is described next.

Let $\mathbf{X}$ be the $N\times\overline{N}$ matrix formed from $\{x(n, \overline{n})\}$. We break $\mathbf{X}$ into $M \times \overline{M}$ overlapping submatrices: $\mathbf{X}_{l,\bar{l}} = \{x(k, \overline{k}),$ $k = l, \ldots, l+M-1; \overline{k} = \bar{l}, \ldots, \bar{l}+\overline{M}-1\}, l = 0, \ldots, L-1;$ $\bar{l} = 0, \ldots, \overline{L}-1$, where $L := N-M+1$ and $\overline{L} := \overline{N}-\overline{M}+1$ (see [8] for how to choose M and $\overline{M}$). Let $\boldsymbol{x}_{l,\bar{l}} := \mathrm{vec}\{\mathbf{X}_{l,\bar{l}}\}$, where $\mathrm{vec}\{\cdot\}$ stacks the columns of the argument on top of one another [9], and $\boldsymbol{\xi}_k$ be the normalized 2-D DFT (discrete Fourier transform) of $\{\boldsymbol{x}_{l,\bar{l}}\}$ at frequencies $\{f_k, \overline{f}_k\}$, i.e.,

$$\boldsymbol{\xi}_k := \frac{1}{L\overline{L}} \sum_{l=0}^{L-1} \sum_{\bar{l}=0}^{\overline{L}-1} \boldsymbol{x}_{l,\bar{l}} e^{-j2\pi(f_k l + \overline{f}_k \bar{l})}.$$

The 2-D APES amplitude estimator is given by [7], [8]

$$\hat{\alpha}_k^{APES} = \frac{\left[\bar{\boldsymbol{a}}_{\overline{M}}(\overline{f}_k)\otimes\boldsymbol{a}_M(f_k)\right]^H\left(\hat{\mathbf{R}}-\boldsymbol{\xi}_k\boldsymbol{\xi}_k^H\right)^{-1}\boldsymbol{\xi}_k}{\left[\bar{\boldsymbol{a}}_{\overline{M}}(\overline{f}_k)\otimes\boldsymbol{a}_M(f_k)\right]^H\left(\hat{\mathbf{R}}-\boldsymbol{\xi}_k\boldsymbol{\xi}_k^H\right)^{-1}\left[\bar{\boldsymbol{a}}_{\overline{M}}(\overline{f}_k)\otimes\boldsymbol{a}_M(f_k)\right]}, \quad k = 1, \ldots, K \tag{6}$$

where $\boldsymbol{a}_M(f_k) := [1 \quad e^{j2\pi f_k} \quad \cdots \quad e^{j2\pi(M-1)f_k}]^T$ and $\bar{\boldsymbol{a}}_{\overline{M}}(\overline{f}_k) := [1 \quad e^{j2\pi\overline{f}_k} \quad \cdots \quad e^{j2\pi(\overline{M}-1)\overline{f}_k}]^T$ are the *steering vectors*, $\hat{\mathbf{R}} := (1/L\overline{L})\sum_{l=0}^{L-1}\sum_{\bar{l}=0}^{\overline{L}-1} \boldsymbol{x}_{l,\bar{l}}\boldsymbol{x}_{l,\bar{l}}^H$ denotes the *sample covariance matrix*, and $(\cdot)^T$, $(\cdot)^H$ and $\otimes$ denotes the transpose, conjugate transpose, and matrix Kronecker product [9], respectively.

IV. CRB for 2-D System Identification

Next, we derive the Cramér–Rao bound (CRB) for the 2-D system identification problem, which is not available in the literature. Let $a_{R_{i,j}} := \mathrm{Re}\{a_{i,j}\}$ and $a_{I_{i,j}} := \mathrm{Im}\{a_{i,j}\}$; let $b_{R_{i,j}}$ and $b_{I_{i,j}}$ be similarly defined from $b_{i,j}$. Let $\boldsymbol{\theta} \in \mathbb{R}^{2J\times 1}$ collect the real and imaginary parts of all unknown parameters: $\boldsymbol{\theta} := [\boldsymbol{a}_R^T \quad \boldsymbol{b}_R^T \quad \boldsymbol{a}_I^T \quad \boldsymbol{b}_I^T]^T$, where $\boldsymbol{a}_R := [a_{R_{0,1}} \quad \cdots \quad a_{R_{0,q-1}} \quad \cdots \quad a_{R_{p-1,q-1}}]^T \in \mathbb{R}^{(pq-1)\times 1}$, $\boldsymbol{b}_R := [b_{R_{0,1}} \quad \cdots \quad b_{R_{0,s-1}} \quad \cdots \quad b_{R_{r-1,s-1}}]^T \in \mathbb{R}^{(rs-1)\times 1}$, and $\boldsymbol{a}_I$ and $\boldsymbol{b}_I$ are similarly formed from $\{a_{I_{i,j}}\}$ and $\{b_{I_{i,j}}\}$, respectively. Let

$$w(n, \overline{n}) := H\left(z^{-1}, \overline{z}^{-1}\right) u(n, \overline{n}).$$

It is easy to verify that

$$w(n, \overline{n}) = \boldsymbol{\phi}_{n,\overline{n}}^T \boldsymbol{\theta}, \qquad n = 0, \ldots, N-1; \overline{n} = 1, \ldots, \overline{N}-1$$

where $\boldsymbol{\phi}_{n,\overline{n}} := [\boldsymbol{w}_{n,\overline{n}}^T \quad \boldsymbol{u}_{n,\overline{n}}^T \quad j\boldsymbol{w}_{n,\overline{n}}^T \quad j\boldsymbol{u}_{n,\overline{n}}^T]^T \in \mathbb{C}^{2J\times 1}$, and where

$$\begin{aligned}\boldsymbol{w}_{n,\overline{n}} := [&-w(n, \overline{n}-1) \cdots -w(n, \overline{n}-q+1) \cdots \\ &-w(n-p+1, \overline{n}) \cdots -w(n-p+1, \overline{n}-q+1)]^T \\ \boldsymbol{u}_{n,\overline{n}} := [&u(n, \overline{n}-1) \cdots u(n, \overline{n}-s+1) \cdots \\ &u(n-r+1, \overline{n}) \cdots u(n-r+1, \overline{n}-s+1)]^T.\end{aligned}$$

Let $\mathbf{X}$, $\mathbf{W}$ and $\mathbf{V}$ be $N \times \overline{N}$ matrices formed from $\{x(n, \overline{n})\}$, $\{w(n, \overline{n})\}$, and $\{v(n, \overline{n})\}$, respectively. Let $\boldsymbol{x} := \mathrm{vec}\{\mathbf{X}\}$, $\boldsymbol{w} := \mathrm{vec}\{\mathbf{W}\}$, and $\boldsymbol{v} := \mathrm{vec}\{\mathbf{V}\}$. Then, we have

$$\boldsymbol{x} = \boldsymbol{w} + \boldsymbol{v} = \boldsymbol{\Phi}\boldsymbol{\theta} + \boldsymbol{v}$$

where $\boldsymbol{\Phi} := [\boldsymbol{\phi}_{0,0} \quad \boldsymbol{\phi}_{1,0} \quad \cdots \quad \boldsymbol{\phi}_{N-1,\overline{N}-1}]^T \in \mathbb{C}^{N\overline{N}\times 2J}$. Under the assumption that the measurement noise $\boldsymbol{v}$ is circularly Gaussian with zero mean and covariance matrix $\boldsymbol{\Gamma}$, the CRB is given by (e.g., [10, App. B]

$$\mathrm{CRB}^{-1}(\boldsymbol{\theta}) = 2\mathrm{Re}\left(\frac{\partial^H \boldsymbol{w}}{\partial\boldsymbol{\theta}}\boldsymbol{\Gamma}^{-1}\frac{\partial\boldsymbol{w}}{\partial\boldsymbol{\theta}^T}\right).$$

To calculate $\partial\boldsymbol{w}/\partial\boldsymbol{\theta}^T$, we first note that

$$\begin{aligned}\frac{\partial w(n, \overline{n})}{\partial a_{R_{i,j}}} &= -\frac{B(z^{-1}, \overline{z}^{-1})}{A^2(z^{-1}, \overline{z}^{-1})} u(n-i, \overline{n}-j) \\ &= -\frac{1}{A(z^{-1}, \overline{z}^{-1})} w(n-i, \overline{n}-j), \\ \frac{\partial w(n, \overline{n})}{\partial b_{R_{i,j}}} &= \frac{1}{A(z^{-1}, \overline{z}^{-1})} u(n-i, \overline{n}-j)\end{aligned}$$

furthermore, we have $(\partial w(n, \overline{n})/\partial a_{I_{i,j}}) = j\left(\partial w(n, \overline{n})/\partial a_{R_{i,j}}\right)$ and $(\partial w(n, \overline{n})/\partial b_{I_{i,j}}) = j\left(\partial w(n, \overline{n})/\partial b_{R_{i,j}}\right)$. Hence,

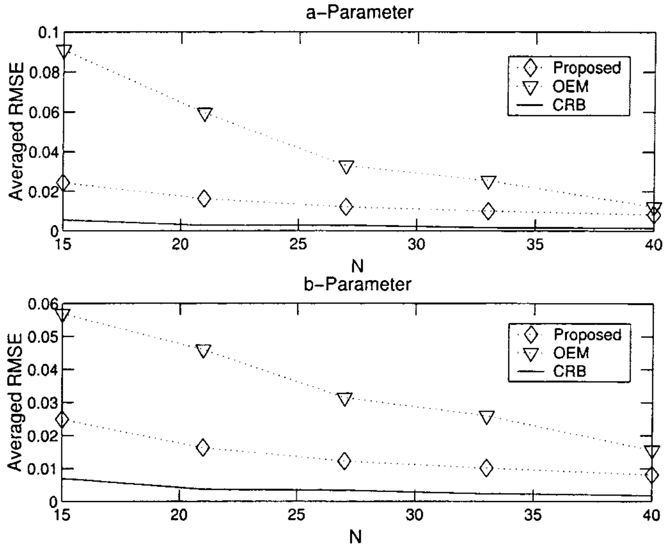
$$\frac{\partial w(n, \overline{n})}{\partial\boldsymbol{\theta}^T} = \frac{1}{A\left(z^{-1}, \overline{z}^{-1}\right)} \boldsymbol{\phi}_{n,\overline{n}}^T := \boldsymbol{\zeta}_{n,\overline{n}}^T. \tag{7}$$

Let $\zeta_{n,\overline{n}}(i)$ and $\phi_{n,\overline{n}}(i)$ denote the ith element of $\boldsymbol{\zeta}_{n,\overline{n}}$ and $\boldsymbol{\phi}_{n,\overline{n}}$, respectively. Equation (7) implies that $\zeta_{n,\overline{n}}(i)$ is obtained from $\phi_{n,\overline{n}}(i)$ through linear regression:

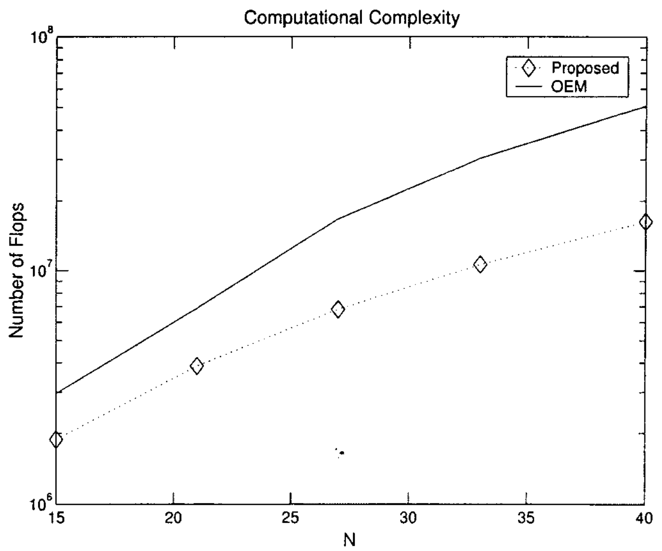
$$A\left(z^{-1}, \overline{z}^{-1}\right)\zeta_{n,\overline{n}}(i) = \phi_{n,\overline{n}}(i), \qquad i = 1, \ldots, 2J. \tag{8}$$

Consequently, we have

$$\frac{\partial\boldsymbol{w}}{\partial\boldsymbol{\theta}^T} = \frac{1}{A\left(z^{-1}, \overline{z}^{-1}\right)}\boldsymbol{\Phi} := \boldsymbol{\Delta}$$



(a)



(b)

Fig. 1. (a) ARMSE and (b) number of flops versus $N = \overline{N}$ when $M = \overline{M} = 5$ and an AR observation noise is present.

where $\boldsymbol{\Delta} \in \mathbb{C}^{N\overline{N} \times 2J}$ with each element obtained in the same manner as in (8). It follows that the CRB can be written as

$$\mathrm{CRB}(\boldsymbol{\theta}) = \frac{1}{2}\left[\mathrm{Re}\left(\boldsymbol{\Delta}^H \boldsymbol{\Gamma}^{-1} \boldsymbol{\Delta}\right)\right]^{-1}. \tag{9}$$

Noting that $\boldsymbol{\Delta}$ has the form: $\boldsymbol{\Delta} = [\tilde{\boldsymbol{\Delta}} \quad j\tilde{\boldsymbol{\Delta}}]$, where $\tilde{\boldsymbol{\Delta}}$ consists of the first J columns of $\boldsymbol{\Delta}$, we can readily verify that the CRB corresponding to the complex vector, $\tilde{\boldsymbol{\theta}} := [\boldsymbol{a}^T \quad \boldsymbol{b}^T]^T \in \mathbb{C}^{J \times 1}$, is given by

$$\mathrm{CRB}(\tilde{\boldsymbol{\theta}}) = \left(\tilde{\boldsymbol{\Delta}}^H \boldsymbol{\Gamma}^{-1} \tilde{\boldsymbol{\Delta}}\right)^{-1}. \tag{10}$$

The evaluation of the CRB (9) and (10) requires initial values of $w(n, \overline{n})$ for $1 - p \leq n \leq -1$, $1 - q \leq \overline{n} \leq -1$, and $u(n, \overline{n})$ for $1 - r \leq n \leq -1$, $1 - s \leq \overline{n} \leq -1$, which are set to zero in our simulations in Section V.

V. Numerical Results

Consider a system described by $A(z^{-1}, \overline{z}^{-1}) = 1 - j0.8\overline{z}^{-1} - (0.529 + j0.7281)z^{-1} + (-0.5825 + j0.4232)z^{-1}\overline{z}^{-1}$ and $B(z^{-1}, \overline{z}^{-1}) = (0.2014 - j0.7846)\overline{z}^{-1} - (0.2194 + j0.6753)z^{-1} + (-0.574 + j0.0361)z^{-1}\overline{z}^{-1}$. The noise is generated by a 2-D autoregressive (AR) process: $v(n, \overline{n}) = 0.99v(n-1, \overline{n}-1) + e(n, \overline{n})$, where $e(n, \overline{n})$ is a complex white Gaussian noise with zero-mean and variance 0.01. The probing signal consists of a sum of $K = 8$ unit-amplitude sinusoids at frequencies: (−0.45, 0.48), (−0.3167, 0.3467), (−0.1833, 0.2133), (−0.05, 0.08), (0.05, −0.08), (0.1833, −0.2133), (0.3167, −0.3467), and (0.45, −0.48). The performance criteria are the averaged root mean squared errors (ARMSE) of the parameter estimates and the number of MATLAB flops associated with each method. The ARMSE for $\boldsymbol{a}$ is defined as $\mathrm{ARMSE}\{\hat{\boldsymbol{a}}\} = (1/(pq-1)) \sum_{i,\, j,\, (i,\, j) \neq (0,\, 0)} \mathrm{RMSE}\ \{\hat{a}_{i,\, j}\}$; the ARMSE for $\boldsymbol{b}$ is similarly defined. Fig. 1(a) depicts the ARMSE performance of OEM as well as the proposed approach used with the 2-D APES amplitude estimator; also shown is the CRB derived in Section IV. Fig. 1(b) shows the number of flops required by each method. It is seen that the proposed 2-D system identification scheme not only yields more accurate parameter estimates, but also is computationally more efficient than the iterative OEM method.

References

[1] P. Stoica, H. Li, and J. Li, “Amplitude estimation of sinusoidal signals: Survey, new results, and an application,” *IEEE Trans. Signal Processing*, vol. 48, pp. 338–352, Feb. 2000.

[2] L. Ljung, *System Identification: Theory for the User*, 2 ed. Upper Saddle River, NJ: Prentice-Hall, 1999.

[3] T. Söderström and P. Stoica, *System Identification*. London, U.K.: Prentice-Hall, 1989.

[4] M. Wax and T. Kailath, “Detection of signals by information theoretic criteria,” *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 33, pp. 387–392, Apr. 1985.

[5] P. V. Kabaila, “On output-error methods for system identification,” *IEEE Trans. Automat. Contr.*, vol. 28, pp. 12–23, Jan. 1983.

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

[7] J. Li and P. Stoica, “An adaptive filtering approach to spectral estimation and SAR imaging,” *IEEE Trans. Signal Processing*, vol. 44, pp. 1469–1484, June 1996.

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

[9] A. Graham, *Kronecker Products and Matrix Calculus With Applications*. Chichester, U.K.: Ellis Horwood, 1981.

[10] P. Stoica and R. L. Moses, *Introduction to Spectral Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1997.