STRONG CONSISTENCY OF THE OVER- AND UNDER-DETERMINED LSE OF 2-D EXPONENTIALS IN WHITE NOISE

Joseph M. Francos and Mark Kliger

Electrical & Computer Engineering Dept.
Ben-Gurion University
Beer Sheva 84105, Israel

ABSTRACT

We consider the problem of least squares estimation of the parameters of 2-D exponential signals observed in the presence of an additive noise field, when the assumed number of exponentials is incorrect. We consider both the case where the number of exponential signals is under-estimated, and the case where the number of exponential signals is over-estimated. In the case where the number of exponential signals is under-estimated we prove the almost sure convergence of the least squares estimates to the parameters of the dominant exponentials. In the case where the number of exponential signals is over-estimated, the estimated parameter vector obtained by the least squares estimator contains a sub-vector that converges almost surely to the correct parameters of the exponentials.

1. INTRODUCTION

In this paper we consider the problem of estimating the parameters of 2-D exponential signals, observed in the presence of an additive noise field. 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, Francos *et. al.* [2]. This modelling and estimation problem has fundamental theoretical importance, as well as various applications in texture estimation of images (see, e.g., [3] and the references therein) and in wave propagation problems (see, e.g., [9] and the references therein).

The problem of estimating 2-D exponential signals has been intensively investigated in the literature (see, e.g., [6] and the references therein). Recently, Rao *et. al.* [8] have studied the asymptotic properties of the maximum likelihood estimator (MLE) of 2-D exponential signals observed in noise. In this framework they also proved the strong consistency of the least squares estimates (LSE) of the parameters of 2-D exponentials observed in the presence of complex white circular Gaussian noise. Kundu and Gupta [7] extended the result of [8] to the case where the observation noise is not necessarily Gaussian. In both papers, as well as in most of the previous studies it is assumed that the number of exponentials is known. However this assumption does not always hold in practice.

In this paper we consider the problem of *least squares* estimation of the parameters of 2-D exponential signals observed in the presence of an additive noise field, when the assumed number of exponentials is incorrect. Let P denote the number of exponential signals in the observed field and let k denote their assumed number. In the case where the number of exponential signals is under-estimated, *i.e.*, $k < P$, we prove the almost sure convergence of the least squares estimates to the parameters of the k dominant exponentials. In the case where the number of exponential signals is over-estimated, *i.e.*, $k > P$, we prove the almost sure convergence of the estimates obtained by the least squares estimator to the parameters of the P exponentials in the observed field. The extra $k-P$ components assumed to exist, are assigned by the least squares estimator to the dominant components of the periodogram of the noise field.

A solution to the problem addressed here, is an essential component in the error analysis of the LS algorithm for estimating 2-D exponentials in noise and in analyzing the performance of the model order selection criteria [5].

2. NOTATIONS, DEFINITIONS AND ASSUMPTIONS

Let $\{y(n,m)\}$ be a complex valued field,

$$y(n,m) = \sum_{i=1}^{P} a_i^0 e^{j(\omega_i^0 n + \upsilon_i^0 m)} + u(n,m), \tag{1}$$

where $0 \leq n \leq S-1$, $0 \leq m \leq T-1$ and for each i, a_i^0 is non-zero. Due to physical considerations it is further assumed that for each i, $|a_i^0|$ is bounded.

We make the following assumptions:

Assumption 1: The field $\{u(n,m)\}$ is an i.i.d. complex valued zero-mean random field. Let $u(n,m) = \Re(u(n,m)) + j\Im(u(n,m))$ where $u_R(n,m) = \Re(u(n,m))$ and $u_I(n,m) = \Im(u(n,m))$ are the real and imaginary parts of $u(n,m)$ respectively. Both $u_R(n,m)$ and $u_I(n,m)$ are zero mean with finite second order moment, $\frac{\sigma^2}{2}$. The real and imaginary parts are independent.

Assumption 2: The spatial frequencies $(\omega_i^0, \upsilon_i^0) \in (0, 2\pi) \times (0, 2\pi)$, $1 \leq i \leq P$ are pairwise different. In other words, $\omega_i^0 \neq \omega_j^0$ or $\upsilon_i^0 \neq \upsilon_j^0$, when $i \neq j$.

Define the loss function due to the error of the k-th order regression model

$$\mathcal{L}(a_1, \ldots, a_k, \omega_1, \upsilon_1, \ldots, \omega_k, \upsilon_k) = \frac{1}{ST} \sum_{n=0}^{S-1} \sum_{m=0}^{T-1} \left| y(n,m) - \sum_{i=1}^{k} a_i e^{j(n\omega_i + m\upsilon_i)} \right|^2. \tag{2}$$

Let $\{\Psi_i\}$ be a sequence of rectangles such that

$$\Psi_i = \{(n,m) \in \mathcal{Z}^2 \mid 0 \leq n \leq S_i - 1, 0 \leq m \leq T_i - 1\}.$$

Definition 1: The sequence of subsets $\{\Psi_i\}$ is said to tend to infinity (we adopt the notation $\Psi_i \to \infty$) as $i \to \infty$ if

$$\lim_{i\to\infty} \min(S_i, T_i) = \infty$$

and

$$0 < \lim_{i\to\infty} (S_i/T_i) < \infty.$$

To simplify notations, we shall omit in the following the subscript i. Thus, the notation $\Psi(S,T) \to \infty$ implies that both S and T tend to infinity as functions of i, and at roughly the same rate.

Definition 2: Let Θ_k be a bounded and closed subset of the $3k$ dimensional space $\mathrm{C}^k \times ((0,2\pi) \times (0,2\pi))^k$ where for any vector

$$\theta_k = (a_1, \ldots, a_k, \omega_1, \upsilon_1, \ldots, \omega_k, \upsilon_k) \in \Theta_k$$

the coordinate a_i is non-zero and absolutely bounded for every $1 \le i \le k$ while the pairs (ω_i, υ_i) are pairwise different, so that no two regressors coincide. We shall refer to Θ_k as the *parameter space*.

From the model definition and the above assumptions it is clear that

$$\theta_k^0 = (a_1^0, \ldots, a_k^0, \omega_1^0, \upsilon_1^0, \ldots, \omega_k^0, \upsilon_k^0) \in \Theta_k.$$

A vector $\hat{\theta}_k \in \Theta_k$ that minimizes $\mathcal{L}$ is called the *Least Square Estimate* (LSE). In the case where $k = P$, the LSE is a *strongly consistent* estimator of θ_P^0 (see, *e.g.*, [7] and the references therein). In the following sections we establish the strong consistency of this LSE when the number of exponentials is under-estimated, or over-estimated.

3. CONSISTENCY OF THE LSE FOR AN UNDER-ESTIMATED MODEL ORDER

Let k denote the assumed number of observed 2-D exponentials, where $k < P$. For any $\delta > 0$, define the set Δ_δ to be a subset of the parameter space Θ_k such that each vector $\theta_k \in \Delta_\delta$ is different from the vector θ_k^0 by at least δ, at least in one of its coordinates, *i.e.*,

$$\Delta_\delta = \left[\bigcup_{i=1}^{k} A_{i\delta}\right] \cup \left[\bigcup_{i=1}^{k} W_{i\delta}\right] \cup \left[\bigcup_{i=1}^{k} V_{i\delta}\right], \qquad (3)$$

where

$$A_{i\delta} = \left\{\theta_k \in \Theta_k : \ |a_i - a_i^0| \ge \delta; \delta > 0 \ \right\}, \qquad (4)$$

$$W_{i\delta} = \left\{\theta_k \in \Theta_k : \ |\omega_i - \omega_i^0| \ge \delta; \delta > 0 \ \right\}, \qquad (5)$$

$$V_{i\delta} = \left\{\theta_k \in \Theta_k : \ |\upsilon_i - \upsilon_i^0| \ge \delta; \delta > 0 \ \right\}. \qquad (6)$$

To prove the main result of this section we shall need an additional assumption and the following lemmas:

Assumption 3: For convenience, and without loss of generality, we assume that the exponentials are indexed according to a descending order of their amplitudes, *i.e.*,

$$|a_1^0| \ge |a_2^0| \ge \ldots |a_k^0| > |a_{k+1}^0| \ldots \ge |a_P^0| > 0, \qquad (7)$$

where we assume that for a given k, $|a_k^0| > |a_{k+1}^0|$ to avoid trivial ambiguities resulting from the case where the k-th dominant component is not unique.

Lemma 1

$$\liminf_{\Psi(S,T)\to\infty} \inf_{\theta_k \in \Delta_\delta} \left(\mathcal{L}(\theta_k) - \mathcal{L}(\theta_k^0)\right) > 0 \ \ a.s. \qquad (8)$$

See [6] for the proof.

Lemma 2 *Let $\{x_n, n \ge 1\}$ be a sequence of random variables. Then*

$$\Pr\{x_n \le 0 \ i.o.\} \le \Pr\{\liminf_{n\to\infty} x_n \le 0\}, \qquad (9)$$

where the abbreviation *i.o.* stands for *infinitely often*.

See [6] for the proof.

The next theorem establishes the strong consistency of the least squares estimator in the case where the number of the regressors is lower than the actual number of exponentials.

Theorem 1 *Let Assumptions 1-3 be satisfied. Then, the k-regressors parameter vector*

$$\hat{\theta}_k = (\hat{a}_1, \ldots, \hat{a}_k, \hat{\omega}_1, \hat{\upsilon}_1, \ldots, \hat{\omega}_k, \hat{\upsilon}_k),$$

that minimizes (2) is a strongly consistent estimator of

$$\theta_k^0 = (a_1^0, \ldots, a_k^0, \omega_1^0, \upsilon_1^0, \ldots, \omega_k^0, \upsilon_k^0),$$

as $\Psi(S,T) \to \infty$. That is,

$$\hat{\theta}_k \to \theta_k^0, \ \ a.s. \ \ as \ \ \Psi(S,T) \to \infty.$$

Proof: The proof follows an argument proposed by Wu [10], Lemma 1. Let $\hat{\theta}_k = (\hat{a}_1, \ldots, \hat{a}_k, \hat{\omega}_1, \hat{\upsilon}_1, \ldots, \hat{\omega}_k, \hat{\upsilon}_k)$ be a parameter vector that minimizes (2). Assume that the proposition $\hat{\theta}_k \to \theta_k^0 \ a.s. \ as \ \Psi(S,T) \to \infty$ is not true. Then, there exists some $\delta > 0$, such that ([1], Theorem 4.2.2),

$$\Pr(\hat{\theta}_k \in \Delta_\delta \ \ i.o.) > 0. \qquad (10)$$

This inequality together with the definition of $\hat{\theta}_k$ as a vector that minimizes $\mathcal{L}$ implies

$$\Pr\left(\inf_{\theta_k \in \Delta_\delta} \left(\mathcal{L}(\theta_k) - \mathcal{L}(\theta_k^0)\right) \le 0 \ \ i.o.\right) > 0. \qquad (11)$$

Using Lemma 2 we obtain

$$\Pr\left(\liminf_{\Psi(S,T)\to\infty} \inf_{\theta_k \in \Delta_\delta} \left(\mathcal{L}(\theta_k) - \mathcal{L}(\theta_k^0)\right) \le 0\right)$$
$$\ge \Pr\left(\inf_{\theta_k \in \Delta_\delta} \left(\mathcal{L}(\theta_k) - \mathcal{L}(\theta_k^0)\right) \le 0 \ \ i.o.\right) > 0, \qquad (12)$$

which contradicts (8). Hence,

$$\hat{\theta}_k \to \theta_k^0 \ \ a.s. \ \ as \ \ \Psi(S,T) \to \infty. \qquad (13)$$

■

4. CONSISTENCY OF THE LSE FOR AN OVER-ESTIMATED MODEL ORDER

Let k denote the assumed number of observed 2-D exponentials, where $k > P$. Without loss of generality, we can assume that $k = P + 1$, (as the proof for $k \geq P + 1$ follows immediately by repeating the same arguments). Let the periodogram of the field $\{u(n, m)\}$ be given by

$$I_u(\omega, \upsilon) = \frac{1}{ST} \left| \sum_{n=0}^{S-1} \sum_{m=0}^{T-1} u(n, m) e^{-j(n\omega + m\upsilon)} \right|^2 . \tag{14}$$

The parameter spaces Θ_P, Θ_{P+1} are defined as in Definition 2.

To prove the main result of this section we need an additional assumption:

Assumption 4: The real and imaginary components of $u(n, m)$ are such that

$$E[u_R(0, 0)^2 \log |u_R(0, 0)|] < \infty$$

and

$$E[u_I(0, 0)^2 \log |u_I(0, 0)|] < \infty.$$

(For example, a white Gaussian noise field satisfies this assumption).

Theorem 2 *Let Assumptions 1, 2, and 4 be satisfied. Then, the parameter vector*

$$\hat{\theta}_{P+1} = (\hat{a}_1, \ldots, \hat{a}_P, \hat{a}_{P+1}, \hat{\omega}_1, \hat{\upsilon}_1, \ldots \\ \ldots, \hat{\omega}_P, \hat{\upsilon}_P, \hat{\omega}_{P+1}, \hat{\upsilon}_{P+1}) \in \Theta_{P+1},$$

that minimizes (2) with $k = P + 1$ regressors as $\Psi(S, T) \to \infty$ is composed of the vector

$$\hat{\theta}_P = (\hat{a}_1, \ldots, \hat{a}_P, \hat{\omega}_1, \hat{\upsilon}_1, \ldots, \hat{\omega}_P, \hat{\upsilon}_P),$$

which is a strongly consistent estimator of

$$\theta_P^0 = (a_1^0, \ldots, a_P^0, \omega_1^0, \upsilon_1^0, \ldots, \omega_P^0, \upsilon_P^0),$$

as $\Psi(S, T) \to \infty$; of the pair of spatial frequencies $(\hat{\omega}_{P+1}, \hat{\upsilon}_{P+1})$ that maximizes the periodogram of the observed realization of the field $\{u(n, m)\}$, i.e.,

$$(\hat{\omega}_{P+1}, \hat{\upsilon}_{P+1}) = \underset{(\omega,\upsilon)\in(0,2\pi)^2}{\arg\max} \; I_u(\omega, \upsilon) \tag{15}$$

and of the element $\hat{a}_{P+1}$ that satisfies

$$|\hat{a}_{P+1}|^2 = \frac{1}{ST} I_u(\hat{\omega}_{P+1}, \hat{\upsilon}_{P+1}) \, . \tag{16}$$

Proof: Let

$$\theta_{P+1} = (a_1, \ldots, a_P, a_{P+1}, \omega_1, \upsilon_1, \ldots, \omega_P, \upsilon_P, \omega_{P+1}, \upsilon_{P+1}),$$

be some vector in the parameter space Θ_{P+1}. The LS function with $P + 1$ regressors will be denoted $\mathcal{L}_{P+1}$ and the LS function with P regressors will be denoted $\mathcal{L}_P$. We have,

$$\begin{aligned}
\mathcal{L}_{P+1}(\theta_{P+1}) &= \tfrac{1}{ST} \sum_{n=0}^{S-1} \sum_{m=0}^{T-1} \left| y(n, m) - \sum_{i=1}^{P+1} a_i e^{j(n\omega_i + m\upsilon_i)} \right|^2 \\
&= \tfrac{1}{ST} \sum_{n=0}^{S-1} \sum_{m=0}^{T-1} \left| y(n, m) - \sum_{i=1}^{P} a_i e^{j(n\omega_i + m\upsilon_i)} \right|^2 \\
&+ \tfrac{1}{ST} \sum_{n=0}^{S-1} \sum_{m=0}^{T-1} \left| a_{P+1} e^{j(n\omega_{P+1} + m\upsilon_{P+1})} \right|^2 \\
&- 2\Re\Big\{ \tfrac{1}{ST} \sum_{n=0}^{S-1} \sum_{m=0}^{T-1} \Big(y(n, m) - \sum_{i=1}^{P} a_i e^{j(n\omega_i + m\upsilon_i)} \Big) \cdot \\
&\cdot \Big(a_{P+1} e^{j(n\omega_{P+1} + m\upsilon_{P+1})} \Big)^* \Big\} \\
&= H_1(\theta_{P+1}) + H_2(\theta_{P+1}) + H_3(\theta_{P+1}).
\end{aligned} \tag{17}$$

where,

$$\begin{aligned}
H_1(\theta_{P+1}) &= \mathcal{L}_P(a_1, \ldots, a_P, \omega_1, \upsilon_1, \ldots, \omega_P, \upsilon_P) \\
&= \mathcal{L}_P(\theta_P),
\end{aligned} \tag{18}$$

where, $\theta_P = (a_1, \ldots, a_P, \omega_1, \upsilon_1, \ldots, \omega_P, \upsilon_P) \in \Theta_P$ and,

$$\begin{aligned}
&H_2(\theta_{P+1}) = |a_{P+1}|^2 \\
&- 2\Re\Big\{ \frac{1}{ST} \sum_{n=0}^{S-1} \sum_{m=0}^{T-1} u(n, m) \left(a_{P+1} e^{j(n\omega_{P+1} + m\upsilon_{P+1})} \right)^* \Big\},
\end{aligned} \tag{19}$$

$$\begin{aligned}
H_3(\theta_{P+1}) = -2\Re\Big\{ \frac{1}{ST} \sum_{n=0}^{S-1} \sum_{m=0}^{T-1} \Big(\sum_{i=1}^{P} a_i^0 e^{j(n\omega_i^0 + m\upsilon_i^0)} \\
- \sum_{i=1}^{P} a_i e^{j(n\omega_i + m\upsilon_i)} \Big) \left(a_{P+1} e^{j(n\omega_{P+1} + m\upsilon_{P+1})} \right)^* \Big\}.
\end{aligned} \tag{20}$$

Let $\hat{\theta}_P = (\hat{a}_1, \ldots, \hat{a}_P, \hat{\omega}_1, \hat{\upsilon}_1, \ldots, \hat{\omega}_P, \hat{\upsilon}_P)$ be a vector in Θ_P that minimizes $H_1(\theta_{P+1}) = \mathcal{L}_P(\theta_P)$. From [7] (or using Theorem 1 in the previous section),

$$\hat{\theta}_P \to \theta_P^0 \quad a.s. \quad \text{as} \quad \Psi(S, T) \to \infty. \tag{21}$$

The function H_2 is a function of $a_{P+1}, \omega_{P+1}, \upsilon_{P+1}$ only. Evaluating the partial derivatives of H_2 with respect to these variables, it is easy to verify that the extremum points of H_2 are also the extremum points of the periodogram of the realization of the noise field. Moreover, let $a^e, \omega^e, \upsilon^e$ denote an extremum point of H_2. Then at this point

$$H_2(a^e, \omega^e, \upsilon^e) = -\frac{I_u(\omega^e, \upsilon^e)}{ST}. \tag{22}$$

Hence, the minimal value of H_2 is obtained at the coordinates $a_{P+1}, \omega_{P+1}, \upsilon_{P+1}$ where the periodogram of $\{u(n, m)\}$ is maximal. Let $\hat{a}_{P+1}, \hat{\omega}_{P+1}, \hat{\upsilon}_{P+1}$ denote the coordinates that minimize H_2. Then we have

$$\begin{aligned}
(\hat{\omega}_{P+1}, \hat{\upsilon}_{P+1}) &= \underset{(\omega,\upsilon)\in(0,2\pi)^2}{\arg\min} \; H_2(a_{P+1}, \omega_{P+1}, \upsilon_{P+1}) \\
&= \underset{(\omega,\upsilon)\in(0,2\pi)^2}{\arg\max} \; I_u(\omega, \upsilon),
\end{aligned} \tag{23}$$

and

$$\hat{a}_{P+1} = \frac{1}{ST} \sum_{n=0}^{S-1} \sum_{m=0}^{T-1} u(n, m) e^{-j(n\hat{\omega}_{P+1} + m\hat{\upsilon}_{P+1})}. \tag{24}$$

By Assumption 4 and Theorem 2.2, [4], we have

$$\sup_{\omega,\upsilon} I_u(\omega,\upsilon) = O(\log ST). \tag{25}$$

Therefore,

$$H_2(\hat{a}_{P+1},\hat{\omega}_{P+1},\hat{\upsilon}_{P+1}) = O\left(\frac{\log ST}{ST}\right). \tag{26}$$

Let $\hat{\theta}_{P+1} \in \Theta_{P+1}$ be the vector composed of the elements of the vector $\hat{\theta}_P \in \Theta_P$ and of $\hat{a}_{P+1},\hat{\omega}_{P+1},\hat{\upsilon}_{P+1}$, defined above, *i.e.*,

$$\hat{\theta}_{P+1} = (\hat{a}_1,\ldots,\hat{a}_P,\hat{a}_{P+1},\hat{\omega}_1,\hat{\upsilon}_1,\ldots,\hat{\omega}_P,\hat{\upsilon}_P,\hat{\omega}_{P+1},\hat{\upsilon}_{P+1}).$$

We need to verify that this vector minimizes $\mathcal{L}_{P+1}(\theta_{P+1})$ on Θ_{P+1} as $\Psi(S,T)\to\infty$.

Recall that for $\omega \in (0,2\pi)$

$$\sum_{n=0}^{N-1} \exp(j\omega n) = O(1). \tag{27}$$

Hence, as $N\to\infty$

$$\frac{1}{\log N}\sum_{n=0}^{N-1} \exp(j\omega n) = o(1) \tag{28}$$

and consequently

$$\frac{1}{N}\sum_{n=0}^{N-1} \exp(j\omega n) = o\left(\frac{\log N}{N}\right) \tag{29}$$

Next, we evaluate H_3. Consider the first term in (20). By (29) and unless there exists some i, $1\le i\le P$, such that $(\omega_{P+1},\upsilon_{P+1}) = (\omega_i^0,\upsilon_i^0)$, we have as $\Psi(S,T)\to\infty$,

$$\frac{1}{ST}\sum_{n=0}^{S-1}\sum_{m=0}^{T-1}\sum_{i=1}^{P} a_i^0 e^{j(n\omega_i^0+m\upsilon_i^0)}\left(a_{P+1}e^{j(n\omega_{P+1}+m\upsilon_{P+1})}\right)^* = o\left(\frac{\log ST}{ST}\right), \tag{30}$$

for *any* set of values $a_{P+1},\omega_{P+1},\upsilon_{P+1}$ may assume.

Assume now that there exists some i, $1\le i\le P$, such that $(\omega_{P+1},\upsilon_{P+1}) = (\omega_i^0,\upsilon_i^0)$. Since by assumption there are no two different regressors with identical spatial frequencies, it follows that one of the estimated frequencies (ω_i,υ_i) is due to noise contribution. As indexing of the components is arbitrary, by interchanging the roles of $(\omega_{P+1},\upsilon_{P+1})$ and (ω_i,υ_i), and repeating the above argument we conclude that this term has the same order as in (30). Similarly, for the second term in (20): By (29) and unless there exists some i, $1\le i\le P$, such that $(\omega_{P+1},\upsilon_{P+1}) = (\omega_i,\upsilon_i)$, we have as $\Psi(S,T)\to\infty$,

$$\frac{1}{ST}\sum_{n=0}^{S-1}\sum_{m=0}^{T-1}\sum_{i=1}^{P} a_i e^{j(n\omega_i+m\upsilon_i)}\left(a_{P+1}e^{j(n\omega_{P+1}+m\upsilon_{P+1})}\right)^* = o\left(\frac{\log ST}{ST}\right). \tag{31}$$

However such i for which $(\omega_{P+1},\upsilon_{P+1}) = (\omega_i,\upsilon_i)$ cannot exist, as this amounts to reducing the number of regressors from $P+1$ to P, as two of them coincide. Hence, for *any* $\theta_{P+1}\in\Theta_{P+1}$ as $\Psi(S,T)\to\infty$

$$H_3(\theta_{P+1}) = o\left(\frac{\log ST}{ST}\right). \tag{32}$$

On the other hand, the strong consistency (21) of the LSE under the correct model order assumption implies that as $\Psi(S,T)\to\infty$ the minimal value of $\mathcal{L}_P(\theta_P)=\sigma^2$ a.s., while from (26) we have for the minimal value of H_2 that $H_2(\theta_{P+1}) = O\left(\frac{\log ST}{ST}\right)$. Hence, the value of $H_3(\theta_{P+1})$ at *any* point in Θ_{p+1} is negligible even relative to the values $\mathcal{L}_P(\theta_P)$ and $H_2(\theta_{P+1})$ assume at their respective *minimum* points. Therefore, evaluating (17) as $\Psi(S,T)\to\infty$ we have

$$\begin{aligned}&\mathcal{L}_{P+1}(\theta_{P+1})\\ &= \mathcal{L}_P(\theta_P) + H_2(a_{P+1},\omega_{P+1},\upsilon_{P+1}) + H_3(\theta_{P+1})\\ &= \mathcal{L}_P(\theta_P) + H_2(a_{P+1},\omega_{P+1},\upsilon_{P+1}) + o\left(\frac{\log ST}{ST}\right).\end{aligned} \tag{33}$$

Since $\mathcal{L}_P(\theta_P)$ is a function of the parameter vector θ_P and is independent of a_{P+1}, ω_{P+1}, υ_{P+1}, while H_2 is a function of $a_{P+1},\omega_{P+1},\upsilon_{P+1}$ and is independent of θ_P, the problem of minimizing $\mathcal{L}_{P+1}(\theta_{P+1})$ becomes *separable* as $\Psi(S,T)\to\infty$. Thus minimizing (33) is equivalent to separately minimizing $\mathcal{L}_P(\theta_P)$ and $H_2(a_{P+1},\omega_{P+1},\upsilon_{P+1})$ as $\Psi(S,T)\to\infty$. Using the foregoing conclusions, the theorem follows. ■

5. REFERENCES

[1] K. L. Chung, *A Course in Probability Theory*, Harcout, Brace & World, 1968.

[2] J. M. Francos, A. Narashimhan, and J. W. Woods, "Maximum-Likelihood Parameter Estimation of Discrete Homogeneous Random Fields with Mixed Spectral Distribution," *IEEE Trans. Signal Processing*, 44 (1996), 1242-1255.

[3] J. M. Francos, A. Narashimhan, and J. W. Woods, "Maximum-Likelihood Estimation of Textures Using a Wold Decomposition Model," *IEEE Trans. Image Processing*, 4 (1995), 1655-1666.

[4] S. He, "Uniform Convergency for Weighted Periodogram of Stationary Linear Random Fields," *Chin. Ann. of Math.*, 16B (1995), 331-340.

[5] M. Kliger and J. M. Francos, "MAP Model Order Selection Rule for 2-D Sinusoids in White Noise," accepted for publication, *IEEE Trans. Signal Processing*.

[6] M. Kliger and J. M. Francos, "Strong Consistency of the Over- and Under-Determined LSE of 2-D Exponentials in White Noise," submitted to *IEEE Trans. Inform. Theory*.

[7] D. Kundu and R. D. Gupta, "Asymptotic Properties of the Least Squares Estimates of a Two Dimensional Model," *Metrika*, 48 (1998), 83-97.

[8] C. R. Rao, L. Zhao and B. Zhou, "Maximum Likelihood Estimation of 2-D Superimposed Exponential Signals," *IEEE Trans. Signal Processing*, 42 (1994), 1795-1802.

[9] J. Ward, *Space-Time Adaptive Processing for Airborne Radar*, Technical Report 1015, Lincoln Laboratory, Massachusetts Institute of Technology, 1994.

[10] C.F.J. Wu, "Asymptotic Theory of Non-Linear Least Squares Estimation," *Ann. Stat.*, 9 (1981), 510-513.