

# Least-squares estimators in a stationary random field\*

SWAGATA NANDI AND DEBASIS KUNDU

Department of Mathematics, Indian Institute of Technology Kanpur, Kanpur 208 016, India.

## Abstract

A particular two-dimensional model in a stationary random field, which has wide applications in statistical signal processing and texture classifications, is considered. We prove the consistency and also obtain asymptotic distributions of the least-squares estimators of different model parameters. It is observed that the asymptotic distribution of the least-squares estimators is multivariate normal. Some numerical experiments are performed to see how the asymptotic results work for finite samples. We propose some open problems at the end.

**Keywords:** Strong consistency, texture classification, statistical signal processing, stationary random field.

## 1. Introduction

We consider the following two-dimensional model:

$$y(m, n) = \sum_{k=1}^q A_k^0 \cos(m\lambda_k^0 + n\mu_k^0) + X(m, n); \text{ for } m = 1, \dots, M, n = 1, \dots, N, \quad (1)$$

where  $A_k^0$ 's are unknown real numbers,  $\lambda_k^0$ 's and  $\mu_k^0$ 's are unknown frequencies. For identifiability, we need to assume  $\lambda_k^0 \in (-\pi, \pi)$  and  $\mu_k^0 \in (0, \pi)$  and they are distinct.  $X(m, n)$  is a two-dimensional (2-D) stationary random field described as follows:

$$X(m, n) = \sum_{i=-P}^P \sum_{j=-Q}^Q b(i, j) e(m-i, n-j). \quad (2)$$

Here  $\{e(m, n)\}$  is a two-dimensional sequence of independent and identically distributed (i.i.d.) random variable with mean zero and finite variance.  $P$  and  $Q$  are arbitrary positive integers,  $q$ , the number of components, is assumed to be a known integer. Given a sample  $y(m, n); m = 1, \dots, M, n = 1, \dots, N$ , the problem is to estimate  $A_k^0$ 's,  $\lambda_k^0$ 's,  $\mu_k^0$ 's for  $k = 1, \dots, q$ .

$X(m, n)$  and  $y(m, n)$  are stationary and non-stationary random fields, respectively. To see how this model represents different textures, the reader is referred to Mandrekar and Zhang<sup>1</sup> or Francos *et al.*,<sup>2</sup> who have provide nice 2-D image plots of  $y(m, n)$  for (i) grey level at  $(m, n)$  proportional to  $y(m, n)$  and (ii) when it is corrupted by independent Gaussian

\*The paper is dedicated to Professor C. R. Rao on his 80th birthday.

noise field. So this model represents mixed textures of regular textures with noise pictures. Our problem is to extract the regular textures from the contaminated  $y(m, n)$ . The problem is of interest in spectrograph and is studied using group-theoretic methods by Malliavan,<sup>3,4</sup> Francos *et al.*<sup>2</sup> considered the Wold-type decomposition of the random fields due to Helson and Lowdenslager,<sup>5,6</sup> but no concrete mathematical results were presented in it. Mandrekar and Zhang<sup>1</sup> also considered the spectral analysis of this problem under the following stationary assumptions on  $X(m, n)$

$$X(m, n) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} b(i, j)e(m-i, n-j), \quad (3)$$

where  $\{e(m, n)\}$  is a double-array sequence of independent random variables such that

$$\sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} |b(i, j)| < \infty, \quad E(e(m, n)) = 0, \quad E(|e(m, n)|^r) < \infty, \quad (4)$$

for some constant  $r > 2$ . They proved that the spectral estimators of  $\lambda$ 's and  $\mu$ 's are consistent estimators of the corresponding parameters when  $X(m, n)$  satisfies (3) and (4). Unfortunately, the corresponding estimators of the linear parameters ( $A$ 's) are not consistent. Moreover, they could not obtain the asymptotic distributions of the different estimators. Therefore, the rates of convergence of those estimators are not known. Their results are mainly based on the work of Lai and Wei<sup>7</sup> which is quite involved mathematically. In this paper, we mainly consider the least-squares estimators (LSEs) of the different parameters and study their large sample properties.

In the particular case, when  $\{X(m, n)\}$ 's are i.i.d. random variables on a 2-D plane, the problem can be interpreted as 'signal extraction'. It has wide applications in multidimensional signal processing. See, for example, the works of Barbieri and Barone,<sup>8</sup> Cabrera and Bose,<sup>9</sup> Chun and Bose,<sup>10</sup> Hua,<sup>11</sup> Lang and McClellan,<sup>12</sup> Kundu and Gupta<sup>13</sup> and the references therein for different estimation procedures and their applications. It is interesting to observe that model (1) is the 2-D extension of the one-dimensional frequency model which is a well-studied model in time series analysis; see, for example, the works of Hannan<sup>14</sup> and Walker<sup>15</sup> in this context.

In this paper, we consider the LSEs of the unknown parameters of model (1) under Assumption (2) on  $X(m, n)$ . It is well known that the LSEs play an important role in estimation theory. It has lots of desirable properties like consistency, asymptotic normality, asymptotic unbiasedness, etc. (see Rao<sup>16</sup>). But, nowhere, at least as known to the authors, the properties of the LSEs of this model have been discussed under this general set up. It is important to observe that it is a nonlinear regression model, but unfortunately it does not satisfy the standard sufficient conditions stated by Jennrich<sup>17</sup> or Wu<sup>18</sup> for the LSEs to be consistent. It may be noted that when  $q = 1$ ,  $M = 1$  and  $\lambda_k^0 = 0$ , this model coincides with the one-dimensional frequency model discussed in Hannan,<sup>14</sup> Walker,<sup>15</sup> Kundu<sup>19</sup> and Kundu and Mitra.<sup>20</sup> It was shown in Kundu<sup>19</sup> that even the one-dimensional model does not satisfy the sufficient conditions of Jennrich<sup>17</sup> or Wu.<sup>18</sup> Therefore, it is not immediately clear how the LSEs will behave

in this particular case under this general set up. In this paper, it is observed that the LSEs are consistent, unlike the spectral estimation method proposed by Mandrekar and Zhang,<sup>1</sup> where the estimators of the linear parameters are not consistent. We obtain the asymptotic distributions of the least-squares estimators, which was not attempted before under these general conditions for the two-dimensional model. The asymptotic distributions of the LSEs are multivariate normal and are useful to obtain the rates of convergence of LSEs of the unknown parameters.

It may be argued that the assumption of Mandrekar and Zhang<sup>1</sup> on  $X(m, n)$  is somewhat weaker than ours, because in our case  $P < \infty$  and  $Q < \infty$  as defined in (2). But since  $P$  and  $Q$  are arbitrary, therefore (3) can be closely approximated arbitrarily by (2) with sufficiently large  $P$  and  $Q$  (see Fuller<sup>21</sup>). Therefore, for all practical purposes they are equivalent. Moreover, Mandrekar and Zhang<sup>1</sup> use higher-order moment assumptions ( $r > 2$ ) on  $e(m, n)$  to prove the necessary consistency results, whereas we assume only the finite second moment of  $e(m, n)$  to prove consistency and asymptotic normality of the LSEs of all the unknown parameters. In this paper, almost sure convergence means the usual Lebesgue measure and is denoted by a.s. We will denote the set of positive integers by  $Z$ . Also, the notation  $a = O(b(M, N))$  means  $|a/b(M, N)|$  is bounded for all  $M$  and  $N$ .

The rest of the paper is organized as follows. In Section 2, we prove strong consistency and in Section 3 we obtain asymptotic distributions of the LSEs of the parameters of the model (1), when  $q = 1$ . For  $q > 1$ , the results are obtained in Section 4. We perform some numerical experiments and present those results in Section 5 and finally draw conclusions and propose some open problems in Section 6.

## 2. Consistency of the LSEs

In this section, we obtain the consistency of the LSEs of the unknown parameters of the model (1), when  $q = 1$ , i.e.

$$y(m, n) = A^0 \cos(m\lambda^0 + n\mu^0) + X(m, n); \text{ for } m = 1, \dots, M, n = 1, \dots, N.$$

The LSEs are obtained by minimizing  $Q(\theta)$ , where

$$Q(\theta) = \sum_{m=1}^M \sum_{n=1}^N (y(m, n) - A \cos(m\lambda + n\mu))^2. \quad (5)$$

Here,  $\theta = (A, \lambda, \mu)$ , the true parameter value and the LSE of  $\theta$  are denoted by  $\theta^0 = (A^0, \lambda^0, \mu^0)$  and  $\hat{\theta} = (\hat{A}, \hat{\lambda}, \hat{\mu})$ , respectively. We make the assumptions explicit on  $X(m, n)$  as follows.

**Assumption 1:** Let  $\{X(m, n); m, n \in Z\}$  be a stationary random field and each  $X(m, n)$  can be represented as (2),  $\{e(m, n); m, n \in Z\}$  is a double array sequence of i.i.d. random variables with mean zero and variance  $\sigma^2$ .

We use the following lemma to prove the necessary results.

**Lemma 1:** If the double array sequence  $\{X(m, n); m, n \in Z\}$  satisfies Assumption 1, then

$$\sup_{\alpha, \beta} \left| \frac{1}{N} \frac{1}{M} \sum_{m=1}^M \sum_{n=1}^N X(m, n) \cos(m\alpha) \cos(n\beta) \right| \xrightarrow{\text{a.s.}} 0 \text{ when } \min\{M, N\} \rightarrow \infty.$$

**Proof:** See Appendix I.

Note that Lemma 1 is a very strong result. It extends some of the existing one-dimensional results of Hannan,<sup>14</sup> Walker,<sup>15</sup> Rao and Zhao,<sup>22</sup> Kundu,<sup>19</sup> and Kundu and Mitra<sup>20, 23</sup> to the 2-D case. It also generalizes the multidimensional results of Bai *et al.*,<sup>24</sup> Rao *et al.*,<sup>25</sup> Kundu and Mitra,<sup>26</sup> and Kundu and Gupta<sup>13</sup> in some sense.

Consider the following assumption on the parameters of the model (1), when  $q = 1$ .

**Assumption 2:** Let  $A^0$  be an arbitrary real number not identically equal to zero,  $\lambda^0 \in (-\pi, \pi)$  and  $\mu^0 \in (0, \pi)$ .

Now we state the consistency result as the following theorem.

**Theorem 1:** Under Assumptions 1 and 2, the LSEs of the parameters of model (1) are strongly consistent, when  $q = 1$ .

**Proof:** Expanding (5), with the help of Lemma 1 and using the similar technique of Bai *et al.*,<sup>24</sup> the results can be obtained.

It is interesting to observe that although the errors are correlated the usual LSEs provide consistent solutions. For the general linear or nonlinear models, the usual LSEs are inconsistent if the errors are correlated.<sup>16, 27</sup> In the correlated case, we need to consider the generalized LSEs which are consistent. On the other hand, Theorem 1 may not be too surprising, because it is known<sup>28</sup> that the LSEs are consistent for one-dimensional frequency model, even if the errors are correlated. In this respect, one or higher-dimensional frequency models are quite different than the usual nonlinear models.

### 3. Asymptotic normality of the LSEs

In this section, we obtain the asymptotic distributions of the LSEs of the parameters of model (1) when  $q = 1$ . We use the following notations. The first derivative of  $Q(\theta)$  is a  $1 \times 3$  vector as

$$Q'(\theta) = \left[ \frac{\delta Q(\theta)}{\delta A}, \frac{\delta Q(\theta)}{\delta \lambda}, \frac{\delta Q(\theta)}{\delta \mu} \right]$$

and the second derivative is a  $3 \times 3$  matrix as follows;

$$Q''(\theta) = \begin{bmatrix} \frac{\delta^2 Q(\theta)}{\delta A^2} & \frac{\delta^2 Q(\theta)}{\delta A \delta \lambda} & \frac{\delta^2 Q(\theta)}{\delta A \delta \mu} \\ \frac{\delta^2 Q(\theta)}{\delta \lambda \delta A} & \frac{\delta^2 Q(\theta)}{\delta \lambda^2} & \frac{\delta^2 Q(\theta)}{\delta \lambda \delta \mu} \\ \frac{\delta^2 Q(\theta)}{\delta \mu \delta A} & \frac{\delta^2 Q(\theta)}{\delta \mu \delta \lambda} & \frac{\delta^2 Q(\theta)}{\delta \mu^2} \end{bmatrix}$$

Therefore, expanding  $Q'(\hat{\theta})$  around  $\theta^0$ , we obtain

$$Q'(\hat{\theta}) - Q'(\theta^0) = (\hat{\theta} - \theta^0)Q''(\bar{\theta}) \quad (6)$$

where  $\bar{\theta}$  is a point on the line joining the points  $\hat{\theta}$  and  $\theta^0$ . Note that  $Q'(\hat{\theta}) = 0$  and consider the  $3 \times 3$  diagonal matrix  $\mathbf{D}$  as follows.

$$\mathbf{D} = \begin{pmatrix} M^{-\frac{1}{2}}N^{-\frac{1}{2}} & 0 & 0 \\ 0 & M^{-\frac{1}{2}}N^{-\frac{1}{2}} & 0 \\ 0 & 0 & M^{-\frac{1}{2}}N^{-\frac{1}{2}} \end{pmatrix}.$$

Now (6) can be written as

$$(\hat{\theta} - \theta^0) = -Q'(\theta^0)[Q''(\bar{\theta})]^{-1} \quad (7)$$

if  $Q''(\bar{\theta})$  is a full-rank matrix (see at the end of this section). Equivalently,

$$(\hat{\theta} - \theta^0)\mathbf{D}^{-1} = -[Q'(\theta^0)\mathbf{D}][\mathbf{D}Q''(\bar{\theta})\mathbf{D}]^{-1}. \quad (8)$$

Now, let us consider different elements of  $[Q'(\theta^0)\mathbf{D}]$ .

$$\begin{aligned} \frac{1}{M^{\frac{1}{2}}N^{\frac{1}{2}}} \frac{\delta Q(\theta^0)}{\delta A} &= \frac{-2}{M^{\frac{1}{2}}N^{\frac{1}{2}}} \sum_{m=1}^M \sum_{n=1}^N X(m, n) \cos(m\lambda + n\mu), \\ \frac{1}{M^{\frac{3}{2}}N^{\frac{1}{2}}} \frac{\delta Q(\theta^0)}{\delta \lambda} &= \frac{2}{M^{\frac{3}{2}}N^{\frac{1}{2}}} \sum_{m=1}^M \sum_{n=1}^N X(m, n) Am \sin(m\lambda + n\mu), \\ \frac{1}{M^{\frac{1}{2}}N^{\frac{3}{2}}} \frac{\delta Q(\theta^0)}{\delta \mu} &= \frac{2}{M^{\frac{1}{2}}N^{\frac{3}{2}}} \sum_{m=1}^M \sum_{n=1}^N X(m, n) An \sin(m\lambda + n\mu). \end{aligned}$$

Using the central limit theorem of the stochastic process (see Fuller<sup>21</sup>), and using the following results of Mangulis<sup>29</sup> for  $\beta \neq 0$ ,

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n \cos^2(t\beta) &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n \sin^2(t\beta) = \frac{1}{2} \\ \lim_{n \rightarrow \infty} \frac{1}{n^2} \sum_{t=1}^n t \cos^2(t\beta) &= \lim_{n \rightarrow \infty} \frac{1}{n^2} \sum_{t=1}^n t \sin^2(t\beta) = \frac{1}{4} \\ \lim_{n \rightarrow \infty} \frac{1}{n^3} \sum_{t=1}^n t^2 \cos^2(t\beta) &= \lim_{n \rightarrow \infty} \frac{1}{n^3} \sum_{t=1}^n t^2 \sin^2(t\beta) = \frac{1}{6} \end{aligned}$$

$$\lim_{n \rightarrow \infty} \frac{1}{n^2} \sum_{t=1}^n t \sin(t\beta) \cos(t\beta) = 0$$

it follows that  $[Q'(\theta^0)\mathbf{D}]$  tends to a 3-variate normal distribution with mean vector zero and the dispersion matrix  $2\sigma^2 c\Sigma$ , where

$$c = \begin{matrix} \sum_{|i=-P|}^P \sum_{|j=-Q}^Q b(i, j) \cos(i\lambda^0) \cos(j\mu^0) & \sum_{|i=-P|}^P \sum_{|j=-Q}^Q b(i, j) \cos(i\lambda^0) \sin(j\mu^0) \\ \sum_{|i=-P|}^P \sum_{|j=-Q}^Q b(i, j) \sin(i\lambda^0) \cos(j\mu^0) & \sum_{|i=-P|}^P \sum_{|j=-Q}^Q b(i, j) \sin(i\lambda^0) \sin(j\mu^0) \end{matrix} \quad (9)$$

and

$$\Sigma = \begin{matrix} 1 & 0 & 0 \\ 0 & \frac{1}{3}A^{0^2} & \frac{1}{4}A^{0^2} \\ 0 & \frac{1}{4}A^{0^2} & \frac{1}{3}A^{0^2} \end{matrix} \quad (10)$$

Observe that because of Theorem 1,  $\bar{\theta}$  converges to  $\theta^0$  a.s. and

$$\lim_{M, N \rightarrow \infty} (\mathbf{D}Q''(\bar{\theta})\mathbf{D}) = \lim_{M, N \rightarrow \infty} (\mathbf{D}Q''(\theta^0)\mathbf{D}) = \Sigma. \quad (11)$$

Therefore, from (8), we have the following result.

**Theorem 2:** Under Assumptions 1 and 2, the limiting distribution of  $\left\{ M^{\frac{1}{2}} N^{\frac{1}{2}} (\hat{A} - A^0), M^{\frac{3}{2}} N^{\frac{1}{2}} (\hat{\lambda} - \lambda^0), M^{\frac{1}{2}} N^{\frac{3}{2}} (\hat{\mu} - \mu^0) \right\}$  as  $\text{Min}(M, N) \rightarrow \infty$  is a 3-variate normal with mean vector zero and covariance matrix  $2\sigma^2 c\Sigma^{-1}$ , when  $\Sigma^{-1}$  has the following structure:

$$\Sigma^{-1} = \begin{matrix} 1 & 0 & 0 \\ 0 & \frac{48}{7} \frac{1}{A^{0^2}} & -\frac{36}{7} \frac{1}{A^{0^2}} \\ 0 & -\frac{36}{7} \frac{1}{A^{0^2}} & \frac{48}{7} \frac{1}{A^{0^2}} \end{matrix} \quad (12)$$

Note that to prove Theorem 2, we use the fact that  $Q'(\bar{\theta})$  is a full-rank matrix a.s. for large  $M$  and  $N$ . In fact, we have used  $\mathbf{D}Q''(\bar{\theta})\mathbf{D}$  as of full-rank a.s. (see (8)). Now from (11), it is clear that for large  $M$  and  $N$ ,  $\mathbf{D}Q''(\theta^0)\mathbf{D}$  is a full-rank matrix. Since the elements of the matrix  $Q''(\theta)$  are continuous functions of  $\theta$  and  $\bar{\theta}$  converges to  $\theta^0$  a.s., therefore  $\mathbf{D}Q''(\bar{\theta})\mathbf{D}$  is a full-rank matrix a.s. for large  $M$  and  $N$ .

From Theorem 2, it is clear that the LSE of the amplitude ( $A$ 's) is asymptotically independent of the LSEs of the frequencies, whereas the LSEs of the two frequencies have a high negative correlation. The asymptotic variances of the LSEs of  $A$ ,  $\lambda$  and  $\mu$  are proportional to  $1/MN$ ,  $1/M^3NA^{0^2}$  and  $1/MN^3A^{0^2}$ , respectively. Therefore, it is immediate that the convergence rates of  $\hat{\lambda}$  and  $\hat{\mu}$  are of the orders  $O(M^3N^{-1})$  and  $O(M^{-1}N^3)$ , respectively, and both of them are faster than the convergence rate of  $\hat{A}$ , which is  $O((MN)^{-1})$ . Moreover, the asymptotic variances of  $\hat{\lambda}$  and  $\hat{\mu}$  are inversely proportional to  $A^{0^2}$ . This may not be very surprising because if  $A^{0^2}$  is small then it is difficult to estimate the frequencies.

#### 4. Multiparameter case

In this section, we consider model (1) for any integer  $q$ . We use the following notations

$$\theta_1 = (A_1, \lambda_1, \mu_1), \dots, \theta_q = (A_q, \lambda_q, \mu_q), \Psi = (\theta_1, \dots, \theta_q).$$

The true parameter value and the LSEs of  $\Psi$  will be denoted by  $\Psi^0$  and  $\hat{\Psi}$ , respectively. We investigate the consistency and the asymptotic properties of  $\hat{\Psi}$ , which is obtained by minimizing

$$R(\Psi) = \sum_{m=1}^M \sum_{n=1}^N \left( y(m, n) - \sum_{k=1}^q A_k \cos(m\lambda_k + n\mu_k) \right)^2 \quad (13)$$

with respect to  $\Psi$ . We need the following assumption.

**Assumption 3:** Let  $A_1^0, \dots, A_p^0$  be arbitrary real numbers, none of them being identically equal to zero;  $\lambda_1^0, \dots, \lambda_q^0 \in (-\pi, \pi)$  which are distinct; similarly,  $\mu_1^0, \dots, \mu_q^0 \in (0, \pi)$  are also distinct.

The following result provides the consistency results of the LSEs of the model parameter for the general case.

**Theorem 3:** Under Assumptions 1 and 3,  $\hat{\Psi}$  is a strongly consistent estimator of  $\Psi^0$ .

**Proof:** It is quite similar to the proof of Theorem 1, so it is omitted.

To establish the asymptotic distribution of  $\hat{\Psi}$ , we use the following notations. The  $3q \times 3q$  diagonal matrix  $\mathbf{V}$  and the  $3q \times 3q$  block diagonal matrix  $\Phi^{-1}$  are defined as follows.

$$\mathbf{V} = \begin{bmatrix} \mathbf{D} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \mathbf{D} \end{bmatrix}, \Phi^{-1} = \begin{bmatrix} c_1 \Sigma_1^{-1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & c_q \Sigma_q^- \end{bmatrix}$$

where  $\Sigma_j^{-1}$  can be obtained from  $\Sigma^{-1}$  defined in (12) by replacing  $A^{0^2}$  with  $A_j^{0^2}$  and similarly  $c_k$ 's can be obtained from  $c$ 's defined in (9) by replacing  $\lambda^0$  and  $\mu^0$  with  $\lambda_k^0$  and  $\mu_k^0$ , respectively.

**Theorem 4:** Under the same assumptions as Theorem 3,  $(\hat{\Psi} - \Psi^0)\mathbf{V}^{-1}$  converges to a  $3q$ -variate normal distribution with mean vector zero and the dispersion matrix  $2\sigma^2\Phi^{-1}$ , where  $\mathbf{V}^{-1}$  and  $\Phi^{-1}$  are as defined above.

**Proof:** The proof can be obtained quite similarly as Theorem 2, so it is omitted.

## 5. Numerical experiments and discussion

In this section, we present some results of the numerical experiments performed to see how the asymptotic results behave for finite sample sizes. We performed all the experiments in Silicon Graphics, using the random deviate generator of Press *et al.*<sup>30</sup> We considered the following model:

$$y(m, n) = 4.0 \cos(2.0m + 1.0n) + 5.0 \cos(2.5m + 1.5n) + X(m, n). \quad (14)$$

$X(m, n)$  has the following form:

$$X(m, n) = e(m, n) + 0.25e(m-1, n) + 0.25e(m+1, n) + 0.25e(m, n-1) + 0.25e(m, n+1)$$

$\{e(m, n); m = 1, \dots, M, n = 1, \dots, N\}$  are i.i.d. Gaussian random variables with mean zero and finite variance  $\sigma^2$ . The stationary random field  $X(m, n)$  has that particular structure which indicates that the error at the point  $(m, n)$  is equally influenced by the four equidistant points from  $(m, n)$ . We considered  $M = N = 10, 20, 30, 40, 50$  and  $\sigma = 25, 0.50, 0.75, 1.0$ . For each sample size and for each  $\sigma$  we computed the LSEs of  $A_1, A_2, \lambda_1, \lambda_2, \mu_1$  and  $\mu_2$  and observed the average estimates and the average mean-squared errors (MSEs) over 500 replications (Table I). We also report the asymptotic variances (ASV) for each parameter for comparison purposes.

From the simulations it becomes very clear that as sample size increases or the variance decreases, the average MSEs and biases of all the estimators decrease. It shows that all the estimators are consistent and asymptotically unbiased. Biases are quite small even when the sample sizes are quite small. It is clear that the MSEs of the estimators of the nonlinear parameters are smaller than that of the linear parameters even for small sample sizes. From the experimental study also it is clear that the estimation of linear parameters is more difficult (in terms of accuracy) compared to nonlinear parameters. Some of the asymptotic behaviors are present even at small sample sizes. For example, if  $A_1 < A_2$ , then it is observed that the MSEs of  $\hat{\mu}_2$  and  $\hat{\lambda}_2$  are smaller than that of  $\hat{\mu}_1$  and  $\hat{\lambda}_1$ , respectively. It is also observed that as the sample size increases, the MSEs become closer to the asymptotic variances, i.e.  $|\text{ASV} - \text{MSE}|$  decreases. Therefore, it is evident from the behavior of the MSEs that the asymptotic results can be used to draw small sample inferences for the different model parameters. In some cases, it is observed that the ASV is lower than the corresponding MSE. This may be



Table I

		Parameters		$\lambda_2$	$\mu_1$	$\mu_2$	
<b><math>M = N = 10</math></b>							
0.25	LSE	3.998	5.001	2.000	2.500	1.000	1.500
	MSE	1.14E-3	6.30E-4	7.22E-6	2.15E-6	6.60E-6	2.05E-6
	ASV	1.41E-3	5.04E	6.04E-6	1.38E-6	6.04E-6	1.38E-6
0.50	LSE	4.0000	4.998	2.000	2.500	1.000	1.500
	MSE	5.75E-3	2.60E-3	3.12E-5	8.96E-6	3.12E-5	8.05E-6
	ASV	5.64E-3	2.01E-3	2.42E-5	5.53E-6	2.42E-5	5.53E-6
0.75	LSE	3.992	4.999	2.000	2.500	1.000	1.500
	MSE	1.25E-2	6.02E-3	6.63E-5	1.87E-5	6.27E-5	1.82E-5
	ASV	1.27E-2	4.53E-3	5.44E-5	1.24E-5	5.44E-5	1.24E-5
1.0	LSE	3.989	5.005	2.000	2.500	1.000	1.500
	MSE	2.15E-2	1.00E-2	1.16E-4	3.55E-5	1.08E-4	3.38E-5
	ASV	2.26E-2	8.06E-3	9.67E-5	2.21E-5	9.67E-5	2.21E-5
<b><math>M = N = 20</math></b>							
0.25	LSE	4.000	5.001	2.000	2.500	1.000	1.500
	MSE	1.24E-4	5.48E-4	3.81E-7	1.11E-7	3.69E-7	1.16E-7
	ASV	3.53E-4	1.26E-4	3.78E-7	8.64E-8	3.78E-7	8.64E-8
0.50	LSE	3.999	5.002	2.000	2.500	1.000	1.500
	MSE	5.48E-4	3.38E-4	1.57E-6	3.77E-7	1.65E-6	3.61E-7
	ASV	1.41E-3	5.04E-4	1.51E-6	3.45E-7	1.51E-6	3.45E-7
0.75	LSE	3.997	5.004	2.000	2.500	1.000	1.500
	MSE	1.65E-3	8.00E-4	3.15E-6	8.81E-7	2.94E-6	9.91E-7
	ASV	3.17E-3	1.13E-3	3.40E-6	7.77E-7	3.40E-6	7.77E-7
1.0	LSE	4.000	5.005	2.000	2.500	1.000	1.500
	MSE	2.33E-3	1.37E-3	6.82E-6	1.72E-6	6.63E-6	1.71E-6
	ASV	5.64E-3	2.01E-3	6.04E-6	1.38E-6	6.04E-6	1.38E-6
<b><math>M = N = 30</math></b>							
0.25	LSE	4.000	5.000	2.000	2.500	1.000	1.500
	MSE	5.37E-5	3.93E-5	8.19E-8	2.15E-8	7.96E-8	1.99E-8
	ASV	1.57E-4	5.60E-5	7.46E-8	1.71E-8	7.46E-8	1.71E-8
0.50	LSE	4.000	5.000	2.000	2.500	1.000	1.500
	MSE	2.16E-4	1.51E-4	3.09E-7	8.27E-8	2.80E-7	7.79E-8
	ASV	6.27E-4	2.24E-4	2.98E-7	6.82E-8	2.98E-7	6.82E-8
0.75	LSE	4.001	5.003	2.000	2.500	1.000	1.500
	MSE	5.31E-4	3.76E-4	6.55E-7	1.75E-7	7.19E-7	1.69E-7
	ASV	1.41E-3	5.04E-4	6.71E-7	1.54E-7	6.71E-7	1.54E-7
1.0	LSE	4.001	5.001	2.000	2.500	1.000	1.500
	MSE	1.04E-3	7.13E-4	1.21E-6	3.22E-7	1.24E-6	2.97E-7
	ASV	2.51E-3	8.95E-4	1.19E-6	2.73E-7	1.19E-6	2.73E-7
<b><math>M = N = 40</math></b>							
0.25	LSE	4.001	5.001	2.000	2.500	1.000	1.500
	MSE	1.14E-5	1.60E-5	2.49E-8	5.98E-9	2.37E-8	6.07E-9
	ASV	8.81E-5	3.15E-5	2.36E-8	5.40E-9	2.36E-8	5.40E-9
0.50	LSE	4.001	5.000	2.000	2.500	1.000	1.500
	MSE	6.47E-5	7.66E-5	9.57E-8	2.53E-8	9.58E-8	2.43E-8
	ASV	3.53E-4	1.26E-4	9.44E-8	2.16E-8	9.44E-8	2.16E-8
0.75	LSE	4.000	5.001	2.000	2.500	1.000	1.500
	MSE	1.20E-4	1.78E-4	2.11E-7	5.69E-8	2.04E-7	5.94E-8
	ASV	7.93E-4	2.83E-4	2.12E-7	4.86E-8	2.12E-7	4.86E-8
1.0	LSE	4.001	5.000	2.000	2.500	1.000	1.500
	MSE	1.74E-4	2.81E-4	3.70E-7	8.61E-8	3.77E-7	9.05E-8
	ASV	1.41E-3	5.04E-4	3.78E-7	8.64E-8	3.78E-7	8.64E-8

(contd)

$\sigma$	Parameters	$A_1$	$A_2$	$\lambda_1$	$\lambda_2$	$\mu_1$	$\mu_2$
<b><math>M = N = 50</math></b>							
0.25	LSE	3.999	5.000	2.000	2.500	1.000	1.500
	MSE	2.00E-5	1.37E-5	9.05E-9	2.48E-9	8.65E-9	2.33E-9
	ASV	5.64E-5	2.01E-5	9.67E-9	2.21E-9	9.67E-9	2.21E-9
0.50	LSE	4.000	5.000	2.000	2.500	1.000	1.500
	MSE	5.89E-5	4.85E-5	3.70E-8	1.02E-8	3.59E-8	9.97E-9
	ASV	2.26E-4	8.06E-5	3.87E-8	8.84E-9	3.87E-8	8.84E-9
0.75	LSE	4.000	5.000	2.000	2.500	1.000	1.500
	MSE	1.67E-4	1.21E-4	9.29E-8	2.12E-8	9.14E-8	2.01E-8
	ASV	5.08E-4	1.81E-4	8.70E-8	1.99E-8	8.70E-9	1.99E-8
1.0	LSE	3.999	4.999	2.000	2.500	1.000	1.500
	MSE	3.15E-4	2.34E-4	1.80E-7	3.76E-8	1.59E-7	3.89E-8
	ASV	9.02E-4	3.22E-4	1.55E-7	3.54E-8	1.55E-7	3.54E-8

due to sampling error as we have considered only 500 replications. (see Karian and Dudewicz.<sup>31</sup>)

## 6. Conclusions

In this paper, we consider the estimation of the parameters of a two-dimensional model which has wide applicability in statistical signal processing and in texture classification. We study the asymptotic properties of the LSEs of the model parameters and show that the LSEs are strongly consistent. We also obtain the asymptotic distributions of the LSEs which provide the rate of convergence of the LSEs. This paper generalizes some of the existing one-dimensional results to the 2D case. It generalizes some of the multidimensional results also in a certain way. Numerical experiments suggest that asymptotic results can be used to draw small sample inferences for linear and nonlinear parameters. We do not address one important problem, namely, the estimation of  $q$ , which is very important in practice. We may have to use certain information-theoretic criteria like AIC, BIC or use cross validation-type technique as proposed by Rao<sup>32</sup> for the one-dimensional case. Another important problem is to obtain an efficient estimator of the different parameters by some non-iterative technique. Non-iterative techniques are important for online implementation or to use as initial guesses for any iterative procedure. More work is needed in these directions.

## Acknowledgments

Most of the work was done when the second author was visiting the Department of Statistics, The Pennsylvania State University as a visiting associate professor. He would like to thank the Department and particularly Professor C. R. Rao for many helpful discussions. The authors would like to thank two referees for constructive comments.

## References

1. MANDREKAR, V. AND ZHANG, H.

*Determination of discrete spectrum of random fields with application to texture classification*, Technical Report, Michigan State University, USA, 1995.

2. FRANCOS, J. M., MIERI, A. Z. AND PORAT, B. On a Wold type decomposition of 2-D discrete random field. *Proc. Int. Conf. Acoustics Speech and Signal Processing*, Albuquerque, 1990.
3. MALLIAVAN, P. Sur la norme d'une matrice circulante gaussienne, *C.R. Acad. Sci. Paris t. 319, Serie I*, 1994, 745-749.
4. MALLIAVAN, P. Estimation d'un signal lorentzien, *C.R. Acad. Sci. Paris t. 319, Serie I*, 1994, 991-997.
5. HELSON, H. AND LOWDENSLAGER, D. Prediction theory and Fourier series in several variables, *Acta Math.*, 1958, **99**, 165-202.
6. HELSON, H. AND LOWDENSLAGER, D. Prediction theory and Fourier series in several variables, *Acta Math.*, 1962, **106**, 175-213.
7. LAI, T. L. AND WEI, Z. C. A law of iterated logarithm for double array of independent random variables with applications to regression and time series models, *Ann. Prob.*, 1982, **10**, 320-335.
8. BARBIERI, M. M. AND BARONE, P. A two dimensional Prony's method for spectral estimation, *IEEE Trans.*, 1992, **SP-40**, 2747-2756.
9. CABRERA, S. D. AND BOSE, N. K. Prony's method for two dimensional complex exponentials modeling, *Applied control* (S. G. Tzafestas, ed.), Ch. 15, Dekker, 1993, pp. 401-411.
10. CHUN, J. AND BOSE, N. K. Parameter estimation via signal selectivity of signal subspaces (PESS) and its applications to 2-D wave number estimation, *Digital Signal Processing*, 1995, **5**, 58-76.
11. HUA, Y. Estimating two dimensional frequencies by matrix enhancement and matrix pencil, *IEEE Trans.*, 1992, **SP-40**, 2267-2280.
12. LANG, S. W. AND MCCLELLAN, J. H. The extension of Pisarenko's method to multiple dimensions, *Proc. ICASSP-82*, 1982, pp. 125-128.
13. KUNDU, D. AND GUPTA, R. D. Asymptotic properties of the least squares estimators of a two dimensional texture model, *Metrika*, 1998, **48**, 83-97.
14. HANNAN, E. J. Nonlinear time series regression, *J. Appl. Prob.*, 1971, **8**, 767-780.
15. WALKER, A. M. On the estimation of a harmonic component in a time series with stationary residuals, *Biometrika*, 1971, **58**, 21-36.
16. RAO, C. R. *Linear statistical inferences and its applications*, 2nd edition, Wiley, 1973.
17. JENNRICH, R. I. Asymptotic properties of the non-linear least squares estimation, *Ann. Math. Stat.*, 1969, **40**, 633-643.
18. WU, C. F. J. Asymptotic theory of non-linear least squares estimation, *Ann. Stat.*, 1981, **9**, 510-513.
19. KUNDU, D. Asymptotic theory of the least squares estimator of a particular non-linear regression model, *Stat. Prob. Lett.*, 1993, **18**, 13-17.
20. KUNDU, D. AND MITRA, A. Asymptotic theory of the least squares estimator of a non-linear regression model, *Commun. Stat., Theory Meth.*, 1996, **25**, 133-141.
21. FULLER, W. *Introduction to statistical time series*, Wiley, 1976.
22. RAO, C. R. AND ZHAO, L. C. Asymptotic behavior of the maximum likelihood estimators of superimposed exponential signals, *IEEE Trans.*, 1993, **SP-41**, 1461-1463.

23. KUNDU, D. AND MITRA, A. On asymptotic properties and confidence intervals of superimposed exponential signals, *Signal Processing*, 1999, **22**, 129–139.
24. BAI, Z. D., KUNDU, D. AND MITRA, A. A theorem in probability and its applications in multidimensional signal processing, *IEEE Trans.*, 1996, **SP-44**, 3167–3169.
25. RAO, C. R., ZHAO, L. C. AND ZHOU, B. Maximum likelihood estimation of 2-D superimposed exponential signals, *IEEE Trans.*, 1994, **ASSP-42**, 1795–1802.
26. KUNDU, D. AND MITRA, A. Asymptotic properties of the least squares estimators of 2-D exponential signals, *Multidimensional Systems Signal Processing*, 1996, **7**, 135–150.
27. SEBER, G. A. F. AND WILD, C. J. *Nonlinear regression*, Wiley, 1989.
28. KUNDU, D. Asymptotic theory of the least squares estimators of sinusoidal signal, *Statistics*, 1997, **30**, 221–238.
29. MANGULIS, V. *Handbook of series for scientists and engineers*, Academic Press, 1965.
30. PRESS, W. H. *Numerical recipes in Fortran*, Cambridge University Press, 1993.
31. KARIAN, Z. A. AND DUDEWICZ, E. J. *Modern statistical systems and GPSS simulations*, 2nd edition, CRC Press, Washington, DC, 1999.
32. RAO, C. R. Some recent developments in signal detection, *Statistical decision theory and related topics*, (J. O. Berger and S. S. Gupta, eds) Springer-Verlag, 1988, pp. 319–332.

## Appendix I

**Proof of Lemma 1:** First we prove the result when  $X(m, n)$  is replaced by  $e(m, n)$

Consider the following random variables;

$$\begin{aligned} Z(m, n) &= e(m, n) \quad \text{if } |X(m, n)| < (mn)^{\frac{3}{4}} \\ &= 0 \quad \text{otherwise.} \end{aligned}$$

First we will show that  $Z(m, n)$  and  $e(m, n)$  are equivalent sequences. Consider

$$\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} P\{e(m, n) \neq Z(m, n)\} = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} P\{|e(m, n)| > (mn)^{\frac{3}{4}}\}.$$

Now observe that there are at most  $2^k$   $k$  combinations of  $(m, n)$ 's such that  $mn < 2^k$ ; therefore, we have

$$\begin{aligned} &\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} P\{|e(m, n)| \geq (mn)^{\frac{3}{4}}\} \\ &\leq \sum_{k=1}^{\infty} \sum_{2^{k-1} \leq r < 2^k} P\{|e(m, n)| \geq r^{\frac{3}{4}}\} \quad \{\text{here } r = mn\} \end{aligned}$$

$$\begin{aligned} &\leq \sum_{k=1}^{\infty} k 2^k P\left\{|e(1,1)| \geq 2^{(k-1)\frac{3}{4}}\right\} \\ &\leq C \sum_{k=1}^{\infty} k 2^k \frac{E|e(1,1)|^2}{2^{(k-1)\frac{3}{2}}} \leq C \sum_{k=1}^{\infty} \frac{k}{2^{\frac{k}{2}}} < \infty. \end{aligned}$$

Here,  $C$  is a constant and note that it may represent different constants at different places. Therefore,  $e(m, n)$  and  $Z(m, n)$  are equivalent sequences. So

$$P\{e(m, n) \neq Z(m, n) \text{ i.o.}\} = 0.$$

Here i.o. means infinitely often. Let  $U(m, n) = Z(m, n) - E(Z(m, n))$ , then

$$\sup_{\alpha, \beta} \left| \frac{1}{N} \frac{1}{M} \sum_{m=1}^M \sum_{n=1}^N E(Z(m, n)) \cos(m\alpha) \cos(n\beta) \right| \leq \frac{1}{N} \frac{1}{M} \sum_{m=1}^M \sum_{n=1}^N |E(Z(m, n))|.$$

Since  $E(Z(m, n)) \rightarrow 0$  as  $M, N \rightarrow \infty$ , therefore, as  $M, N \rightarrow \infty$

$$\frac{1}{N} \frac{1}{M} \sum_{m=1}^M \sum_{n=1}^N |E(Z(m, n))| \rightarrow 0.$$

Therefore, it is enough to prove that

$$\sup_{\alpha, \beta} \left| \frac{1}{N} \frac{1}{M} \sum_{m=1}^M \sum_{n=1}^N U(m, n) \cos(m\alpha) \cos(n\beta) \right| \rightarrow 0.$$

Now, for any fixed  $\epsilon > 0$ ,  $-\pi < \alpha, \beta < \pi$  and  $0 < h \leq \frac{1}{2(MN)^{\frac{3}{4}}}$ , we have

$$\left| \frac{1}{N} \frac{1}{M} \sum_{m=1}^M \sum_{n=1}^N U(m, n) \cos(m\alpha) \cos(n\beta) \right| \geq \epsilon \left\} \leq 2e^{-hMN\epsilon} \prod_{m=1}^M \prod_{n=1}^N E e^{hU(m, n) \cos(m\alpha) \cos(n\beta)}.$$

Since  $|hU(m, n) \cos(m\alpha) \cos(n\beta)| \leq 1/2$ , using  $e^x < 1 + x + x^2$  for  $|x| < 1/2$ , we have

$$2e^{-hMN\epsilon} \prod_{m=1}^M \prod_{n=1}^N E e^{hU(m, n) \cos(m\alpha) \cos(n\beta)} \leq 2e^{-hMN\epsilon} (1 + h^2 \sigma^2)^{MN}.$$

Now, choose  $h = \frac{1}{2(MN)^{\frac{3}{4}}}$ ; therefore, for large  $M$  and  $N$

$$\left| \frac{1}{N} \frac{1}{M} \sum_{m=1}^M \sum_{n=1}^N U(m, n) \cos(m\alpha) \cos(n\beta) \right| \geq \epsilon \left\} \leq C e^{-(MN)^{\frac{1}{4}} \epsilon / 2} \quad (C \text{ is a constant}).$$

Let  $K = M^2 N^2$ , choose  $K$  points,  $\theta_1 = (\alpha_1, \beta_1), \dots, \theta_K = (\alpha_K, \beta_K)$ , such that for each point  $\theta = (\alpha, \beta) \in (-\pi, \pi)$ , we have a point  $\theta_j$  satisfying

$$|\alpha_j - \alpha| + |\beta_j - \beta| \leq \frac{2\pi}{M^2 N^2}$$

Note that

$$\begin{aligned} & \left| \frac{1}{N} \frac{1}{M} \sum_{m=1}^M \sum_{n=1}^N U(m, n) \left\{ \cos(m\alpha) \cos(n\beta) - \cos(m\alpha_j) \cos(n\beta_j) \right\} \right| \\ & \leq C \frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N \frac{MN}{M^2 N^2} [m+n] \rightarrow 0 \text{ as } M, N \rightarrow \infty. \end{aligned}$$

Therefore, for large  $M$  and  $N$ , we have

$$\begin{aligned} & P \left\{ \sup_{\alpha, \beta} \left| \frac{1}{N} \frac{1}{M} \sum_{m=1}^M \sum_{n=1}^N U(m, n) \cos(m\alpha) \cos(n\beta) \right| \geq 2\varepsilon \right\} \\ & \leq P \left\{ \max_{j \leq M^2 N^2} \left| \frac{1}{N} \frac{1}{M} \sum_{m=1}^M \sum_{n=1}^N U(m, n) \cos(m\alpha_j) \cos(n\beta_j) \right| \geq \varepsilon \right\} \leq CM^2 N^2 e^{-(MN)^{\frac{1}{4}}} \varepsilon / 2 \end{aligned}$$

Since  $\sum_{i=1}^{\infty} t^2 e^{-t^{\frac{1}{4}}} < \infty$ , from Borel Cantelli's lemma, we have

$$\sup_{\alpha, \beta} \left| \frac{1}{N} \frac{1}{M} \sum_{m=1}^M \sum_{n=1}^N U(m, n) \cos(m\alpha) \cos(n\beta) \right| \xrightarrow{\text{a.s.}} 0.$$

Therefore,

$$\sup_{\alpha, \beta} \left| \frac{1}{N} \frac{1}{M} \sum_{m=1}^M \sum_{n=1}^N e(m, n) \cos(m\alpha) \cos(n\beta) \right| \xrightarrow{\text{a.s.}} 0.$$

Since  $P < \infty$ ,  $Q < \infty$  and  $|b(i, j)| < \infty$ , it proves the lemma.