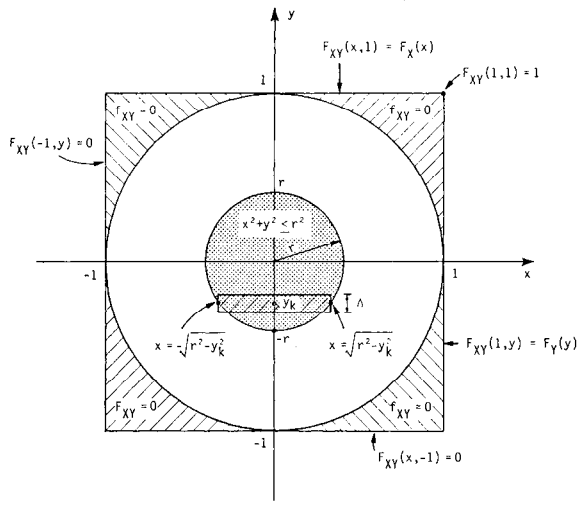

Fig. 5. Illustrations of regions of interest for $F_{XY}(x, y)$ and $f_{XY}(x, y)$.

where Δ is the grid spacing. Then

$$\Pr\{x^2 + y^2 \le r^2\} \approx \sum_{k=0}^{K} \left[F_{XY}\left(\sqrt{r^2 - y_k^2}, y_k + \frac{\Delta}{2}\right) - F_{XY}\left(\sqrt{r^2 - y_k^2}, y_k - \frac{\Delta}{2}\right) - F_{XY}\left(-\sqrt{r^2 - y_k^2}, y_k + \frac{\Delta}{2}\right) + F_{XY}\left(-\sqrt{r^2 - y_k^2}, y_k - \frac{\Delta}{2}\right)\right], \quad \text{(A.16)}$$

where $y_0 = -r + \Delta/2$, and $y_K = r - \Delta/2$. Note that the support of the joint pdf.

$$f_{XY}(x, y) \triangleq \frac{\partial^2}{\partial x\, \partial y} F_{XY}(x, y)$$

corresponds to the area enclosed by the unit circle, i.e., $f_{XY}(x, y) \equiv 0$ outside this circle. For the H_0 case, we may take advantage of symmetry to reduce (A.16) to

$$P_{fa} \approx \sum_{k=0}^{(K/2)-1} \left[F_{XY}\left(-\sqrt{r^2 - y_k^2}, y_k + \frac{\Delta}{2}\right) - F_{XY}\left(-\sqrt{r^2 - y_k^2}, y_k - \frac{\Delta}{2}\right)\right] + 4F_{XY}(0, -r). \quad \text{(A.17)}$$

References

[1] A. D. Whalen, *Detection of Signals in Noise*. New York: Academic, 1971.

[2] T. W. Anderson, *The Statistical Analysis of Time Series*. New York: Wiley, 1971.

[3] I. S. Reed, "On a moment theorem for complex Gaussian processes," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 194–195, Apr. 1962.

[4] G. H. Robertson, *Normal Error Model Analysis Charts*. Mountainside, NJ: Xerox Reproduction Center, May 1977.

[5] B. O. Steenson, "Detection performance of a mean-level threshold," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-4, pp. 529–534, July 1968.

[6] J. T. Rickard and G. M. Dillard, "Adaptive detection algorithms for multiple-target situations," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-13, pp. 338–343, July 1977.

A Model-Based Approach for Estimation of Two-Dimensional Maximum Entropy Power Spectra

GOVIND SHARMA AND RAMA CHELLAPPA, SENIOR MEMBER, IEEE

***Abstract*—A stochastic model-based approach is presented for estimation of the two-dimensional maximum entropy power spectrum (MEPS) from given finite uniform array data. The method consists of fitting an appropriate two-dimensional noncausal Gaussian–Markov random field (GMRF) model to the given data using the maximum likelihood (ML) technique for parameter estimation. The nonlinear criterion function used for ML estimation is similar in structure to the function arising in the deterministic approach of Lang and McClellan. The model-based approach provides new insights into the two-dimensional MEPS estimation problem. For example, using the asymptotic normality of ML estimates, we derive simultaneous confidence bands for the estimated MEPS. It turns out that when the true correlations are generated by a noncausal GMRF model, the two-dimensional MEPS can be obtained by solving linear equations. This approach also suggests techniques for realizing two-dimensional GMRF models from the given correlation data. Several numerical examples are given to illustrate the usefulness of the approach.**

Manuscript received February 14, 1984; revised July 25, 1984. This work was partially supported by the National Science Foundation under Grant ECS-82-04181, by the Office of Naval Research under contract ONR-N00014-81-K-0191, and by the USC Faculty Research Innovation fund. The paper was presented in part at the International Conference on Acoustics, Speech and Signal Processing, San Diego, CA, March 1984.

Rama Chellappa is with the Department of Electrical Engineering—Systems, University of Southern California, Los Angeles, CA 90089, USA.

Govind Sharma is with the Department of Electrical Engineering, Louisiana State University, Baton Rouge, LA, 70803, USA.

I. Introduction

TWO-DIMENSIONAL spectral estimation is of interest in frequency wavenumber analysis [2], image restoration [3], filtering of radar images [4], and texture

0018-9448/85/0100-0090$01.00 ©1985 IEEE

classification [5]. There are various methods for estimating power spectra in two dimensions, some of which are direct generalizations of one-dimensional techniques. But not all one-dimensional methods can be directly generalized to two dimensions. This is because in contrast to one dimension, where the notion of causality is appropriate, in two-dimensional methods such a notion is restrictive. Furthermore, the two-dimensional power spectrum does not always factorize as it does in one dimension. One of the popular methods of spectrum analysis in one and two dimensions is maximum entropy power spectrum (MEPS) analysis. In this method the spectrum is obtained by maximizing an entropy functional subject to the condition that the inverse Fourier transform of the spectrum exactly matches the given correlations. It turns out that in one dimension the maximum entropy spectrum is the same as the spectrum obtained by assuming an autoregressive (AR) model for the underlying time series data [6]–[8]. We are very often given the raw time series data and not the correlations. In such a case we estimate the MEPS by fitting an AR model to given data. The AR parameter estimates are obtained by calculating autocorrelation estimates and then using Yule–Walker (YW) equations. One can also get the AR parameters by minimizing forward prediction error or forward plus backward prediction error [7]. These estimates differ, depending on the estimates of autocorrelations used. For an admissible solution to exist, the estimated autocorrelations must be extendible. Extendibility requires that the Toeplitz matrix formed by the correlations should be positive-definite. Extendibility is also guaranteed if the magnitudes of the reflection coefficients are less than or equal to one in magnitude. Burg's technique [9] estimates the reflection coefficients such that their magnitude is less than or equal to one. For these estimates of MEPS the inverse Fourier transform of the estimated spectrum matches the estimated correlations.

Now consider the two-dimensional MEPS analysis. Most related literature concerns the deterministic approach, where the true correlations are given [1], [10]–[14]. In one dimension the MEPS algorithm involves the inversion of a Toeplitz matrix. But in two dimensions the algorithms involve minimization of a nonlinear criterion function. Since polynomials in two variables do not always factor into causal and anticausal polynomials, the minimization problem cannot always be converted to a linear prediction problem. The various algorithms are different techniques for maximizing entropy or for solving an equivalent dual problem. The discussion regarding the convergence of these various algorithms and the related question of extendibility can be found in [1] and [14]. The question of extendibility is very important in two-dimensional methods. Even if a given correlation function is positive-definite on a rectangle, it may not be extendible to the whole plane [15].

It has been observed [10] that the two-dimensional MEPS has the same structure as the Gaussian–Markov random field (GMRF) spectrum. Thus, one feasible method for estimating two-dimensional MEPS from random field data is to obtain the GMRF model parameters such that the correlation matching property is satisfied. No such method of estimating model parameters from the data was given in [10]. Another method would be to use estimates of correlations instead of the true correlations and to use the algorithms described in [1] and [10]–[14]. But then one has to decide on the particular sample correlations to be used, and the estimates must be extendible. Even if the data are from a stationary random process with bounded spectrum, the sample correlation estimates may not be extendible. A second question arises as to how many correlations should be used. The correlations with larger lags are less reliable, hence one may get better results by using fewer correlations. Questions of this kind can be satisfactorily answered if one uses the modeling approach.

As mentioned earlier, the similarity between the GMRF model spectrum and the MEPS spectrum was noted in [10]. Based on this similarity, an asymptotic method for obtaining a model based estimate of two-dimensional MEPS was given in [16]. There it was assumed that the given random field data are doubly periodic (the toroidal lattice assumption). Under this assumption it was shown that the spectrum estimates obtained by using the maximum likelihood (ML) estimates of the model parameters have a correlation matching property. That is, the (periodic) sample correlations in the neighborhood of $(0, 0)$ are in perfect agreement with the inverse discrete Fourier transform of the estimated spectrum. Because of the toroidal lattice assumption, the GMRF spectral estimate in [16] is only asymptotically equivalent to the MEPS estimate.

In this paper, we present a method for estimating the two-dimensional MEPS from a given finite array data. The approach consists of fitting an appropriate GMRF model to the given data, using the ML method of parameter estimation.

The Gaussian-Markov Random Field Model

In the GMRF model [17]–[19] the observations $y(s)$ obey a Markovian property

$$p\{y(s)|\text{all } y(r),\ r \neq s\} = p\{y(s)|\text{all } y(s+r),\ r \in N\},$$

where N is a neighbor set that consists of pairs of integers excluding $(0, 0)$. The zero-mean observations $\{y(s)\}$ for the GMRF model satisfy the following difference equation:

$$y(s) = \sum_{r \in N} \theta_r y(s+r) + e(s), \tag{1.1}$$

where $\{\theta_r, r \in N\}$ are GMRF model parameters and $\{e(s)\}$ is a correlated two-dimensional noise sequence with the correlation structure

$$E[e(s)e(r)] = \begin{cases} -\theta_{r-s}\nu, & (r-s) \in N \\ \nu, & r = s \\ 0 & \text{otherwise.} \end{cases}$$

The power spectrum associated with (1.1) is

$$S(\lambda) = \frac{\nu}{1 - \sum_{r \in N} \theta_r \exp(-i\lambda' r)}. \tag{1.2}$$

Since the spectrum must be real and positive, we should have $\theta_r = \theta_{-r}$. Thus, we can use the following representation for the GMRF model

$$y(s) = \sum_{r \in N_s} \theta_r [y(s+r) + y(s-r)] + e(s),$$

where N_s is an asymmetric neighbor set such that if $r \in N_s$, then $-r \notin N_s$ and $N = \{r: r \in N_s \text{ or } -r \in N_s\}$.

The functional form of (1.2) is the same as the maximum entropy power spectrum. Now we need parameter estimates from the given data such that the correlation matching property is satisfied. That is, the inverse Fourier transform of (1.2) matches the "sample correlations" when θ_r is replaced by the estimate $\hat{\theta}_r$ in (1.2). The least squares (LS) estimates of the parameters do not have this property. The ML estimates have the property that the sample correlations are in perfect agreement with the inverse Fourier transform of the GMRF spectrum estimate obtained by substituting the ML parameters into (1.2) [20].

In any estimation problem it is not enough to provide only the point estimates. It is also desirable to obtain some bounds on the estimates. Since the spectrum is a function of frequency, one needs to give bands within which the spectrum lies. Using the asymptotic normality of the ML estimates, we derive simultaneous confidence bands for the estimated spectrum.

In addition to the parameter estimation, the modeling approach requires the choice of appropriate model order. An asymptotically consistent and transitive decision rule for choosing an appropriate neighbor set of the GMRF model is given in Section III.

The GMRF models have been used for two-dimensional spectral analysis in [21]. In this work the model parameters are obtained by solving a set of linear two-dimensional YW equations. In general, the parameters so obtained do not satisfy the correlation matching property. Also these parameters are sometimes inadmissible in the sense that the spectrum determined by these parameters might be negative for some frequencies. Furthermore, the choice of neighbor set is arbitrary. We show that when the given correlations are generated by a GMRF model whose neighbor set is known, the parameters satisfy the correlation matching property. Thus it is only for a certain class of problems that the two-dimensional MEPS solution can be obtained by solving a linear problem.

The contributions of this paper can be summarized as follows. 1) A systematic procedure for estimating the two-dimensional MEPS from raw data is given by using GMRF models. Thus the method provides some theoretical justification for the known approach of using the sample correlations for the MEPS problem. 2) Simultaneous confidence bands are derived for the estimated spectrum. 3) A linear solution applicable under special conditions to the two-dimensional MEPS problem from true correlations is given. 4) A systematic procedure for realizing two-dimensional GMRF models from the given true correlations is given.

The organization of the paper is as follows. In Section II a brief discussion of the MEPS analysis is given. In Section III it is shown that the ML estimates satisfy the sample correlation matching property. Using the asymptotic normality property of the ML estimates, simultaneous confidence bands are derived in the same section. The details of the model based algorithm, including a model selection procedure, are also given. In Section IV some simulation examples are presented. Simulations include examples of parameter estimation, spectrum estimation, and confidence bands. In Section V we discuss the special case where YW equations can be used for determining the two-dimensional MEPS. Examples illustrating this method are also found in Section V. In Section VI we discuss a general procedure for realizing two-dimensional GMRF models from given true correlations.

II. Maximum Entropy Power Spectrum Analysis

Let $R(r)$ be the autocorrelation function of homogeneous (stationary) field $\{y(s)\}$ where $r = (r_1, r_2)$ and $s = (s_1, s_2)$. Let A be a set of lattice points including $(0, 0)$ on which $R(r)$ is known. The maximum entropy power spectrum $S_{\text{MEPS}}(\lambda)$ is obtained by maximizing the entropy functional H defined as

$$H = \int \log S(\lambda)\, d\lambda \tag{2.1}$$

subject to the condition that

$$R(r) = \hat{R}(r) \qquad \text{for } r \in A, \tag{2.2}$$

where

$$\hat{R}(r) = \text{IFT}\{S(\lambda)\},$$

where IFT denotes inverse Fourier transformation. It can be shown that the MEPS estimate, if it exists, has the following form [9]

$$S_{\text{MEPS}}(\lambda) = \frac{\rho}{1 - \sum_{r \in N} a_r \exp(-i\lambda' r)}. \tag{2.3}$$

Thus the MEPS is obtained if parameters $\{\rho, a_r, r \in A\}$ are determined such that (2.2) is satisfied. In general, this is a nonlinear optimization problem. In the one-dimensional case, the denominator polynomial can be factored to convert the nonlinear problem into a linear one involving the inversion of a Toeplitz correlation matrix. However, in two dimensions the polynomial factorization is generally not possible, and the relation between $R(r)$ and a_r and ρ is highly nonlinear. Thus a_r and ρ are found by minimizing a nonlinear criterion function. Some iterative optimization procedures for obtaining the MEPS can be found in [1] and [10]–[14].

In one dimension it is known that the MEPS is the same as the spectrum obtained by assuming an AR model for the underlying data $\{y(s)\}$ [6]–[8]. Very often, we are given the data $\{y(s)\}$ and not the correlations $R(r)$. Thus one obtains an estimate of the MEPS by fitting an AR model to the data. We show that in two dimensions the MEPS is the same as the spectrum obtained by assuming a

GMRF model for the underlying field. When we are given data on a finite lattice, we can obtain an estimate of the MEPS by fitting a GMRF model to the given data.

III. A Model-Based Algorithm for Two-Dimensional MEPS Estimation

An estimate of the two-dimensional MEPS is obtained if, from the given random data, we estimate the model parameters such that when estimates are used in (1.2) the correlation matching constraint (2.2) is satisfied.

There are three methods for estimating the GMRF model parameters. These are the coding method [19], the LS method [22], and the ML method [20], [23]. The coding and LS estimates are easy to compute, but they are not efficient and do not have the sample correlation matching property. The ML estimate is asymptotically consistent and efficient.

A. Maximum Likelihood estimation of parameters

Let $\boldsymbol{y}$ and $\boldsymbol{e}$ be lexicographically ordered arrays of $y(s)$ and $e(s)$, respectively. Neglecting the boundary conditions, (2.4) for $M \times M$ lattice size can be written as

$$\boldsymbol{H}(\boldsymbol{\theta})\boldsymbol{y} = \boldsymbol{e}. \tag{3.1}$$

From (2.5) we have

$$E(\boldsymbol{e}\boldsymbol{e}^t) = \nu\boldsymbol{H}(\boldsymbol{\theta}) \quad \text{and} \quad E(\boldsymbol{e}) = 0, \tag{3.2}$$

where $E(\cdot)$ is the expectation operator and t denotes transposition. Now we have from (3.1) and (3.2) that $E(\boldsymbol{y}) = 0$ and

$$E(\boldsymbol{y}\boldsymbol{y}^t) = \nu\boldsymbol{H}^{-1}(\boldsymbol{\theta}), \tag{3.3}$$

since $\boldsymbol{H}(\boldsymbol{\theta})$ is a symmetric positive definite matrix. Since $\boldsymbol{y}$ is Gaussian, we can write the log likelihood function as

$$\log p(\boldsymbol{y}|\boldsymbol{\theta}, \nu) = \frac{1}{2}\log\det\boldsymbol{H}(\boldsymbol{\theta}) - \frac{M^2}{2}\log 2\pi\nu - \frac{1}{2\nu}\boldsymbol{y}^t\boldsymbol{H}(\boldsymbol{\theta})\boldsymbol{y}. \tag{3.4}$$

The quadratic form $\boldsymbol{y}^t\boldsymbol{H}(\boldsymbol{\theta})\boldsymbol{y}$ can be written as

$$\boldsymbol{y}^t\boldsymbol{H}(\boldsymbol{\theta})\boldsymbol{y} = M^2\left[C(0) - 2\sum_{r\in N_s}\theta_r C(r)\right],$$

where

$$C(r) = 1/M^2\sum y(s)y(s+r). \tag{3.5}$$

To further simplify (3.4) we use the following lemma.

Lemma: The term $1/M^2(\log\det\boldsymbol{H}(\boldsymbol{\theta}) + \log\nu)$ converges to

$$1/2\left((2\pi)^{-2}\int\log S(\lambda,\boldsymbol{\theta})\,d\lambda + \log 2\pi\right)$$

as $M \to \infty$ where $S(\lambda,\boldsymbol{\theta})$ is the spectrum in (1.2).

Proof: For n Gaussian random variables $y_1, y_2, \cdots, y_n$ the entropy is given by [24]

$$H = -\int p(y)\log\left(c^{2n}p(y)\right)dy, \tag{3.6}$$

where c is a constant that fixes the reference level. Thus for the probability density given in (3.4) we get

$$H = -1/2\log\det\boldsymbol{H}(\boldsymbol{\theta}) + M^2/2\log\nu. \tag{3.7}$$

Here we have taken c as $(2\pi)^{1/4}$ and $n = M^2$. The entropy rate is defined as

$$\lim_{M^2\to\infty} H/M^2,$$

which converges to [25]

$$1/2\left((2\pi)^{-2}\int\log S(\lambda,\boldsymbol{\theta})\,d\lambda + \log 2\pi\right).$$

Hence for large M we can write (3.4) as

$$-2M^{-2}\log p(\boldsymbol{y}|\boldsymbol{\theta},\nu) \cong (2\pi)^{-2}\int\log S(\lambda,\boldsymbol{\theta})\,d\lambda + 2\log 2\pi + 1/\nu\left[C(0) - 2\sum_{r\in N_s}\theta_r C(r)\right]. \tag{3.8}$$

A direct and mathematically rigorous proof of (3.8) can be found in [20]. Differentiating (3.8) and equating to zero gives

$$(2\pi)^{-2}\int\cos\lambda^t r S(\lambda,\boldsymbol{\theta})\,d\lambda = C(r), \qquad R \in N_s. \tag{3.9}$$

Equation (3.9) shows that the correlation of the fitted model will be equal to the sample correlation for lags r in N_s. It is interesting to note that (3.8) has the same form as the criterion function minimized for obtaining the two-dimensional MEPS in [26] when true correlations are replaced by sample correlations.

From (3.8) we can prove directly that the ML estimates of the parameters maximize the entropy without invoking the similarity of spectra in (2.3) and (1.2). The right-hand side of (3.8) can be written as

$$-2M^{-2}\log p(\boldsymbol{y}|\boldsymbol{\theta},\nu) = (2\pi)^{-2}\int[\log S(\lambda,\boldsymbol{\theta}) + I(\lambda)/S(\lambda,\boldsymbol{\theta})]\,d\lambda + 2\log 2\pi \tag{3.10}$$

where $I(\lambda) = C(0) + 2\Sigma_{r\in N_s}C(r)\cos\lambda^t r$, which can be interpreted as the likelihood function of independent exponential random variables. Hence from [27] it follows that parameters that minimize (3.8) also maximize the entropy. The likelihood function derived by Whittle [28] also has the form of (3.10). But the likelihood derived in [28] is for spatial autoregressive (SAR) processes. The spatial autoregressive process has the representation

$$y(s) = \sum_{r\in N}\theta_r y(s+r) + \omega(s),$$

where $\omega(s)$ is an independent identically distributed noise sequence. This model is different from the GMRF model. For example, if we have a GMRF model with the neighbor set $\{(1,0), (-1,0), (0,1), (0,-1)\}$, then there is no equiv-

alent SAR model. Although the entropy is maximized by the ML estimates of the parameters, the correlation matching property is not satisfied for all noncausal SAR models. This property will be satisfied only if the correlations correspond to an SAR model with the same neighbor set. If the spectrum factorizes into causal and anticausal functions, then one can use causal models. Spectrum estimation using causal models, which is similar to the one-dimensional case, is discussed in [29]. As we have pointed out earlier, because of the lack of spectral factorization we cannot always find a finite parameter causal model in the two-dimensional case. Thus we need more general noncausal models. The correlation matching property is always satisfied when the fitted model is a GMRF model.

B. Existence, Uniqueness, and Consistency of Estimates

When we are given a portion of a true correlation function, the question arises whether or not this correlation is extendible; i.e., does there exist a positive spectrum that will match these given correlations? It is well known that even if the given correlation function is positive-definite on a rectangle, it may not have a positive-definite extension to the entire plane [15]. Using (3.5), the quadratic form

$$Q = \sum\sum a_m C(r_m - r_n) a_n$$

can be written as

$$\begin{aligned} Q &= \sum\sum a_m 1/M^2 \sum y(s + r_m) y(s + r_n) a_n \\ &= 1/M^2 \sum \Big(\sum a_m y(s + r_m)\Big)^2 \\ &\geq 0. \end{aligned}$$

Hence $\{C(r)\}$ is positive-definite on the entire plane and is thus extendible. It is shown in [10] that if a correlation function $\{C(r)\}$ is positive-definite on the entire plane, then there exists a unique Markov random field $\{y\}$ with associated correlation function R_y such that $R_y = C(r)$ on a neighbor set N. Thus we see that a solution for (3.9) always exists and is unique.

By the ergodic theorem, the $C(k)$'s converge to the $R(k)$'s. Hence as $M \to \infty$, the estimated spectrum matches the true correlations exactly. Since the GMRF spectrum has the same functional form as the MEPS, we conclude that the MEPS is the same as the spectrum obtained by assuming a GMRF model for the underlying data. When the data are given instead of the correlations we get an estimate for which estimated correlations are matched exactly. Also we see that $\hat{\theta}$ converges to θ in probability as $M \to \infty$. The asymptotic error covariance matrix is given by J^{-1}, where the (i, j)-th element of J is given by

$$J_{i,j}(\theta) = \frac{1}{2M^2(2\pi)^2} \int \frac{\partial}{\partial \theta_i} \log S(\lambda, \theta) \frac{\partial}{\partial \theta_i} \log S(\lambda, \theta)\, d\lambda, \tag{3.11}$$

which is obtained by differentiating the log likelihood function twice. The asymptotic normality of the ML estimates can be used to determine confidence bands for power spectrum estimates.

C. Confidence Bands for Power Spectrum Estimates

Since the reciprocal of the power spectrum estimate is a linear combination of random variables whose joint distribution is asymptotically Gaussian, we can obtain confidence bands for the power spectrum estimate by using Scheffé's projection theorem [30, p. 407] of multivariate analysis. Let us define

$$\begin{aligned} h(\lambda) = \frac{1}{S(\lambda)} &= \frac{1}{\nu}\Big(1 - \sum_{r \in N_s} \theta_r \cos \lambda^t r\Big) \\ &= \gamma(0) + 2 \sum_{r \in N_s} \gamma(r) \cos \lambda^T r, \end{aligned} \tag{3.12}$$

where $\gamma(0) = 1/\nu$ and $\gamma(r) = -\theta_r/\nu$. Let the asymptotic covariance matrix for $M(\hat{\gamma} - \gamma)$ be $D(\theta)$. We can write (3.12) as $h(\lambda) = x^T(\lambda)\gamma$, where $x^T(\lambda) = [1, 2\cos \lambda^T r_1, \cdots, 2\cos \lambda^T r_p]$, $\gamma^T = [\gamma(0), \gamma(r_1), \cdots, \gamma(r_p)]$, and $\{r_1, r_2, \cdots, r_p\}$ is any ordering of the elements of N_s. The confidence bands are obtained from the following theorem.

Theorem: As $M \to \infty$ the probability is at least $(1 - \alpha)$ that simultaneously for all λ

$$\frac{1}{\hat{h}(\lambda) + f(\lambda)} \leq S(\lambda) \leq \frac{1}{\hat{h}(\lambda) - f(\lambda)},$$

where if $\hat{h}(\lambda) - f(\lambda) \leq 0$, we use infinity as the upper limit and $f(\lambda)$ is defined by

$$f^2(\lambda) = \frac{\chi^2_{\alpha, p+1} x^T(\lambda) D(\theta) x(\lambda)}{M^2},$$

and where $\chi^2_{\alpha, p+1}$ is the α critical value of the chi-squared distribution having $(p + 1)$ degrees of freedom.

Proof: The proof given here is very similar to the one-dimensional case given in [31]. Since $M(\hat{\gamma} - \gamma)$ is normal with covariance matrix $D(\theta)$, the quantity

$$M^2(\hat{\gamma} - \gamma)^T D^{-1}(\theta)(\hat{\gamma} - \gamma)$$

is an χ^2 variable with $(p + 1)$ degrees of freedom. The asymptotic probability of γ lying within an ellipsoid around $\hat{\gamma}$ is $(1 - \alpha)$ if

$$M^2(\hat{\gamma} - \gamma)^T D^{-1}(\theta)(\hat{\gamma} - \gamma) \leq \chi^2_{\alpha, p+1},$$

where $\chi^2_{\alpha, p+1}$ is as defined above. Using Scheffe's theorem [30], we have that γ is in this ellipsoid if and only if

$$|v^T(\hat{\gamma} - \gamma)| \leq \left(\frac{\chi^2_{\alpha, p+1} v^T D(\theta) v}{M^2}\right)^{1/2}$$

for all $(p + 1)$-dimensional vectors v. In particular we consider $v = x(\lambda)$, which gives

$$|x^t(\lambda)(\hat{\gamma} - \gamma)| \leq \left(\frac{\chi^2_{\alpha, p+1} x^T(\lambda) D(\theta) x(\lambda)}{M^2}\right)^{1/2} = f(\lambda)$$

or $|\hat{h}(\lambda) - h(\lambda)| \leq f(\lambda)$, the reciprocal of which gives the desired result

$$\frac{1}{\hat{h}(\lambda) + f(\lambda)} \leq S(\lambda) \leq \frac{1}{\hat{h}(\lambda) - f(\lambda)}.$$

In this proof, we have assumed that the true power spectrum corresponds to a GMRF and that the neighbor set is known exactly. The proof also holds if we use a neighbor set selected by a consistent neighbor selection rule. Such a neighbor selection rule is considered in the next section.

D. Model-Order Selection Rule

In the above discussion and in parameter estimation we have assumed that the neighbor set N is known. In deterministic algorithms one does not have to address the neighbor selection problem. But in any model-based method the correct choice of N is vital in obtaining meaningful estimates of the power spectrum of the process. Some examples illustrating this in one dimension are given in [7].

The model selection problem comes under the category of multiple decision problems. A method that is well suited for this problem is to compute a test statistic for different models and to choose the model corresponding to the minimum value. Akaike's information criterion (AIC) and the Bayes method are two such procedures. The expression for the AIC text statistic can be found in [20]. In general, the AIC method gives a transitive decision rule but is not consistent even for one-dimensional AR models [32]. The Bayes method gives transitive and asymptotically consistent decision rules. Following [22] we can write the Bayes rule as follows.

Suppose we have k sets $N_{s1}, \cdots N_{sk}$ of neighbors containing $m_1, \cdots m_k$ members, respectively. Corresponding to each N_{si} we write the GMRF model as

$$y(s) = \sum_{r \in N_{si}} \theta_{ir}[y(s+r) + y(s-r)] + \sqrt{\nu_i}\, e(s).$$

The decision rule, then, is to choose neighbor set N_{si^*} if

$$i^* = \text{Argument} \min_n \{g_n\},$$

where

$$g_n = (2\pi)^{-2} \int \log S(\lambda, \hat{\nu}_n, \hat{\theta}_n)\, d\lambda + (m_n/M^2) \log M^2, \tag{3.13}$$

and

$$\hat{\theta}_n = \text{ML estimate of } \theta \text{ for neighbor set } N_{sn}$$

$$\hat{\nu}_n = \text{ML estimate of } \nu \text{ for neighbor set } N_{sn}.$$

Simulation experiments using this and other decision rules are reproted in [33]. The simulations show that the performance of the AIC and Bayes rules are comparable, but the AIC method overestimates the neighbor set size most of the time.

The main difficulty in estimating GMRF model parameters is the calculation of $\det H(\theta)$ in (3.4) or the integral in (3.8). By assuming specific boundary conditions, $\det H(\theta)$ can be calculated very efficiently. One such boundary condition is to assume a doubly periodic extension of the given data. However, the special boundary conditions violate the maximum entropy assumption, so this procedure yields the MEPS only asymptotically. Parameter estimation and MEPS estimation for this case are discussed in [16].

IV. Simulation Results

In this section we give some simulation results using synthetic data. The parameters of the GMRF models were estimated by minimizing the negative of the log likelihood function (3.8) with respect to the parameters θ and ν. The log likelihood function is a complicated function of the parameters and requires numerical optimization procedures. We have used the Fletcher–Powell minimization procedure to find the minimum of (3.8). Synthetic data corresponding to given GMRF model parameters were generated using the scheme given in [34]. Simulation results are summarized below.

Experiment 1 (Spectrum Estimation using a GMRF Model)

A realization of 64×64 data points corresponding to the model parameters given in Table I was obtained. The spectrum corresponding to this model is shown in Fig. 1. The model parameters were estimated by minimizing the negative of the log likelihood function in (3.8). The estimated ML parameters are also given in Table I. Since the estimated parameters are numerically close to the actual parameters, the estimated spectrum should be close to the actual spectrum. The estimated spectrum is shown in Fig. 2.

Experiment 2 (Parameter Estimation in a GMRF Model)

In this experiment the parameter estimation is studied in more detail. The experiments were carried out using first- and second-order models. For the first-order model, $N_s = \{(1,0), (0,1)\}$. The model parameters were $\theta_{(1,0)} = 0.2340$, $\theta_{(0,1)} = 0.1011$, and $\nu = 1.0$. Thirty realizations of the GMRF model with these parameters were obtained. The lattice size used was 32×32. The sample means and variances of the parameter estimates obtained are

$$\langle \theta_{(1,0)} \rangle = 0.230756, \qquad \langle \theta_{(0,1)} \rangle = 0.091894,$$

$$\langle \nu \rangle = 1.009173$$

$$\text{var}(\theta_{(1,0)}) = 0.000505, \qquad \text{var}(\theta_{(0,1)}) = 0.000371,$$

$$\text{var}(\nu) = 0.002285.$$

TABLE I
PARAMETER VALUES IN EXPERIMENT 1

(k,l)	1,0	0,1	1,1	1,−1	0,2	2,0	−1,2	−2,2	1,2	2,1	2,2	2,−1
$\theta_{(k,l)}$	0.138	0.063	−0.067	0.154	0.009	−0.084	−0.040	0.075	0.079	−0.038	−0.023	0.156
$\theta_{(k,l)}$	0.133	0.064	−0.068	0.147	0.003	−0.083	−0.045	0.067	0.071	−0.044	−0.021	0.142

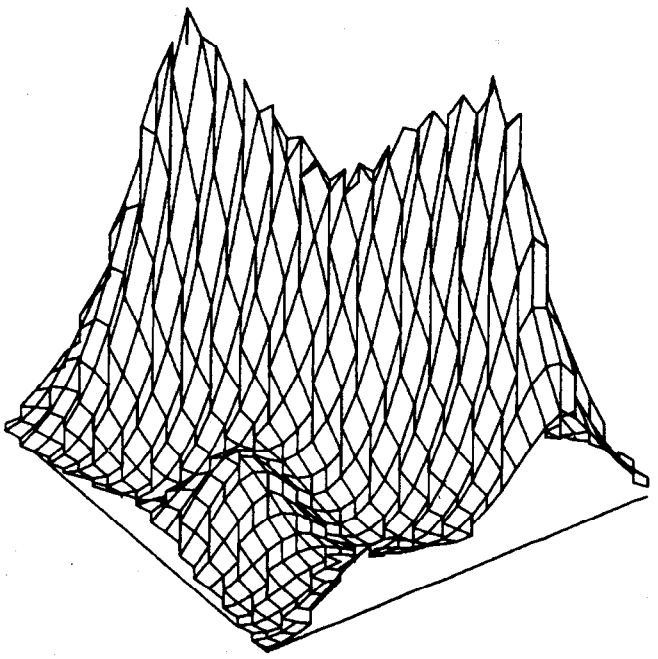

Fig. 1. Original spectrum in Experiment 1.

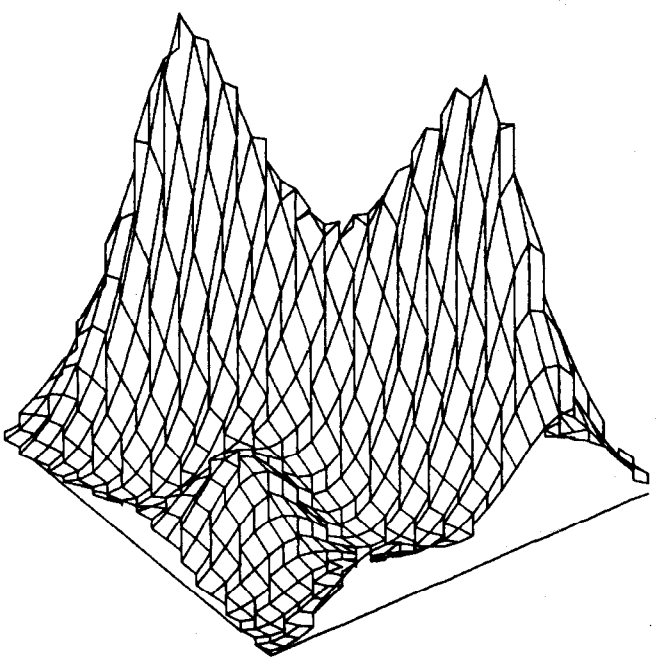

Fig. 2. Estimated spectrum in Experiment 1.

The theoretical asymptotic error covariance matrix from (3.11) is

$$\begin{bmatrix} 0.6067330^{-3} & -0.1697421^{-3} & 0.3124391^{-3} \\ -0.1697421^{-3} & 0.6365225^{-3} & 0.7520876^{-4} \\ 0.3124391^{-3} & 0.7520876^{-4} & 0.2123552^{-2} \end{bmatrix}.$$

The mean values of the estimates are very near the actual parameters, and the sample variances are also comparable to the theoretical values, although the number of realization is only 30. For the second-order model we have

$$N_s = \{(1,0), (0,1), (1,1), (1,-1)\}.$$

The parameters for this model were

$$\theta_{(1,0)} = 0.1950, \quad \theta_{(0,1)} = 0.0500, \quad \theta_{(1,1)} = -0.1350,$$
$$\theta_{(1,-1)} = 0.1010, \quad \nu = 1.0000.$$

Thirty-three simulation runs were made. The means and variances of the parameters were

$$\langle\theta_{(1,0)}\rangle = 0.201745, \qquad \langle\theta_{(0,1)}\rangle = 0.050060,$$
$$\langle\theta_{(1,1)}\rangle = -0.139635, \qquad \langle\theta_{(1,-1)}\rangle = 0.090640,$$
$$\langle\nu\rangle = 0.984636$$
$$\operatorname{var}(\theta_{(1,0)}) = 0.000610, \qquad \operatorname{var}(\theta_{(0,1)}) = 0.000752,$$
$$\operatorname{var}(\theta_{(1,1)}) = 0.000337, \qquad \operatorname{var}(\theta_{(1,-1)}) = 0.000756,$$
$$\operatorname{var}(\nu) = 0.001627$$

The theoretical asymptotic error covariance matrix in this case is

$$\begin{bmatrix} 0.64512^{-3} & 0.56589^{-4} & 0.19553^{-4} & -0.12266^{-3} & 0.29527^{-3} \\ 0.56589^{-4} & 0.94099^{-3} & -0.32675^{-3} & -0.36869^{-3} & 0.11420^{-3} \\ 0.19553^{-4} & -0.32675^{-3} & 0.71194^{-3} & 0.14801^{-3} & -0.21318^{-3} \\ -0.12266^{-3} & -0.36869^{-3} & 0.14801^{-3} & 0.74347^{-3} & 0.73846^{-4} \\ 0.29527^{-3} & 0.11420^{-3} & -0.21318^{-3} & 0.73846^{-4} & 0.22702^{-2} \end{bmatrix}.$$

Experiment 3 (Confidence Bands for a GMRF Spectrum) The confidence bands derived in Section III-C assume the asymptotic normality of the error in the parameter estimation. To investigate the finite parameter behavior of this bound, we considered spectrum estimates corresponding to the parameters in the preceding experiment. The actual spectrum and estimated spectrum were calculated at 32×32 equidistant points in the $[0, 2\pi] \times [0, 2\pi]$ frequency region. Confidence bands for 95 percent confidence level were also calculated at these 32×32 points. Then we counted the number of points at which the actual spectrum was out of the calculated spectral bands at these 32×32 points. In 29 out of 30 simulation runs for the first-order GMRF model this count was zero. For the second-order GMRF model this out-of-band count was zero in 32 out of 33 runs. These numbers are within the 5 percent error allowed by the bound.

V. A Linear Solution to the Two-Dimensional MEPS Problem

In general the determination of the MEPS is a nonlinear problem. But in a very special case the two-dimensional MEPS can be obtained as the solution to a linear problem. Suppose that the given true correlations correspond to a GMRF model. Then we have

$$y(s) = \sum_{r \in N_s} \theta_r [y(s+r) + y(s-r)] + e(s). \quad (5.1)$$

Multiplying both sides by $y(s+t)$ and taking expectations, we get

$$R(t) = \sum_{r \in N_s} \theta_r [R(t-r) + R(t+r)] + \nu\delta(t). \quad (5.2)$$

If we take $t = (0,0)$ and $t \in N_s$, we get

$$\nu = R(0) - 2 \sum_{r \in N_s} \theta_r R(r) \quad (5.3)$$

and

$$A\theta = b, \quad (5.4)$$

where the matrix A and vector b can be easily identified from (5.2). Hence the solution θ is given by $\theta = A^{-1}b$. This solution was suggested in [35]. We prove that this θ is the same as the parameter vector obtained by solving the nonlinear MEPS problem. Suppose we are given $R(r)$ for $r \in N_s$, and some more lags. Using $R(r)$, $r \in N_s$, we can find θ' and ν' by putting $R(r)$ instead of $C(r)$ in (3.8); this satisfies the correlation matching property on N_s. Since there is a unique GMRF model that satisfies the correlation matching property [10] and the correlation corresponding to this GMRF must satisfy (5.2), we conclude that $\theta' = \theta$ and $\nu' = \nu$.

Thus in a special case when the correlations correspond to a GMRF model and the model order is also known exactly, the MEPS can be obtained by solving linear equations. We note that the number of correlations involved in calculating the parameters are more than the number of parameters. This method is not useful in general, as the correlation matching property will not be satisfied if correlations do not correspond to a GMRF model or if the neighbor set is incorrect. We illustrate these points in the following examples.

Example 1 (Solution Using Two-Dimensional YW)

Suppose that we have a GMRF model with eight nearest neighbors

$$\theta_{1,0} = \theta_{-1,0} = 0.381, \qquad \theta_{0,1} = \theta_{0,-1} = 0.160$$
$$\theta_{1,1} = \theta_{-1,-1} = -0.015, \quad \theta_{1,-1} = \theta_{-1,1} = -0.056.$$

$I \backslash k$	0	1	2
−2	0.0950	0.0676	0.0462
−1	0.2923	0.1853	0.1153
0	1.0000	0.5252	0.2818
1	0.2923	0.2354	0.1659
2	0.0950	0.0926	0.0771

Using (5.4) for solving θ we get

$$\theta_{1,0} = \theta_{-1,0} = 0.381, \qquad \theta_{0,1} = \theta_{0,-1} = 0.160$$
$$\theta_{1,1} = \theta_{-1,-1} = -0.015, \quad \theta_{1,-1} = \theta_{-1,1} = -0.056.$$

Thus θ can be calculated using linear equations.

Example 2 (Effect of an Inappropriate Neighbor Set)

If for the same data we fit a first-order model, then the following parameters are obtained:

$$\theta_{1,0} = \theta_{-1,0} = 0.3686, \qquad \theta_{0,1} = \theta_{0,-1} = 0.1253$$
$$\nu = 0.5395.$$

The correlation values corresponding to this model follow.

$I \backslash k$	0	1	2
0	1.1230	0.7756	0.5031
1	0.5346	0.4651	0.3740
2	0.2990	0.2833	0.2540

The correlations $C(r)$, $r \in N_s$ are not matched in this case, as N_s is the wrong neighbor set.

VI. Realization of GMRF Models from the True Correlations

Suppose that we are given true correlations. One way to realize the GMRF model corresponding to these correlations is to substitute these correlations for $C(r)$ into (3.8) and to minimize the resulting expression for θ. The YW equations have been used to realize the GMRF models in [36]. However, from the examples of the previous section and the following examples it is clear that the determination of θ from the YW equations is not appropriate in most of the cases.

Example 3 (Inadmissibilty of YW Solutions)

In some cases the parameters obtained by solving YW equations can give a negative spectrum. For example, if

$$R(k,I) = 0.9^{\sqrt{k^2+I^2}},$$

then the parameters using YW equations are [36]

$$\theta_{1,0} = \theta_{-1,0} = -0.2518, \qquad \theta_{0,1} = \theta_{0,-1} = -0.2518$$
$$\theta_{1,1} = \theta_{-1,-1} = -0.0006, \quad \theta_{1,-1} = \theta_{-1,1} = -0.0006$$
$$\nu = 0.0395.$$

Thus we have

$$S(\pi,\pi) = -8.229.$$

Example 4 (Determination of Parameters when the Correlations Are Not from a GMRF Model): Suppose that the correlations are given by

$$R(k,I) = 0.3^{\sqrt{k^2+I^2}}.$$

This correlation function does not correspond to a GMRF model. If we determine parameters by using nonlinear minimization we get the correlations as

$I \backslash k$	0	1	2
0	0.999	0.299	0.099
1	0.299	0.155	0.067
2	0.099	0.067	0.036

with the corresponding parameters being $\theta_{1,0} = \theta_{-1,0} = 0.212$, $\theta_{0,1} = \theta_{0,-1} = 0.212$, and $\nu = 0.7455$. The true correlation values are

$I \backslash k$	0	1	2
0	1.000	0.300	0.090
1	0.300	0.182	0.068
2	0.090	0.068	0.033

The correlations $R(0,0)$, $R(0,1)$, and $R(1,0)$ are matched (within the numerical error). If we use (5.4) to solve for θ, we get the correlations as

$I \backslash k$	0	1	2
0	0.981	0.277	0.086
1	0.277	0.137	0.056
2	0.086	0.056	0.028

with corresponding parameters $\theta_{1,0} = \theta_{-1,0} = 0.206$, $\theta_{0,1} = \theta_{0,-1} = 0.206$, and $\nu = 0.7524$. Here the correlation matching is not as good as that resulting from the use of the nonlinear criterion function. Correlation matching for more lags can be obtained by using higher order models. For example, using a third-order model gives the correlations

$I \backslash k$	0	1	2
0	1.001	0.301	0.089
1	0.301	0.181	0.074
2	0.089	0.074	0.041

with parameters $\theta_{1,0} = \theta_{-1,0} = 0.195$, $\theta_{0,1} = \theta_{0,-1} = 0.195$, $\theta_{1,1} = \theta_{-1,-1} = 0.034$, $\theta_{-1,1} = \theta_{1,-1} = 0.034$, $\theta_{2,0} = \theta_{-2,0} = -0.017$, $\theta_{0,2} = \theta_{0,-2} = -0.017$, and $\nu = 0.7466$. Thus more correlation values are matched. The correlations using YW equations are given as

$I \backslash k$	0	1	2
0	1.008	0.309	0.101
1	0.309	0.193	0.084
2	0.101	0.084	0.048

with parameters $\theta_{1,0} = \theta_{-1,0} = 0.190$, $\theta_{0,1} = \theta_{0,-1} = 0.190$, $\theta_{1,1} = \theta_{-1,-1} = 0.038$, $\theta_{-1,1} = \theta_{1,-1} = 0.038$, $\theta_{2,0} = \theta_{-2,0} = -0.013$, $\theta_{0,2} = \theta_{0,-2} = -0.013$, and $\nu = 0.7474$. As before, the correlation matching is not good.

VII. Discussion

In this paper we have presented an estimate of two-dimensional MEPS using GMRF models for the given two-dimensional data. We have shown that the GMRF spectrum obtained by using ML estimates of parameters satisfies the sample correlation matching property. This result can be applied for higher dimensions also. However, in three or more dimensions there are some correlation sequences that are extendible but do not correspond to any GMRF model. Examples of such sequences can be found in [1] and [20]. Thus for three or more dimensions the solution to the ML equations may not exist. However, whenever a solution does exist it will be unique and will satisfy the correlation matching property. We have also shown that YW solutions generally do not satisfy the correlation matching property and may give negative spectra. Hence for realizing a two-dimensional noncausal GMRF model from given correlations, one should use the criterion function (3.8), with $C(r)$'s being replaced by true correlations. When data are given we cannot use the YW equations with true correlations replacing sample correlations even if the data are from a GMRF model. This is because statistical variability keeps the sample correlations from satisfying (5.2).

The modeling approach of this paper can be extended to the case when data are sampled nonuniformly [37]. Note that the case of nonuniformly sampled data is different from the case of correlations with nonuniform lags discussed in [1] and [14].

Acknowledgment

We are grateful to Dr. Victor Solo for pointing out reference [27] and to the participants of the panel session on Spectrum Estimation at the Third ASSP Workshop on Multidimensional Signal Processing held in Lake Tahoe, Nevada, for their useful comments.

References

[1] S. W. Lang and J. H. McClellan, "Multi-dimensional MEM spectral estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-30, pp. 880–887, Dec. 1982.

[2] J. Capon, "High resolution frequency-wavenumber spectral analysis," in *Proc. IEEE*, vol. 57, pp. 1408–1418, Aug. 1969.

[3] H. C. Andrews and B. R. Hunt, *Digital Image Restoration*. Englewood Cliffs, NJ: Prentice-Hall, 1977.

[4] V. S. Frost *et al.*, "A system model for imaging radars and its application to minimum mean square filtering," in *Proc. 5th Int. Conf. Patt. Recog.*, Miami, FL, pp. 131–133, Dec. 1980.

[5] J. Weszka, C. R. Dyer, and A. Rosenfeld, "A comparative study of texture measures for terrain classification," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-6, pp. 269–285, Apr. 1976.

[6] A. Vanden Bos, "Alternative interpretation of maximum entropy spectral analysis," *IEEE Trans. Inform. Theory*, vol. IT-17, pp. 493–494, July 1971.

[7] T. J. Ulrych and T. N. Bishop, "Maximum entropy spectral analysis and autoregressive decomposition," *Rev. Geophys. Space Phys.*, vol. 13, pp. 183–200, Feb. 1975.

[8] A. Papoulis, "Maximum entropy spectral estimation: a review," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, pp. 1176–1186, Dec. 1981.

[9] J. P. Burg, "Maximum entropy spectral analysis," Ph.D. dissertation, Stanford Univ., Stanford, CA.

[10] J. W. Woods, "Two-dimensional Markov spectral estimation," *IEEE Trans. Inform. Theory*, vol. IT-22, pp. 552–559, Sept. 1976.

[11] J. S. Lim and N. A. Malik, "A new algorithm for two-dimensional maximum entropy power spectrum estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, pp. 401–413, June 1981.

[12] A. K. Jain and S. Ranganath, "Two-dimensional spectrum estimation," in *Proc. RADC Spectrum Estimation Workshop*, pp. 151–157, May 1978.

[13] S. J. Wernecke and L. R. D'Addario, "Maximum entropy image reconstruction," *IEEE Trans. Comput.*, vol. C-26, pp. 351–364, Apr. 1977.

[14] J. H. McClellan, "Multidimensional spectral estimation," in *Proc. IEEE*, vol. 70, pp. 1029–1038, Sept. 1982.

[15] B. W. Dickinson, "Two-dimensional Markov spectrum estimates need not exist," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 120–121, Jan. 1980.

[16] R. Chellappa, Y. H. Hu, and S. Y. Kung, "On two-dimensional Markov spectral estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 836–841, Aug. 1983.

[17] J. W. Woods, "Two-dimensional discrete Markovian random fields," *IEEE Trans. Inform. Theory*, vol. 18, pp. 232–240, Mar. 1972.

[18] Yu. A. Rosanov, "On Gaussian fields with given conditional distributions," *Theory Prob. Appl.*, vol. 12, pp. 381–391, 1967.

[19] J. E. Besag, "Spatial interaction and statistical analysis of lattice

systems," *J. Roy. Statist. Soc.*, Series B, vol. B-36, pp. 192–236, 1974.

[20] H. Kunsch, "Thermodynamics and statistical analysis of Gaussian random fields," *Z. Wahr. Ver. Geb.*, vol. 58, pp. 407–421, Nov. 1981.

[21] A. K. Jain and S. Ranganath, "High resolution spectrum estimation in two-dimensions," *Proc. First Workshop on Spectral Estimation*, ON, Canada, Aug. 1981.

[22] R. L. Kashyap and R. Chellappa, "Estimation and choice of neighbors in spatial interaction models of images," *IEEE Trans. Inform. Theory*, vol. IT-29, pp. 60–72, Jan. 1983.

[23] P. A. P. Moran and J. E. Besag, "On the estimation and testing of spatial interaction in Gaussian lattices," *Biometrika*, vol. 62, pp. 555–562, 1975.

[24] D. E. Smylie *et al.*, "Analysis of irregularities in the earth's rotation," in *Methods in Computational Physics*, Vol. 13. New York: Academic, 1973, pp. 391–403.

[25] X. Guyon, "Parameter estimation for a stationary process on a d-dimensional lattice," *Biometrika*, vol. 69, pp. 95–105, 1982.

[26] S. W. Lang, "Spectral estimation for sensor array," Ph.D. dissertation, MIT, Cambridge, MA, Aug. 1981.

[27] L. L. Campbell, "Equivalence of Gauss's principle and minimum discrimination information estimation of probabilities," *Ann. Math. Statist.*, vol. 41, pp. 1011–1015, 1970.

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

[29] W. I. Newman, "A new method of multidimensional power spectral analysis," *Astron. Astrophys.*, vol. 54, pp. 369–380, 1977.

[30] H. Scheffe, *The Analysis of Variance*, New York: Wiley, 1959.

[31] H. J. Newton and M. Pagano, "Simultaneous confidence bands for autoregressive spectra," *Biometrika*, vol. 71, pp. 197–202, Mar. 1984.

[32] R. L. Kashyap, "Inconsistency of the AIC rule for estimating the order of autoregressive models," *IEEE Trans. Automat. Cont.*, vol. AC-25, pp. 996–998, Oct. 1980.

[33] G. Sharma, "Modern two-dimensional spectrum estimation using noncausal spatial models," Ph.D. dissertation, Univ. of Southern California, Los Angeles, CA, July 1984.

[34] R. L. Kashyap, "Random field models on finite lattices for finite images," *Proc. Conf. Inform. Sciences and Systems*, Johns Hopkins Univ., Baltimore, MD, May 1981.

[35] A. K. Jain, "Advances in mathematical models for image processing," *Proc. IEEE*, vol. 69, pp. 502–528, May 1981.

[36] S. Ranganath and A. K. Jain, "Two-dimensional linear prediction models, part I: Spectral factorization and realization," Rep. SIPL-83-5, Univ. of California, Davis, May 1983.

[37] K. Ord, "Estimation methods for models of spatial interaction," *J. Amer. Statist. Assoc.*, vol. 70, pp. 120–126, Mar. 1975.