Least Squares Estimation of 2-D Sinusoids in Colored Noise: Asymptotic Analysis

Guy Cohen and Joseph M. Francos, *Senior Member, IEEE*

***Abstract*—This paper considers the problem of estimating the parameters of real-valued two-dimensional (2-D) sinusoidal signals observed in colored noise. This problem is a special case of the general problem of estimating the parameters of a real-valued homogeneous random field with mixed spectral distribution from a single observed realization of it. The large sample properties of the least squares (LS) estimator of the parameters of the sinusoidal components are derived, making no assumptions on the type of the probability distribution of the observed field. It is shown that if the disturbance field satisfies a combination of conditions comprised of a strong mixing condition and a condition on the order of its uniformly bounded moments, the normalized estimation error of the LS estimator is consistent asymptotically normal with zero mean and a normalized asymptotic covariance matrix for which a simple expression is derived. It is further shown that the LS estimator is asymptotically unbiased. The normalized asymptotic covariance matrix is block diagonal where each block corresponds to the parameters of a different sinusoidal component. Assuming further that the colored noise field is Gaussian, the LS estimator of the sinusoidal components is shown to be asymptotically efficient.**

***Index Terms*—Cramer–Rao bound (CRB), least squares (LS) estimation, regression spectrum, strong mixing property, two-dimensional (2-D) colored noise, 2-D random fields, 2-D sinusoids, 2-D Wold decomposition.**

I. Introduction

In this paper we consider the problem of estimating the parameters of two-dimensional (2-D) sinusoids in colored observation noise. This problem is in fact a special case of the more general problem of estimating the parameters of a 2-D regular and homogeneous random field from a single observed realization of it, [2], [7]. This modeling and estimation problem has fundamental theoretical importance, as well as various applications in texture estimation of images (see, e.g., [25] and the references therein) and in wave propagation problems (see, e.g., [26] and the references therein). From the 2-D Wold-like decomposition [1], we have that any 2-D regular and homogeneous discrete random field can be represented as a sum of two mutually orthogonal components: a *purely indeterministic* field and a *deterministic* one. The purely indeterministic component has a unique white innovations-driven moving-average representation. The deterministic component is further orthogonally decomposed into a *harmonic* field and a countable number of mutually orthogonal *evanescent* fields. This decomposition results in a corresponding decomposition of the spectral measure of the regular random field into a countable sum of mutually singular spectral measures. The spectral distribution function of the purely indeterministic component is absolutely continuous. Furthermore, since the random field is regular, the spectral density of the purely indeterministic component is zero at most on a set of Lebesgue measure zero, [5], [6]. Thus, the spectral measure of the deterministic component is singular with respect to the Lebesgue measure, i.e., it is concentrated on a set of Lebesgue measure zero in the frequency plane.

An early discussion on the problem of analyzing 2-D homogeneous random fields with discontinuous spectral distribution functions can be found in [31]. Parameter estimation techniques of sinusoidal signals in additive white noise include the periodogram-based approximation (applicable for widely spaced sinusoids) to the maximum-likelihood (ML) solution [8], extensions to the Pisarenko harmonic decomposition [11], or the singular value decomposition [12]. More recently, a matrix enhancement and matrix pencil method for estimating the parameters of 2-D superimposed, complex-valued exponential signals was suggested in [13], and analyzed in [14]. Assuming the noise field is white, the Cramer–Rao lower bound for this problem was derived as well. The same problem is also considered in [19] where three methods based on the approach of parameter estimation via signal selectivity of signal subspace are derived. In [15], a least squares (LS) estimation algorithm for estimating the parameters of exponentials in complex white Gaussian noise is derived, assuming the signal-to-noise ratio (SNR) is high. For high SNR, the estimator is unbiased, and its error variance achieves the Cramer–Rao bound (CRB). Least squares estimation of 2-D complex sinusoids in circular independent and identically distributed (i.i.d.) residual noise is considered in [10]. Strong consistency and asymptotic normality of the estimator are established. The problem of ML estimation of 2-D superimposed, complex-valued exponential signals in complex white circular Gaussian noise has been recently considered in [9]. In [7], a conditional ML algorithm for jointly estimating the parameters of the harmonic, evanescent, and purely indeterministic components of a real-valued homogeneous random field from a single observed realization of it, is derived. It is shown that by introducing appropriate parameter transformations, the highly nonlinear least squares (NLLS) problem that results from maximizing the conditional likelihood function is transformed into a separable NLLS problem, such that by first estimating the unknown spectral supports of the harmonic and evanescent components, the problem of solving for the transformed parameters of the field is reduced to linear least squares. The problem

Manuscript received October 26, 1999; revised February 27, 2002. This work was supported in part by the Israel Ministry of Science under Grant 8635297.

The authors are with the Department of Electrical and Computer Engineering, Ben-Gurion University, Beer-Sheva 84105, Israel (e-mail: guycohen@ee.bgu.ac.il; francos@ee.bgu.ac.il).

Communicated by J. A. O'Sullivan, Associate Editor for Detection and Estimation.

Publisher Item Identifier 10.1109/TIT.2002.800727.

of estimating the parameters of real-valued one–dimensional (1-D) sinusoids observed in the presence of additive stationary process generated by filtering a white noise process with an exponentially stable and invertible linear filter is considered in [17]. Basic and optimal NLLS methods are introduced and the corresponding normalized error covariance matrices are evaluated. It is shown that the two methods achieve the same normalized asymptotic covariance and both attain the CRB in the Gaussian case. In [20], a decoupled parameter estimation (DPE) algorithm for estimating the parameters of 1-D and 2-D sinusoids corrupted by autoregressive noise is derived. The algorithm consists of two steps. The first step is a RELAX algorithm which employs the fast Fourier transform (FFT) to estimate the sinusoids' parameters, followed by a LS solution for estimating the autoregressive (AR) parameters. The statistical efficiency of the algorithm and its applications to target feature extraction are also considered. In [18], a computationally efficient eigenstructure-based 2-D MODE algorithm for estimating the frequencies of 2-D sinusoids observed in additive white Gaussian noise is derived. It is further shown, that the algorithm is statistically efficient and superior to subspace rotation methods. Under the same modeling assumptions, a computationally efficient 1-D MODE algorithm is employed in [21] to estimate the frequencies of 2-D complex sinusoids. It is shown that the 1-D MODE algorithm is computationally more efficient than the asymptotically statistically efficient 2-D MODE algorithm.

The asymptotic CRB on the parameters of a Gaussian purely indeterministic field was derived by Whittle [3]. More recently, this general derivation was specialized for the case of noncausal AR models, and nonsymmetrical half plane (NSHP) AR models in [4]. In [2], the Cramer–Rao lower bound on the error variance in jointly estimating the parameters of the purely indeterministic, harmonic, and evanescent components of a homogeneous Gaussian random field, based on a *finite dimension, single observed realization* of this field, is established.

In this paper, we consider the problem of *LS* estimation of the parameters of the harmonic component of the field in the presence of the purely indeterministic component. More specifically, using the results of [28]–[30], it is shown that the normalized estimation error of the LS estimator is asymptotically normal with a zero mean and a normalized asymptotic covariance matrix for which a simple expression is derived. This derivation makes no assumptions regarding the type of the probability distribution of the observed field. However, it is concluded that there is a tradeoff between the order of the moments of the purely indeterministic component that need to be uniformly bounded and the mixing rate of the purely indeterministic component. It is also shown that the LS estimator of the parameters of the 2-D sinusoids is asymptotically unbiased. Finally, assuming the purely indeterministic component is Gaussian, a sufficient condition on its mixing rate is derived such that the LS estimator of the sinusoidal components is asymptotically efficient. Using the finite sample expression of the CRB [2], it is further demonstrated that even for modest dimensions of the observed field and low SNRs, the normalized asymptotic error covariance matrix of the LS estimator (and the normalized asymptotic CRB) is rapidly approached (in terms of both SNR and data dimensions) by the exact, finite-sample CRB, when the CRB matrix is normalized by the same normalization matrix employed to normalize the LS estimator covariance matrix.

The paper is organized as follows. In Section II, we introduce our notations and assumptions. In Section III, following [16], [22], [28], [30], we introduce the regression correlation matrix and the corresponding regression spectrum for this problem. In Section IV, it is shown that the normalized estimation error of the LS estimator is asymptotically normal with a zero mean and a normalized asymptotic covariance matrix for which a simple expression is derived. Asymptotic unbiasedness of the estimator is also established. In Section V, asymptotic efficiency of the LS estimator is established for the case where the observed field is Gaussian. Section VI provides our concluding remarks.

II. Problem Definition

In this section, we introduce our notations and assumptions.

Let $\{y(m, n)\}$, $(m, n) \in \mathcal{Z}^2$ be the observed 2-D real-valued random field such that

$$y(m, n) = h(m, n) + \varepsilon(m, n) \tag{1}$$

and

$$h(m, n) = \sum_{p=1}^{P} \alpha_p \cos(\omega_p m + \nu_p n + \varphi_p). \tag{2}$$

Assumption 1: Let $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ denote the parameter vector of the harmonic field, i.e.,

$$\boldsymbol{\theta} = [\alpha_1 \quad \varphi_1 \quad \omega_1 \quad \nu_1 \quad \cdots \quad \alpha_P \quad \varphi_P \quad \omega_P \quad \nu_P]^T \tag{3}$$

where for all k the amplitude $\alpha_k \in [E_1, E_2]$ where $0 < E_1$ and $E_2 < \infty$. Let d be some positive constant. Assume further that $\varphi_k, \omega_k, \nu_k \in [-\pi, \pi]$ where either $\min(|\omega_k - \omega_j|) \geq d$ or $\min(|\nu_k - \nu_j|) \geq d$ for $k \neq j$. Hence, the parameter space $\boldsymbol{\Theta}$ is a bounded and closed subset of the $4P$-dimensional Euclidian space. It is therefore compact, by the Heine–Borel theorem.

Let $(\Omega, \mathcal{F}, P)$ be the probability space on which the random field $\{\varepsilon(m, n)\}$ is defined. Let $\boldsymbol{R}$ be a partially ordered (by inclusion) system of subsets in $\mathcal{Z}^2$. Also, for any $\Delta \in \boldsymbol{R}$ denote by $\mathcal{F}(\Delta)$ the σ-algebra generated by $\{\varepsilon(m, n)\colon (m, n) \in \Delta\}$. We define the distance between two sets $\Delta_1, \Delta_2 \in \boldsymbol{R}$ by

$$\rho(\Delta_1, \Delta_2) = \inf\{\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|_2, \boldsymbol{x}_1 \in \Delta_1, \boldsymbol{x}_2 \in \Delta_2\}.$$

Definition 1 (Rosenblatt Dependence Measure, [27]): Let $\mathcal{F}_1, \mathcal{F}_2$ be two σ-algebras of subsets of Ω. Define the Rosenblatt dependence measure

$$d(\mathcal{F}_1, \mathcal{F}_2) = \sup\{|P(AB) - P(A)P(B)|\colon A \in \mathcal{F}_1, B \in \mathcal{F}_2\}.$$

Let us further define the mixing rate

$$\alpha(\rho) = \sup d(\mathcal{F}(\Delta_1), \mathcal{F}(\Delta_2))$$

where the supremum is taken over all $\Delta_1, \Delta_2 \in \boldsymbol{R}$ whose distance is at least ρ. (The definitions of the mixing properties employed in this paper were originally defined for 1-D processes in [27] and were adapted to the case of multidimensional fields by [28]).

Assumption 2: The purely indeterministic component $\{\varepsilon(m, n)\}$ is a zero-mean, wide-sense-homogeneous field, with a positive spectral density $\phi(\omega, \nu)$. Assume that $\varepsilon(m, n)$ has uniformly bounded $4+\delta$ absolute moments for some $\delta > 0$. Let us further assume that $\{\varepsilon(m, n)\}$ has the strong mixing property, i.e., that there exists a decreasing function $\varphi(\cdot)$ of a positive variable ρ where $\varphi(\rho) \downarrow 0$ as $\rho \to \infty$, such that for any two subsets $\Delta_1, \Delta_2 \in \mathcal{Z}^2$,

$$d(\mathcal{F}(\Delta_1), \mathcal{F}(\Delta_2)) \leq \varphi(\rho(\Delta_1, \Delta_2)).$$

We shall assume that $\varphi(\rho) = O(\rho^{-(2+\gamma)})$ where γ is a positive constant and $(\gamma-2)\delta > 16$. Thus, the foregoing assumptions imply that $\alpha(\rho) \leq k\rho^{-(2+\gamma)}$.

Assumption 3: The number P of harmonic components is *a priori* known.

Let $\{\Delta_k\}$ be a sequence of rectangles in $\boldsymbol{R}$ such that

$$\Delta_k = \{(i, j) \in \mathcal{Z}^2 \,|\, 0 \leq i \leq M_k - 1, 0 \leq j \leq N_k - 1\}.$$

Definition 2: The sequence of subsets $\{\Delta_k\}$ is said to tend to infinity (we adopt the notation $\Delta_k \to \infty$) as $k \to \infty$ if

$$\lim_{k\to\infty} \min(M_k, N_k) = \infty$$

and

$$0 < \lim_{k\to\infty} (M_k/N_k) < \infty.$$

To simplify notations, we shall omit in the following the subscript k. Thus, the notation $\Delta \to \infty$ implies that both N and M tend to infinity as functions of k, and at roughly the same rate.

Let $\boldsymbol{y}$, $\boldsymbol{h}$, $\boldsymbol{\varepsilon}$ denote the observation, harmonic component, and purely indeterministic component column vectors, respectively, where

$$\boldsymbol{y} = [y(0, 0), \ldots, y(M-1, 0), y(0, 1), \ldots, y(M-1, 1), \ldots, y(0, N-1), \ldots, y(M-1, N-1)]^T \quad (4)$$

and $\boldsymbol{h}$, $\boldsymbol{\varepsilon}$ are similarly defined. Let $\boldsymbol{\Gamma}$ denote the covariance matrix of $\boldsymbol{\epsilon}$ and hence of $\boldsymbol{y}$ as well. Thus,

$$\boldsymbol{\Gamma} = \begin{bmatrix} \boldsymbol{\Gamma}^{(0)} & \boldsymbol{\Gamma}^{(-1)} & \cdots & \boldsymbol{\Gamma}^{(1-N)} \\ \boldsymbol{\Gamma}^{(1)} & \boldsymbol{\Gamma}^{(0)} & \cdots & \boldsymbol{\Gamma}^{(2-N)} \\ \vdots & \vdots & \cdots & \vdots \\ \boldsymbol{\Gamma}^{(N-1)} & \boldsymbol{\Gamma}^{(N-2)} & \cdots & \boldsymbol{\Gamma}^{(0)} \end{bmatrix} \quad (5)$$

where

$$\boldsymbol{\Gamma}^{(k)} = \begin{bmatrix} r_{0,k} & r_{-1,k} & \cdots & r_{1-M,k} \\ r_{1,k} & r_{0,k} & \cdots & r_{2-M,k} \\ \vdots & \vdots & \cdots & \vdots \\ r_{M-1,k} & r_{M-2,k} & \cdots & r_{0,k} \end{bmatrix}. \quad (6)$$

III. The Regression Spectrum

Define the $4P \times 4P$ normalization matrix

$$\boldsymbol{D}_\Delta = \begin{bmatrix} \boldsymbol{D} & 0 & \cdots & 0 \\ 0 & \boldsymbol{D} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \boldsymbol{D} \end{bmatrix} \quad (7)$$

where

$$\boldsymbol{D} = \begin{bmatrix} (MN)^{1/2} & 0 & 0 & 0 \\ 0 & (MN)^{1/2} & 0 & 0 \\ 0 & 0 & (M^3N)^{1/2} & 0 \\ 0 & 0 & 0 & (MN^3)^{1/2} \end{bmatrix}. \quad (8)$$

Define also the mean gradient vector with respect to parameter vector $\boldsymbol{\theta}$

$$\tilde{\boldsymbol{\Phi}}(m, n) = \frac{\partial h(m, n)}{\partial \boldsymbol{\theta}} = \begin{bmatrix} \cos(\omega_1 m + \nu_1 n + \varphi_1) \\ -\alpha_1 \sin(\omega_1 m + \nu_1 n + \varphi_1) \\ -\alpha_1 m \sin(\omega_1 m + \nu_1 n + \varphi_1) \\ -\alpha_1 n \sin(\omega_1 m + \nu_1 n + \varphi_1) \\ \vdots \\ \cos(\omega_P m + \nu_P n + \varphi_P) \\ -\alpha_P \sin(\omega_P m + \nu_P n + \varphi_P) \\ -\alpha_P m \sin(\omega_P m + \nu_P n + \varphi_P) \\ -\alpha_P n \sin(\omega_P m + \nu_P n + \varphi_P) \end{bmatrix} \quad (9)$$

and let

$$\boldsymbol{\Phi} = \frac{\partial \boldsymbol{h}}{\partial \boldsymbol{\theta}^T} = \begin{bmatrix} \tilde{\boldsymbol{\Phi}}^T(0, 0) \\ \tilde{\boldsymbol{\Phi}}^T(1, 0) \\ \vdots \\ \tilde{\boldsymbol{\Phi}}^T(M-1, 0) \\ \vdots \\ \tilde{\boldsymbol{\Phi}}^T(M-1, N-1) \end{bmatrix}. \quad (10)$$

Define

$$\tilde{\boldsymbol{\Psi}}(m, n) = \begin{bmatrix} e^{i(\omega_1 m + \nu_1 n + \varphi_1)} \\ i\alpha_1 e^{i(\omega_1 m + \nu_1 n + \varphi_1)} \\ i\alpha_1 m e^{i(\omega_1 m + \nu_1 n + \varphi_1)} \\ i\alpha_1 n e^{i(\omega_1 m + \nu_1 n + \varphi_1)} \\ \vdots \\ e^{i(\omega_P m + \nu_P n + \varphi_P)} \\ i\alpha_P e^{i(\omega_P m + \nu_P n + \varphi_P)} \\ i\alpha_P m e^{i(\omega_P m + \nu_P n + \varphi_P)} \\ i\alpha_P n e^{i(\omega_P m + \nu_P n + \varphi_P)} \end{bmatrix}. \quad (11)$$

Let $\Re\{\boldsymbol{x}\}$ denote the real part of some vector (matrix) $\boldsymbol{x}$. Thus, $\tilde{\boldsymbol{\Phi}}(m, n) = \Re\{\tilde{\boldsymbol{\Psi}}(m, n)\}$.

Next, consider the sequence of matrices

$$\begin{aligned} \boldsymbol{R}_{k,\ell}^{\Delta} &= \boldsymbol{D}_\Delta^{-1} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \tilde{\boldsymbol{\Phi}}(m+k, n+\ell) \tilde{\boldsymbol{\Phi}}^T(m, n) \boldsymbol{D}_\Delta^{-1} \\ &= \frac{1}{2} \boldsymbol{D}_\Delta^{-1} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \Re\{\tilde{\boldsymbol{\Psi}}(m+k, n+\ell) \tilde{\boldsymbol{\Psi}}^T(m, n)\} \boldsymbol{D}_\Delta^{-1} \\ &\quad + \frac{1}{2} \boldsymbol{D}_\Delta^{-1} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \Re\{\tilde{\boldsymbol{\Psi}}(m+k, n+\ell) \tilde{\boldsymbol{\Psi}}^H(m, n)\} \boldsymbol{D}_\Delta^{-1}. \end{aligned} \quad (12)$$

The last equality follows from the identity

$$\Re(\boldsymbol{A})\Re(\boldsymbol{B}) = \frac{1}{2}\Re(\boldsymbol{A}\boldsymbol{B}) + \frac{1}{2}\Re(\boldsymbol{A}\overline{\boldsymbol{B}})$$

where $\boldsymbol{A}$, $\boldsymbol{B}$ are two complex-valued matrices and $\overline{\boldsymbol{A}}$ denotes the matrix whose elements are the conjugates of the elements of $\boldsymbol{A}$, i.e., $\overline{\boldsymbol{A}} = (\boldsymbol{A}^H)^T$.

Note that

$$\lim_{N\to\infty} \frac{1}{N^{k+1}} \sum_{n=0}^{N-1} n^k e^{i\rho n} = \begin{cases} \frac{1}{k+1}, & \rho = 0 \\ 0, & \rho \neq 0 \end{cases} \tag{13}$$

(see, e.g., [33] for the case where $\rho = 0$. The result for $\rho \neq 0$ is easily derived by differentiating k times the geometric series $\sum_n e^{i\rho n}$). Using (13) and some straightforward arithmetic, it can be shown that as $\Delta \to \infty$

$$\frac{1}{2}\boldsymbol{D}_\Delta^{-1} \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} \Re\left\{\tilde{\boldsymbol{\Psi}}(m+k,\, n+\ell)\tilde{\boldsymbol{\Psi}}^T(m,\, n)\right\}\boldsymbol{D}_\Delta^{-1} = 0 \tag{14}$$

and hence that as $\Delta \to \infty$, the sequence $\{\boldsymbol{R}_{k,\,\ell}^\Delta\}$ tends to the limit denoted by $\boldsymbol{R}_{k,\,\ell}$

$$\boldsymbol{R}_{k,\,\ell} = \frac{1}{2}\Re\left\{\begin{bmatrix} e^{i(\omega_1 k+\nu_1\ell)}\boldsymbol{B}_1 & 0 & \cdots & 0 \\ 0 & e^{i(\omega_2 k+\nu_2\ell)}\boldsymbol{B}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{i(\omega_P k+\nu_P\ell)}\boldsymbol{B}_P \end{bmatrix}\right\} \tag{15}$$

where

$$\boldsymbol{B}_p = \begin{bmatrix} 1 & i\alpha_p & \frac{i\alpha_p}{2} & \frac{i\alpha_p}{2} \\ -i\alpha_p & \alpha_p^2 & \frac{\alpha_p^2}{2} & \frac{\alpha_p^2}{2} \\ -\frac{i\alpha_p}{2} & \frac{\alpha_p^2}{2} & \frac{\alpha_p^2}{3} & \frac{\alpha_p^2}{4} \\ -\frac{i\alpha_p}{2} & \frac{\alpha_p^2}{2} & \frac{\alpha_p^2}{4} & \frac{\alpha_p^2}{3} \end{bmatrix} \tag{16}$$

and

$$\boldsymbol{R}_{0,\,0} = \lim_{\Delta\to\infty} \boldsymbol{D}_\Delta^{-1}\boldsymbol{\Phi}^T\boldsymbol{\Phi}\boldsymbol{D}_\Delta^{-1} = \frac{1}{2}\Re\left\{\begin{bmatrix} \boldsymbol{B}_1 & 0 & \cdots & 0 \\ 0 & \boldsymbol{B}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \boldsymbol{B}_P \end{bmatrix}\right\}. \tag{17}$$

Note that in the terms of [16], [30], $\boldsymbol{R}_{k,\,\ell}$ is a regression correlation matrix.

We next show that the double-indexed sequence $\boldsymbol{R}_{k,\,\ell}$ admits a spectral representation. Since for all p, $\boldsymbol{B}_p$ is independent of ℓ and k while $\boldsymbol{B}_p = \boldsymbol{B}_p^H$, it can be easily verified using (15) that $\boldsymbol{R}_{-k,\,-\ell} = \boldsymbol{R}_{k,\,\ell}^T$ and $\boldsymbol{R}_{-k,\,\ell} = \boldsymbol{R}_{k,\,-\ell}^T$, i.e., $\boldsymbol{R}_{k,\,\ell}$ is a symmetric series of operators on $\mathcal{Z}^2$. Furthermore, the sequence $\boldsymbol{R}_{k,\,\ell}$ is nonnegative definite. (See Appendix A for the proof.) Since $\boldsymbol{R}_{k,\,\ell}$ is a double-index positive semidefinite symmetric sequence, we conclude using the spectral representation theorem (see, e.g., [5]) that there exists a nonnegative 2-D distribution function $M_\alpha(\omega,\,\nu)$ such that

$$\boldsymbol{\alpha}^T\boldsymbol{R}_{h,\,k}\boldsymbol{\alpha} = \frac{1}{4\pi^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} e^{i(h\omega+k\nu)}\,dM_\alpha(\omega,\,\nu). \tag{18}$$

However, $M_\alpha(\omega,\,\nu) = \boldsymbol{\alpha}^T\boldsymbol{M}(\omega,\,\nu)\boldsymbol{\alpha}$, where $\boldsymbol{M}(\omega,\,\nu)$ is a matrix-valued function of ω and ν taking as values Hermitian $4P \times 4P$ positive semidefinite matrices whose elements are functions of bounded variation, while the functions on the diagonal are nondecreasing. It can be shown following similar arguments to those in [16, p. 45], that $\boldsymbol{R}_{k,\,\ell}$ has a spectral representation of the form

$$\boldsymbol{R}_{k,\,\ell} = \frac{1}{4\pi^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} e^{i(k\omega+\ell\nu)}\,d\boldsymbol{M}(\omega,\,\nu). \tag{19}$$

Thus, $\boldsymbol{M}$, the measure induced by the matrix-valued function $\boldsymbol{M}(\omega,\,\nu)$ is given by

$$\boldsymbol{M} = \pi^2\begin{bmatrix} \delta_{\omega_1,\,\nu_1}\boldsymbol{B}_1 & 0 & \cdots & 0 \\ 0 & \delta_{\omega_2,\,\nu_2}\boldsymbol{B}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \delta_{\omega_P,\,\nu_P}\boldsymbol{B}_P \end{bmatrix} + \pi^2\begin{bmatrix} \delta_{-\omega_1,\,-\nu_1}\boldsymbol{B}_1^T & 0 & \cdots & 0 \\ 0 & \delta_{-\omega_2,\,-\nu_2}\boldsymbol{B}_2^T & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \delta_{-\omega_P,\,-\nu_P}\boldsymbol{B}_P^T \end{bmatrix} \tag{20}$$

where $\delta_{\omega_p,\,\nu_p}$ denotes the Dirac measure concentrated at $\omega_p,\,\nu_p$.

IV. LS Estimation of the Parameters of the Harmonic Component

Let

$$f_L(\boldsymbol{\theta}) = [\boldsymbol{y} - \boldsymbol{h}(\boldsymbol{\theta})]^T[\boldsymbol{y} - \boldsymbol{h}(\boldsymbol{\theta})] \tag{21}$$

be the quadratic objective function to be minimized with respect to the parameter vector $\boldsymbol{\theta}$ and let $\hat{\boldsymbol{\theta}}_\Delta$ be its global minimum over $\boldsymbol{\Theta}$. Define the normalized estimation error

$$\boldsymbol{D}_\Delta(\hat{\boldsymbol{\theta}}_\Delta - \boldsymbol{\theta}) \tag{22}$$

and let $F_\Delta(\boldsymbol{y},\,\boldsymbol{\theta})$ denote the distribution function of $\boldsymbol{D}_\Delta(\hat{\boldsymbol{\theta}}_\Delta - \boldsymbol{\theta})$.

Theorem 1: Assume the observations are given by (1) and (2). Assume further that the parameter space $\boldsymbol{\Theta}$ satisfies Assumption 1, and that the purely indeterministic component $\{\varepsilon(m,\,n)\}$ satisfies the conditions of Assumption 2. Then uniformly in $\boldsymbol{\Theta}$

$$F_\Delta(\boldsymbol{y},\,\boldsymbol{\theta}) \quad \text{is} \quad \mathrm{AN}(\boldsymbol{0},\,\boldsymbol{\Sigma}) \quad \text{as} \quad \Delta\to\infty \tag{23}$$

where

$$\boldsymbol{\Sigma} = (\boldsymbol{R}_{0,\,0})^{-1}\frac{1}{4\pi^2}\left[\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\phi(\omega,\,\nu)\,d\boldsymbol{M}(\omega,\,\nu)\right](\boldsymbol{R}_{0,\,0})^{-1}. \tag{24}$$

Proof: The proof follows immediately from the general result in [29, Theorem 3.4.2]. This is because it is shown in

Section III that the gradient of the regression functions admits a spectral representation, while the remaining conditions that the gradient of the regression functions must satisfy in order for [29, Theorem 3.4.2] to hold, are satisfied in the special case where the regression functions are the sinusoids in (2). This is because the sinusoids are bounded and infinitely differentiable functions, because of (13), and due to Assumption 1 that guarantees that the uniform convergence conditions required by [29, Theorem 3.4.2] are satisfied. □

Substituting (17) and (20) into (24) we have (25), shown at the bottom of the page, where

$$\boldsymbol{C}_p = 2\begin{bmatrix} \phi(\omega_p, \nu_p) & 0 & 0 & 0 \\ 0 & \frac{7\phi(\omega_p,\nu_p)}{\alpha_p^2} & -\frac{6\phi(\omega_p,\nu_p)}{\alpha_p^2} & -\frac{6\phi(\omega_p,\nu_p)}{\alpha_p^2} \\ 0 & -\frac{6\phi(\omega_p,\nu_p)}{\alpha_p^2} & \frac{12\phi(\omega_p,\nu_p)}{\alpha_p^2} & 0 \\ 0 & -\frac{6\phi(\omega_p,\nu_p)}{\alpha_p^2} & 0 & \frac{12\phi(\omega_p,\nu_p)}{\alpha_p^2} \end{bmatrix}. \tag{26}$$

We note that the conditions on the disturbance field $\{\varepsilon(n, m)\}$ specified by Assumption 2 imply that for Theorem 1 to hold there is a tradeoff between the order of the disturbance moments that need to be uniformly bounded and the measure of dependence of the field. More specifically, choosing a larger (smaller) value of δ implies that higher (lower) order moments of the disturbance field need to be uniformly bounded, while the measure of dependence $d(\mathcal{F}(\Delta_1), \mathcal{F}(\Delta_2))$ of any two subsets $\Delta_1, \Delta_2 \in \mathcal{Z}^2$ can be higher (lower) such that the inequality $(\gamma - 2)\delta > 16$ is satisfied. Since the mixing condition implies a similar type of condition for the covariance sequence, this tradeoff can be similarly stated in terms of the rate of decay of the covariance sequence versus the order of the disturbance field moments that need to be uniformly bounded. Examples of homogeneous random fields that satisfy the conditions of Assumption 2 (and hence Theorem 1) include finite support Markov random fields (MRF) and finite support Gaussian moving average (MA) fields. For MRFs, the strong mixing condition holds since samples with sufficiently large distance separating them, are independent. For Gaussian MA fields, the covariance sequence is zero for sufficiently large lags, while the distribution is completely determined by the second-order moments.

Theorem 2: Assume the observations are given by (1) and (2). Assume further that the parameter space $\boldsymbol{\Theta}$ satisfies Assumption 1, and that the purely indeterministic component $\{\varepsilon(m, n)\}$ satisfies the conditions of Assumption 2. Let $\boldsymbol{\theta}(i)$ denote the ith element of $\boldsymbol{\theta}$. Then, for each element of $\boldsymbol{\theta}$,

$$\lim_{\Delta\to\infty} E|\hat{\boldsymbol{\theta}}_\Delta(i) - \boldsymbol{\theta}(i)|^q = 0, \qquad 1 \le q < \infty.$$

The special case of $q = 1$ implies that $\hat{\boldsymbol{\theta}}_\Delta$ is an asymptotically unbiased estimate of $\boldsymbol{\theta}$.

To prove Theorem 2 we will need the following lemma (see Appendix B for its proof):

Lemma 1: Let $\{a_n\}$ be a positive monotone sequence tending to infinity, and let $\{X_n\}$ be a sequence of random variables such that $\{a_n X_n\}$ converges in distribution to an integrable random variable X, then $\{X_n\}$ converges in probability to zero.

Proof of Theorem 2: Applying Lemma 1 to the elements of $\boldsymbol{D}_\Delta(\hat{\boldsymbol{\theta}}_\Delta - \boldsymbol{\theta})$ we conclude using Theorem 1 that all the entries of $\hat{\boldsymbol{\theta}}_\Delta - \boldsymbol{\theta}$ converge to zero in probability. Hence, $|\hat{\boldsymbol{\theta}}_\Delta(i) - \boldsymbol{\theta}(i)|^q$ converges to zero in probability, for any $1 \le q < \infty$, as well. By assumption all the entries of $\hat{\boldsymbol{\theta}}_\Delta - \boldsymbol{\theta}$ are bounded. Hence, using Lebesgue bounded convergence theorem (see, e.g., [34, p. 96], [35]) we have $\lim_{\Delta\to\infty} E(|\hat{\boldsymbol{\theta}}_\Delta(i) - \boldsymbol{\theta}(i)|^q) = 0$. □

Remark: Since all the entries of $\hat{\boldsymbol{\theta}}_\Delta - \boldsymbol{\theta}$ converge to zero in probability, while $\boldsymbol{D}_\Delta(\hat{\boldsymbol{\theta}}_\Delta - \boldsymbol{\theta})$ converges in distribution to a multivariate Gaussian random variable, $\hat{\boldsymbol{\theta}}_\Delta$ is *consistent asymptotically normal* (CAN). The covariance of the limiting distribution $\boldsymbol{\Sigma}$ is the *normalized asymptotic covariance* of $\hat{\boldsymbol{\theta}}_\Delta$, [36].

Note that the normalized asymptotic covariance matrix in (25) and (26) is block diagonal where each block corresponds to the parameters of a different sinusoidal component. Moreover, its element that expresses the normalized asymptotic error variance in estimating the amplitude parameter of each sinusoid is decoupled and independent of *all* other model parameters. It is a function *only* of the colored noise spectral density at the sinusoid's frequency. It should be emphasized that this derivation of the large sample properties of the LS estimator is independent of the type of the probability distribution function of the observed field. Finally, we note that related results for the special cases, where the purely indeterministic component is a white noise field of some unknown distribution, and for the case where the white noise field is Gaussian, were derived in [10] and [9], respectively.

V. Asymptotic Efficiency of the Least Squares Estimator

A. Asymptotic Results

The CRB provides a lower bound on the error variance in estimating the model parameters for any unbiased estimator of these parameters. In [2], the finite-sample Cramer–Rao lower bound on the error variance in jointly estimating the parameters of the different components of a 2-D homogeneous Gaussian random field with mixed spectral distribution, from a single observed

$$\boldsymbol{\Sigma} = 2\begin{bmatrix} \phi(\omega_1, \nu_1)[\Re(\boldsymbol{B}_1)]^{-1} & & \cdots & 0 \\ & \phi(\omega_2, \nu_2)[\Re(\boldsymbol{B}_2)]^{-1} & & \\ \vdots & & \ddots & \vdots \\ 0 & & \cdots & \phi(\omega_P, \nu_P)[\Re(\boldsymbol{B}_P)]^{-1} \end{bmatrix} = \begin{bmatrix} \boldsymbol{C}_1 & & \cdots & 0 \\ & \boldsymbol{C}_2 & & \\ & & \ddots & \vdots \\ 0 & & \cdots & \boldsymbol{C}_P \end{bmatrix} \tag{25}$$

realization of it, is derived. The model defined in (1)–(3) is a special case of the model addressed in [2], as it assumes the observed field contains only the harmonic and purely indeterministic components of the 2-D Wold-like decomposition. Since the LS estimator of the harmonic component parameters was shown to be consistent asymptotically normal and asymptotically unbiased we investigate in this section its asymptotic efficiency. The problem of deriving the normalized asymptotic CRB (and the normalized asymptotic information matrix) for the Gaussian complex-valued model which is equivalent to the model defined in (1)–(3), using the normalization matrix $\boldsymbol{D}_\Delta$ defined in (7) has been recently investigated in [23]. Following similar arguments for the real-valued case, and assuming that the covariance sequence $\{r_{k,l}\}$ of the purely indeterministic component satisfies the conditions

$$\lim_{\Delta\to\infty} \frac{1}{\sqrt{MN}} \sum_{k=-M}^{M} \sum_{l=-N}^{N} |klr_{k,l}| = 0 \tag{27}$$

$$\lim_{\Delta\to\infty} \frac{1}{\sqrt{M}} \sum_{k=-M}^{M} \sum_{l=-N}^{N} |kr_{k,l}| = 0 \tag{28}$$

$$\lim_{\Delta\to\infty} \frac{1}{\sqrt{N}} \sum_{k=-M}^{M} \sum_{l=-N}^{N} |lr_{k,l}| = 0 \tag{29}$$

it can be shown that the normalized large sample CRB is given by (25) and (26). (We note that conditions (27)–(29) are trivially satisfied by finite support 2-D MA models). A CAN estimate for which the normalized asymptotic covariance matrix is equal to the normalized asymptotic CRB is asymptotically efficient (or best asymptotically normal (BAN), see e.g., [36]). Thus, we next show that in the case where $\{\varepsilon(n, m)\}$ is Gaussian, a sufficient condition for the LS estimator of the harmonic component parameters to be asymptotically efficient is that $\{\varepsilon(n, m)\}$ satisfies Assumption 2.

Since $\{\varepsilon(n, m)\}$ has uniformly bounded $4+\delta$ absolute moments, using the strong mixing assumption it can be shown (see, [28, Lemma 2]) that there exists some positive constant c such that

$$|r_{k,l}| \le c(\varphi(\rho(S, T)))^{\frac{2+\delta}{4+\delta}} \tag{30}$$

for any pair of sets $S, T \subseteq \mathcal{Z}^2$ for which $\varepsilon(0, 0)$ and $\varepsilon(k, l)$ are measurable with respect to $\mathcal{F}(S)$ and $\mathcal{F}(T)$, respectively. Hence, letting $S = \{(0, 0)\}, T = \{(k, l)\}$, and using Assumption 2 we have that as $\sqrt{k^2+l^2} \to \infty$

$$|r_{k,l}| \le \eta_1 \left(\sqrt{k^2+l^2}\right)^{-(2+\gamma)\frac{2+\delta}{4+\delta}} \tag{31}$$

where η_1 is a positive constant. We can write

$$(2+\gamma)\frac{2+\delta}{4+\delta} = 2+\eta_2$$

where

$$\eta_2 = \frac{\gamma\delta + 2\gamma - 4}{4+\delta} > 1$$

since we assume that $\gamma\delta - 2\delta > 16$. Note that (31) implies that the spectral density of the purely indeterministic field $\phi(\omega, \nu)$ is continuous and bounded.

Theorem 3: Assume the observations are given by (1) and (2). Assume further that the parameter space $\boldsymbol{\Theta}$ satisfies Assumption 1, and that the purely indeterministic component $\{\varepsilon(m, n)\}$ satisfies the conditions of Assumption 2. Assume also that $\{\varepsilon(m, n)\}$ is Gaussian. Then, the LS estimator of the harmonic component parameters is asymptotically efficient (BAN).

Proof: It has been shown that $\eta_2 > 1$. We next show that this condition is sufficient for (27)–(29) to hold. Since for all $(k, l), |r_{k,l}| \le r_{0,0}$, and since in the following we are interested only in bounding sums of the type found on the left-hand side of (32), we can assume without loosing the generality of this proof that (31) holds for all (k, l). (In fact, this assumption holds for all (k, l) except for a finite number of pairs.) Thus,

$$\sum_{k=-M}^{M} \sum_{l=-N}^{N} |klr_{k,l}| \le 4\eta_1 \sum_{k=1}^{M} \sum_{l=1}^{N} \frac{kl}{\left(\sqrt{k^2+l^2}\right)^{(2+\eta_2)}}. \tag{32}$$

Next, consider the function

$$f(x, y) = \frac{xy}{(x^2+y^2)^{1+\eta_2/2}}, \qquad x \ge 1, y \ge 1.$$

This positive function is continuous and strictly decreasing to zero in any direction. Hence, employing the Cauchy integral test we conclude that the right-hand side (RHS) of (32) converges or diverges with the same rate as the integral

$$\int_1^M \int_1^N \frac{xy}{(x^2+y^2)^{1+\eta_2/2}}\, dx\, dy.$$

Hence, we obtain using the transformation to polar coordinates that

$$\begin{aligned} &\sum_{k=-M}^{M} \sum_{l=-N}^{N} |klr_{k,l}| \\ &\le 4\eta_1 \int_1^M \int_1^N \frac{xy}{(x^2+y^2)^{1+\eta_2/2}}\, dx\, dy \\ &\le 4\eta_1 \int_1^{\sqrt{M^2+N^2}} \int_0^{\pi/2} \frac{r\cos\theta\, r\sin\theta\, r}{r^{2+\eta_2}}\, dr\, d\theta \\ &\le K \int_1^{\sqrt{M^2+N^2}} r^{1-\eta_2}\, dr \\ &\le K \left(\sqrt{M^2+N^2}\right)^{2-\eta_2} \end{aligned} \tag{33}$$

where K is some positive constant.

Considering now the normalized sum, we conclude that

$$\frac{1}{\sqrt{MN}} \sum_{k=-M}^{M} \sum_{l=-N}^{N} |klr_{k,l}| \le K \frac{\left(\sqrt{M^2+N^2}\right)^{2-\eta_2}}{\sqrt{MN}}. \tag{34}$$

On taking the limit, we conclude that as $\Delta \to \infty$, $\eta_2 > 1$ is a sufficient condition that the RHS of (34) tends to zero.

Following a similar line of proof it can be shown that

$$\frac{1}{\sqrt{M}} \sum_{k=-M}^{M} \sum_{l=-N}^{N} |kr_{k,l}| \le K \frac{\left(\sqrt{M^2+N^2}\right)^{1-\eta_2}}{\sqrt{M}}. \tag{35}$$

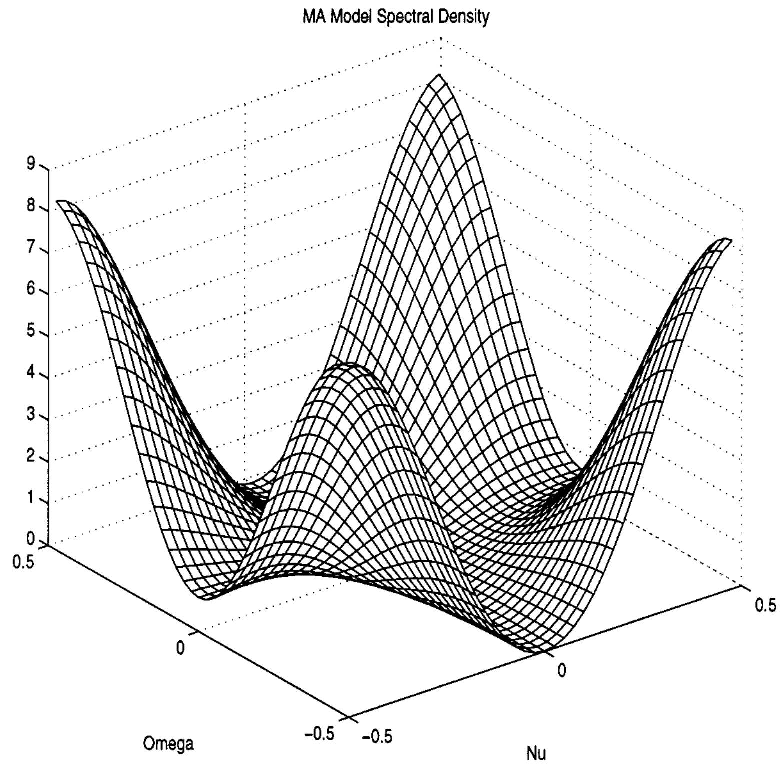

Fig. 1. The spectral density function of the MA field.

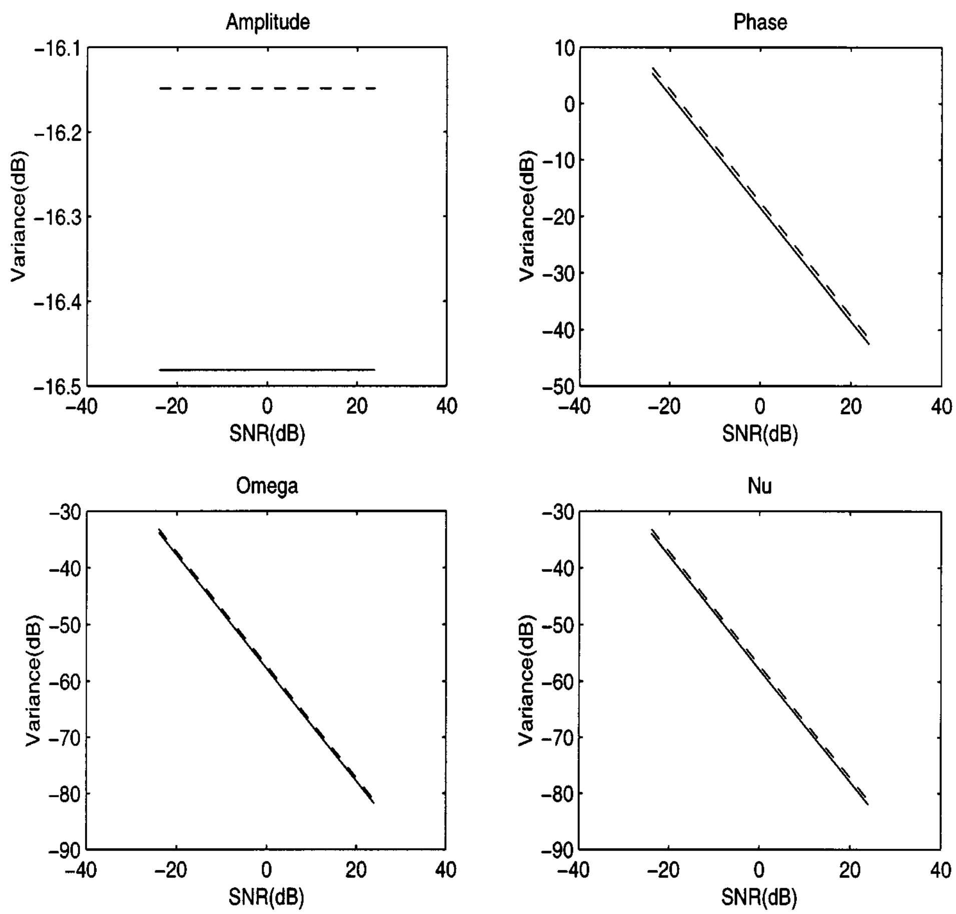

Fig. 2. The rescaled normalized asymptotic error variance of the LS estimate of the amplitude, phase, and spatial frequency as a function of SNR (dashed line), compared with the corresponding exact CRB (solid line).

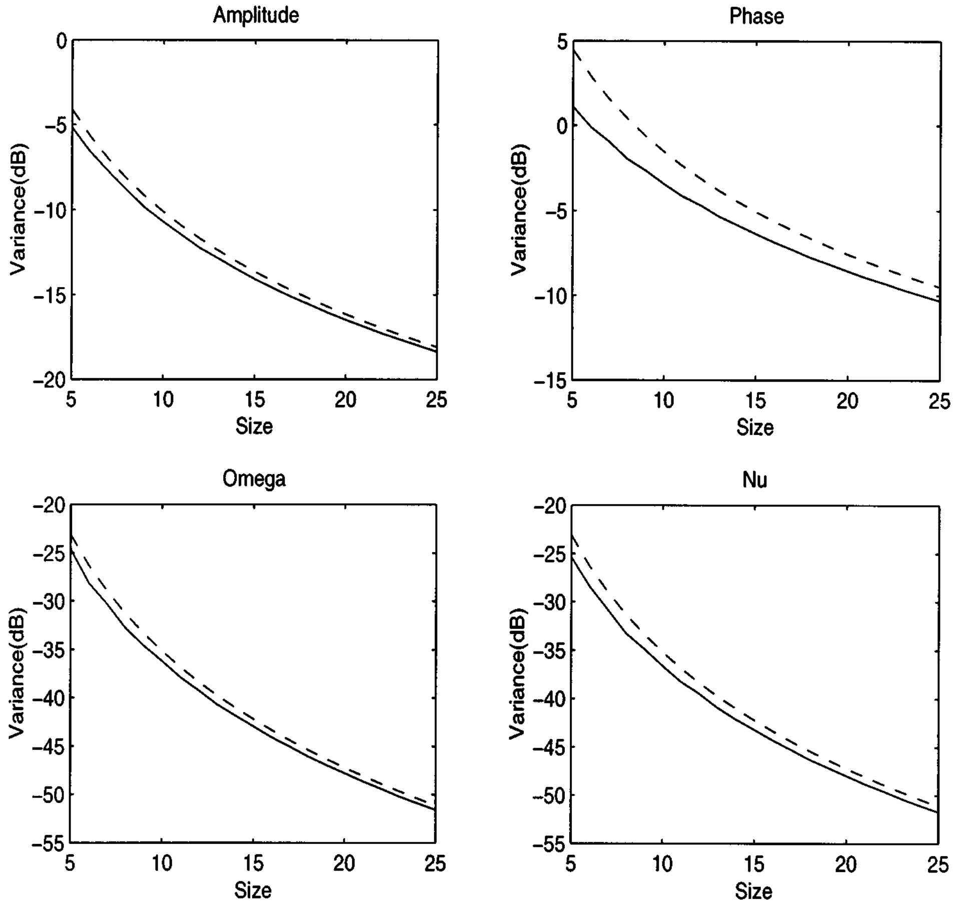

Fig. 3. The rescaled normalized asymptotic error variance of the LS estimate of the amplitude, phase, and spatial frequency as a function of data dimensions for SNR $= -10$ dB (dashed line), compared with the corresponding exact CRB (solid line).

Here, $\eta_2 > \frac{1}{2}$ is a sufficient condition that the RHS of (35) tends to zero. Similar arguments hold in evaluating the limit in (29). Thus, $\eta_2 > 1$ is a sufficient condition for (27)–(29) to hold. □

B. Finite Sample Results

In the following, we shall consider the normalized asymptotic covariance matrix of the LS estimator and the normalized asymptotic CRB matrix, when these matrices (shown in the previous section to be identical) are rescaled by multiplying them on the left and on the right by $\boldsymbol{D}_{\Delta}^{-1}$. Having established the sufficient condition for asymptotic efficiency of the LS estimator of the parameters of 2-D sinusoids observed in a colored Gaussian noise field, it is next shown using numerical evaluation of some specific examples that the rescaled normalized asymptotic error covariance of the LS estimator is rapidly approached (in terms of both SNR and data dimensions) by the exact, finite-sample CRB. In the following experiments, the rescaled normalized asymptotic covariance matrix of the LS estimator of $\boldsymbol{\theta}$ is compared with the *exact* CRB [2]. It should be noted that the performance of the LS estimator depicted in the following examples is not the result of any particular implementation of the estimator, but rather a performance curve of the asymptotic behavior of the estimator, evaluated analytically for various SNRs and data dimensions, based on the foregoing derivations.

1) Estimation Performance as a Function of SNR: In this subsection, we investigate the asymptotic performance of the LS estimator in comparison with the exact CRB, as a function of the local SNR. The local SNR for the kth sinusoid is defined as

$$\mathrm{SNR}_k = 10 \log \frac{\alpha_k^2}{\phi(\omega_k, \nu_k)}. \tag{36}$$

In this example, the purely indeterministic component of the field is an NSHP MA field with support $S_{1,1}$. The MA model parameters are $b(0, 1) = -0.9$, $b(1, -1) = 0.1$, $b(1, 0) = -0.5$, $b(1, 1) = 0.4$. The driving noise of the MA model is a zero mean, white Gaussian noise field with a unit variance. For illustration purposes, the spectral density function of the field is depicted in Fig. 1. The harmonic component of the field comprises a single sinusoid with frequency $(\omega_1, \nu_1) = (0.2, 0.15)$. Its amplitude varies to provide the desired range of SNR values. The dimensions of the observed field are 20×20 .

The results of this example, Fig. 2, indicate that even for modest dimensions of the observed field, and for a wide range of SNR values, the rescaled normalized asymptotic error variances of the LS estimates of the amplitude, phase, and spatial frequency are essentially identical to the corresponding values of the exact CRB. These CRB values are evaluated for the given dimensions of the observed data (20×20 in this case).

2) Estimation Performance as a Function of Data Dimensions: In this subsection, we investigate the effect of the size

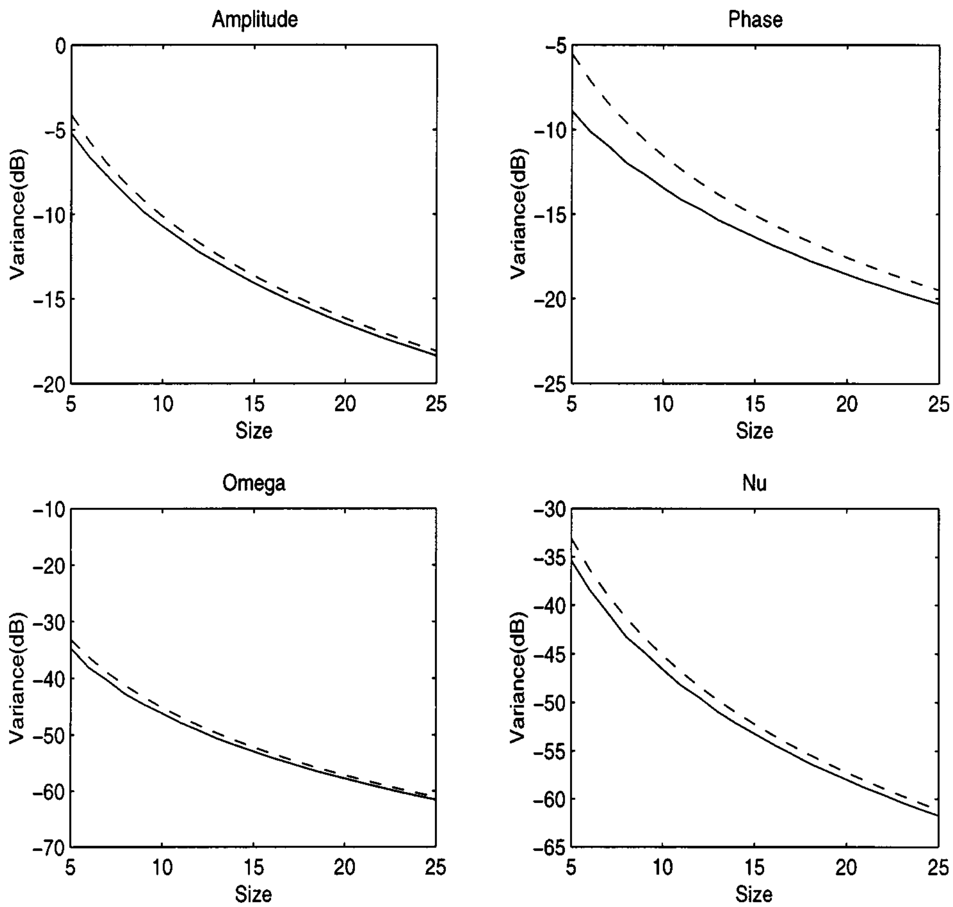

Fig. 4. The rescaled normalized asymptotic error variance of the LS estimate of the amplitude, phase, and spatial frequency as a function of data dimensions for SNR $= 0$ dB (dashed line), compared with the corresponding exact CRB (solid line).

of the observed field on the asymptotic performance of the LS estimator and on the CRB.

In this example, the harmonic component of the field comprises a single sinusoid such that $\omega = 0.1$, $\nu = 0.4$, and $\alpha = 5$. The purely indeterministic component is the same as in the first example. To evaluate the functional dependence of the LS estimator rescaled normalized asymptotic error variance, and of the corresponding CRB, on the dimensions of the observed field we set $N = M$ and let both N and M assume values from 5 to 20. The results of evaluating the rescaled normalized asymptotic error variance of the amplitude, phase, and spatial frequency estimates, and the corresponding exact CRB, as a function of the field dimensions are depicted in Figs. 3 and 4 for SNR values of -10 and 0 dB, respectively. The results indicate that the rescaled normalized asymptotic error variance of the LS estimator of each of the sinusoids parameters is nearly identical to the corresponding exact CRB, even for modest data dimension and relatively low SNR values.

VI. Conclusion

We have investigated the problem of LS estimation of the parameters of 2-D sinusoids observed in colored noise. Making no assumptions about the specific type of the probability distribution of the observed field, it is shown that if the disturbance field satisfies a combination of conditions comprised of a strong mixing condition and a condition on the order of its uniformly bounded moments, the normalized estimation error of the LS estimator is consistent asymptotically normal with zero mean and a normalized asymptotic covariance matrix for which a simple expression is derived. It is concluded that there is a tradeoff between the order of the disturbance moments that need to be uniformly bounded and the mixing rate of the disturbance field. It is also shown that the LS estimator of the parameters of the 2-D sinusoids is asymptotically unbiased.

The normalized asymptotic covariance matrix is block diagonal where each block corresponds to the parameters of a different sinusoidal component. Moreover, for each sinusoid, the matrix element that expresses the normalized asymptotic error variance in estimating the amplitude parameter of that sinusoid is decoupled and independent of all other model parameters. It is a function only of the colored noise spectral density at the sinusoid's frequency.

Assuming further that the colored noise field is Gaussian, the LS estimator of the sinusoidal components is shown to be asymptotically efficient. It is also demonstrated that the normalized asymptotic error covariance of the LS estimator is rapidly approached (in terms of both SNR and data dimensions) by the exact, finite-sample CRB, when the CRB matrix is normalized by the same normalization matrix employed to normalize the LS estimator covariance matrix.

APPENDIX A

We next show that the sequence $\boldsymbol{R}_{k,l}$ is nonnegative definite. Indeed, let $\boldsymbol{\alpha}$ be an arbitrary real-valued vector $\boldsymbol{\alpha} = [\boldsymbol{\alpha}_1^T, \ldots, \boldsymbol{\alpha}_P^T]^T$ where $\boldsymbol{\alpha}_l = [\alpha_1^l, \ldots, \alpha_4^l]^T$ and $l = 1, 2, \ldots, P$. Consider now the quadratic form

$$\sigma_{\delta-\mu,\,\varepsilon-\eta} = \boldsymbol{\alpha}^T \boldsymbol{R}_{\delta-\mu,\,\varepsilon-\eta} \boldsymbol{\alpha} = \frac{1}{2} \Re \left\{ \sum_{l=1}^{P} e^{i[\omega_l(\delta-\mu)+\nu_l(\varepsilon-\eta)]} \boldsymbol{\alpha}_l^T \boldsymbol{B}_l \boldsymbol{\alpha}_l \right\}. \quad (37)$$

Since all principle minors of $\boldsymbol{B}_p$ are nonnegative, $\boldsymbol{B}_p$ is positive semidefinite (see, e.g., [24, p. 405]). Let also $\boldsymbol{k}$ be any m-vector where m is arbitrary and let $r_1, r_2, \ldots, r_m$ and $s_1, s_2, \ldots, s_m$ be some positive integers. Then

$$\sum_{\mu,\,\varepsilon=1}^{m} k_\mu \sigma_{r_\mu - r_\varepsilon,\, s_\mu - s_\varepsilon} k_\varepsilon = \frac{1}{2} \Re \left\{ \sum_{l=1}^{P} \boldsymbol{\alpha}_l^T \boldsymbol{B}_l \boldsymbol{\alpha}_l \left| \sum_{\mu=1}^{m} k_\mu e^{i(\omega_l r_\mu + \nu_l s_\mu)} \right|^2 \right\} \ge 0 \quad (38)$$

where the last inequality results from the positive semidefiniteness of the blocks $\boldsymbol{B}_l$.

APPENDIX B

Lemma 1: Let $\{a_n\}$ be a positive monotone sequence tending to infinity, and let $\{X_n\}$ be a sequence of random variables such that $\{a_n X_n\}$ converges in distribution to an integrable random variable X, then $\{X_n\}$ converges in probability to zero.

Proof: Fix $\delta > 0$. The set of continuity points of the distributions of X, $\{a_n X_n\}$ is a dense set (as its complement is a countable union of countable sets). Hence, by Chebyshev inequality, for any $\epsilon > 0$ we can choose $C > 0$ such that $P(\{|X| \ge C\}) < \epsilon$, where $\pm C$ are continuity points of all the above distributions. For that C, there exists N such that for any $n \ge N$, $\delta > \frac{C}{a_n}$; hence, $\{|X_n| \ge \delta\} \subset \{a_n |X_n| \ge C\}$. By the continuity of the distribution functions at $\pm C$

$$\limsup_n P(\{|X_n| \ge \delta\}) \le \lim_n P(\{a_n |X_n| \ge C\}) = P(\{|X| \ge C\}) < \epsilon. \qquad \square$$

REFERENCES

[1] J. M. Francos, A. Z. Meiri, and B. Porat, "A Wold-like decomposition of 2-D discrete homogeneous random fields," *Ann. Appl. Probab.*, vol. 5, pp. 248–260, 1995.

[2] J. M. Francos, "Bounds on the accuracy of estimating the parameters of discrete homogeneous random fields with mixed spectral distributions," *IEEE Trans. Inform. Theory*, vol. 43, pp. 908–922, May 1997.

[3] P. Whittle, "On stationary processes in the plane," *Biometrika*, vol. 41, pp. 434–449, 1954.

[4] A. J. Isaksson, "Analysis of identified 2-D noncausal models," *IEEE Trans. Inform. Theory*, vol. 39, pp. 525–534, Mar. 1993.

[5] H. Helson and D. Lowdenslager, "Prediction theory and Fourier series in several variables," *Acta Math.*, vol. 99, pp. 165–202, 1958.

[6] ——, "Prediction theory and Fourier series in several variables," *Acta Math.*, vol. 106, pp. 175–213, 1962.

[7] J. M. Francos, A. Narasimhan, and J. W. Woods, "Maximum likelihood parameter estimation of discrete homogeneous random fields with mixed spectral distributions," *IEEE Trans. Signal Processing*, vol. 44, pp. 1242–1255, May 1996.

[8] A. M. Walker, "On the estimation of a harmonic component in a time series with stationary independent residuals," *Biometrika*, vol. 58, pp. 21–36, 1971.

[9] C. R. Rao, L. Zhao, and B. Zhou, "Maximum likelihood estimation of 2-D superimposed exponential signals," *IEEE Trans. Signal Processing*, vol. 42, pp. 1795–1802, July 1994.

[10] D. Kundu and A. Mitra, "Asymptotic properties of the least squares estimates of 2-D exponential signals," *Multidim. Syst. Signal Processing*, vol. 7, pp. 135–150, 1996.

[11] S. W. Lang and J. H. McClellan, "The extension of Pisarenko's method to multiple dimensions," in *Proc. Int. Conf. Acoustics, Speech, and Signal Processing*, 1982, pp. 125–128.

[12] R. Kumaresan and D. W. Tufts, "A two-dimensional technique for frequency-wavenumber estimation," *Proc. IEEE*, vol. 69, pp. 1515–1517, Nov. 1981.

[13] Y. Hua, "Estimating two-dimensional frequencies by matrix enhancement and matrix pencil," *IEEE Trans. Signal Processing*, vol. 40, pp. 2267–2280, Sept. 1992.

[14] H. Yang and Y. Hua, "Statistical decomposition of an eigendecomposition based method for 2-D frequency estimation," *Automatica*, vol. 30, pp. 157–168, 1994.

[15] S. Kay and R. Nekovei, "An efficient two-dimensional frequency estimator," *IEEE Trans. Signal Processing*, vol. 38, pp. 1807–1809, Oct. 1990.

[16] U. Grenander and M. Rosenblatt, *Statistical Analysis of Stationary Time Series*. New York: Wiley, 1957.

[17] P. Stoica and A. Nehorai, "Statistical analysis of two nonlinear least-squares estimators of sine-wave parameter in the colored-noise case," *Circuits Syst. Signal Processing*, vol. 8, no. 1, pp. 3–15, 1989.

[18] J. Li, P. Stoica, and D. Zheng, "An efficient algorithm for two-dimensional frequency estimation," *Multidim. Syst. Signal Processing*, vol. 7, pp. 151–178, 1996.

[19] J. Chun and N. K. Bose, "Parameter estimation via Signal-Selectivity of Signal-Subspace (PESS) and its application to 2-D wavenumber estimation," *Dig. Signal Processing*, vol. 5, pp. 58–76, 1995.

[20] J. Li and P. Stoica, "Efficient mixed-spectrum estimation with applications to target feature extraction," *IEEE Trans. Signal Processing*, vol. 44, pp. 281–295, Feb. 1996.

[21] J. Li, P. Stoica, and D. Zheng, "One-dimensional MODE algorithm for two-dimensional frequency estimation," *Multidim. Syst. Signal Processing*, vol. 8, pp. 449–468, 1997.

[22] P. Stoica, A. Jakobsson, and J. Li, "Cisoid parameter estimation in the colored noise case: Asymptotic Cramer–Rao bound, maximum likelihood and nonlinear least-squares," *IEEE Trans. Signal Processing*, vol. 45, pp. 2048–2059, Aug. 1997.

[23] A. Mitra and P. Stoica, "The asymptotic Cramer–Rao bound for 2-D superimposed exponential signals," manuscript, submitted for publication.

[24] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1990.

[25] J. M. Francos, A. Narasimhan, and J. W. Woods, "Maximum likelihood parameter estimation of textures using a Wold decomposition based nodel," *IEEE Trans. Image Processing*, vol. 4, pp. 1655–1666, Dec. 1995.

[26] J. Ward, "Space–time adaptive processing for airborne radar," Lincoln Lab., MIT, Lexington, MA, Tech. Rep. 1015, 1994.

[27] M. Rosenblatt, *Markov Processes, Structure and Asymptotic Behavior*. Berlin, Germany: Springer, 1971.

[28] N. N. Leonenko, "Estimates of linear regression coefficients on a homogeneous random field," *Ukrain. Mat. Zh.*, vol. 30, pp. 749–756. English translation in *Ukrainian Math. J.*, vol. 30, 1978.

[29] A. V. Ivanov and N. N. Leonenko, *Statistical Analysis of Random Fields*. Norwell, MA: Kluwer Academic, 1989.

[30] G. Cohen and J. M. Francos, "Least squares estimation of two-dimensional regression models in colored noise," *J. Multiv. Anal.*, to be published.

[31] M. B. Priestley, *Spectral Analysis and Time Series*. New York: Academic, 1981.

[32] E. J. Hannan, *Multiple Time Series*. New York: Wiley, 1970.

[33] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*. New York: Academic, 1980.

[34] H. L. Royden, *Real Analysis*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1988.

[35] K. L. Chung, *A Course in Probability Theory*. New York: Academic, 1974.

[36] B. Porat, *Digital Processing of Random Signals*. Englewood Cliffs, NJ: Prentice-Hall, 1994.