Strongly Consistent Model Order Selection for Estimating 2-D Sinusoids in Colored Noise

Mark Kliger and Joseph M. Francos *

Abstract

We consider the problem of jointly estimating the number as well as the parameters of two-dimensional sinusoidal signals, observed in the presence of an additive colored noise field. We begin by elaborating on the least squares estimation of 2-D sinusoidal signals, when the assumed number of sinusoids is incorrect. In the case where the number of sinusoidal signals is under-estimated we show the almost sure convergence of the least squares estimates to the parameters of the dominant sinusoids. In the case where this number 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 sinusoids. Based on these results, we prove the strong consistency of a new model order selection rule.

Keywords: Two-dimensional random fields; model order selection; least squares estimation; strong consistency.

*M. Kliger is with the Department of Electrical Engineering and Computer Science , University of Michigan, Ann Arbor, MI, 48109-2122, USA. Tel: (734) 647-8389 FAX: (734) 763-8041, email: mkliger@umich.edu.

J. M. Francos is with the Department of Electrical and Computer Engineering, Ben-Gurion University, Beer-Sheva 84105, Israel. Tel: +972 8 6461842, FAX: +972 8 6472949, email: francos@ee.bgu.ac.il

1 Introduction

We consider the problem of jointly estimating the number as well as the parameters of two-dimensional sinusoidal signals, observed in the presence of an additive noise field. This problem is, in fact, a special case of a much more general problem, [5]: From the 2-D Wold-like decomposition we have that any 2-D regular and homogeneous discrete random field can be represented as a sum of two mutually orthogonal components: a purely-indeterministic field and a deterministic one. In this paper we consider the special case where the deterministic component consists of a finite (unknown) number of sinusoidal components, while the purely-indeterministic component is an infinite order non-symmetrical half plane, (or a quarter-plane), moving average field. This modeling and estimation problem has fundamental theoretical importance, as well as various applications in texture estimation of images (see, e.g., [4] and the references therein) and in wave propagation problems (see, e.g., [14] and the references therein).

Many algorithms have been devised to estimate the parameters of sinusoids observed in white noise and only a small fraction of the derived methods has been extended to the case where the noise field is colored (see, *e.g.*, Francos *et. al.* [3], He [8], Kundu and Nandi [11], Li and Stoica [12], Zhang and Mandrekar [13], and the references therein). Most of these assume the number of sinusoids is *a-priori* known. However this assumption does not always hold in practice. In the past three decades the problem of model order selection for 1-D signals has received considerable attention. In general, model order selection rules are based (directly or indirectly) on three popular criteria: Akaike information criterion (AIC), the minimum description length (MDL), and the maximum a-posteriori probability criterion (MAP). All these criteria have a common form composed of two terms: a data term and a penalty term, where the data term is the log-likelihood function evaluated for the assumed model. The problem of modelling multidimensional fields has received much less attention. In [9], a MAP model order selection criterion for jointly estimating the number and the parameters of two-dimensional sinusoids observed in the presence of an additive white Gaussian noise field, is derived. In [10], we proved the strong consistency of a large family of model order selection rules, which includes the MAP based rule in [9] as a special case.

In this paper we derive a strongly consistent model order selection rule, for jointly estimating the number of sinusoidal components and their parameters in the presence of colored noise. This derivation extends the results of [10] to the case where the additive noise is colored, modeled by an infinite order non-symmetrical half-plane or quarter-plane moving average representation, such that the noise field is not necessarily Gaussian. To the best of our knowledge this is the most general result available in the area of model-order selection rules of 2-D random fields with mixed spectrum.

The proposed criterion has the usual form of a data term and a penalty term, where the first is the *least squares estimator* evaluated for the assumed model order and the latter is proportional to the logarithm of the data size.

Since we evaluate the data term for any assumed model order, including incorrect ones, we should consider the problem of least squares estimation of the parameters of 2-D sinusoidal signals when the assumed number of sinusoids is incorrect. Let P denote the number of sinusoidal signals in the observed field and let k denote their assumed number. In the case where the number of sinusoidal 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 sinusoids. In the case where the number of sinusoidal 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 sinusoids in the observed field. The additional $k - P$ components assumed to exist, are assigned by the least squares estimator to the dominant components of the periodogram of the noise field.

Finally, using this result, we prove the strong consistency of a new model order selection criterion and show how different assumptions regarding a noise field parameters affect the penalty term of the criterion. The proposed criterion completely generalized the previous results [9], [10], and provides a strongly consistent estimator of the number as well as of the parameters of the sinusoidal components.

2 Notations, Definitions and Assumptions

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

$$y(n, m) = \sum_{i=1}^{P} \rho_i^0 \cos(\omega_i^0 n + \upsilon_i^0 m + \varphi_i^0) + w(n, m), \tag{1}$$

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

Recall that the *non-symmetrical half-plan total-order* is defined by

$$(i, j) \succeq (s, t) \text{ iff } (i, j) \in \{(k, l)|k = s, l \geq t\} \cup \{(k, l)|k > s, -\infty \leq l \leq \infty\}. \tag{2}$$

Let D be an *infinite* order non-symmetrical half-plane support, defined by

$$D = \left\{(i, j) \in \mathbb{Z}^2 : i = 0, 0 \leq j \leq \infty\right\} \cup \left\{(i, j) \in \mathbb{Z}^2 : 0 < i \leq \infty, -\infty \leq j \leq \infty\right\}. \tag{3}$$

Hence the notations $(r, s) \in D$ and $(r, s) \succeq (0, 0)$ are equivalent.

We assume that $\{w(n, m)\}$ is an infinite order non-symmetrical half-plane MA noise field, *i.e.*,

$$w(n, m) = \sum_{(r,s) \in D} a(r, s)u(n - r, m - s), \tag{4}$$

such that the following assumptions are satisfied:

Assumption 1: The field $\{u(n, m)\}$ is an i.i.d. real valued zero-mean random field with finite variance σ^2, such that $E[|u(n, m)|^\alpha] < \infty$ for some $\alpha > 3$.

Assumption 2: The sequence $a(i, j)$ is an absolutely summable deterministic sequence, *i.e.*,

$$\sum_{(r,s) \in D} |a(r, s)| < \infty. \tag{5}$$

Let $f_w(\omega, v)$ denote the spectral density function of the noise field $\{w(n, m)\}$. Hence,

$$f_w(\omega, v) = \sigma^2 \left| \sum_{(r,s) \in D} a(r, s)e^{j(\omega r + vs)} \right|^2. \tag{6}$$

Assumption 3: The spatial frequencies $(\omega_i^0, v_i^0) \in (0, 2\pi) \times (0, 2\pi)$, $1 \le i \le P$ are pairwise different. In other words, $\omega_i^0 \ne \omega_j^0$ or $v_i^0 \ne v_j^0$, when $i \ne j$.

Let $\{\Psi_i\}$ be a sequence of rectangles such that $\Psi_i = \{(n, m) \in \mathbb{Z}^2 \mid 0 \le n \le N_i - 1, 0 \le m \le M_i - 1\}$.

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

$$\lim_{i \to \infty} \min(N_i, M_i) = \infty,$$

and

$$0 < \lim_{i \to \infty} (N_i / M_i) < \infty.$$

To simplify notations, we shall omit in the following the subscript i. Thus, the notation $\Psi(N, M) \to \infty$ implies that both N and M 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 $4k$ dimensional space $\mathbb{R}^k \times ((0, 2\pi) \times (0, 2\pi))^k \times [0, 2\pi)^k$ where for any vector $\theta_k = (\rho_1, \omega_1, v_1, \varphi_1, \ldots, \rho_k, \omega_k, v_k, \varphi_k) \in \Theta_k$ the coordinate ρ_i is non-zero and bounded for every $1 \le i \le k$ while the pairs (ω_i, v_i) are pairwise different, so that no two regressors coincide. We shall refer to Θ_k as the *parameter space*.

From the model definition (1) and the above assumptions it is clear that

$$\theta_k^0 = (\rho_1^0, \omega_1^0, v_1^0, \varphi_1^0, \ldots, \rho_k^0, \omega_k^0, v_k^0, \varphi_k^0) \in \Theta_k.$$

3

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

$$\mathcal{L}_k(\theta_k) = \frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \left(y(n,m) - \sum_{i=1}^{k} \rho_i^0 \cos(\omega_i^0 n + v_i^0 m + \varphi_i^0) \right)^2. \tag{7}$$

A vector $\hat{\theta}_k \in \Theta_k$ that minimizes $\mathcal{L}_k(\theta_k)$ 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.*, [11] and the references therein).

3   Strong Consistency of the Over- and Under-Determined LSE

In the following subsections we establish the strong consistency of this LSE when the number of sinusoids is under-estimated, or over-estimated. The first 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 sinusoids. The second theorem establishes the strong consistency of the least squares estimator in the case where the number of the regressors is higher than the actual number of sinusoids.

3.1   Consistency of the LSE for an Under-Estimated Model Order

Let k denote the assumed number of observed 2-D sinusoids, 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.*,

$$\Delta_\delta = \left[\bigcup_{i=1}^{k} \mathcal{R}_{i\delta} \right] \cup \left[\bigcup_{i=1}^{k} \Phi_{i\delta} \right] \cup \left[\bigcup_{i=1}^{k} W_{i\delta} \right] \cup \left[\bigcup_{i=1}^{k} V_{i\delta} \right], \tag{8}$$

where

$$\mathcal{R}_{i\delta} = \left\{ \theta_k \in \Theta_k : \ |\rho_i - \rho_i^0| \geq \delta; \delta > 0 \right\},$$
$$\Phi_{i\delta} = \left\{ \theta_k \in \Theta_k : \ |\varphi_i - \varphi_i^0| \geq \delta; \delta > 0 \right\},$$
$$W_{i\delta} = \left\{ \theta_k \in \Theta_k : \ |\omega_i - \omega_i^0| \geq \delta; \delta > 0 \right\},$$
$$V_{i\delta} = \left\{ \theta_k \in \Theta_k : \ |v_i - v_i^0| \geq \delta; \delta > 0 \right\}. \tag{9}$$

4

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

Assumption 4: For convenience, and without loss of generality, we assume that the sinusoids are indexed according to a descending order of their amplitudes, *i.e.*,

$$\rho_1^0 \geq \rho_2^0 \geq \ldots \rho_k^0 > \rho_{k+1}^0 \ldots \geq \rho_P^0 > 0 , \tag{10}$$

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

Lemma 1.

$$\liminf_{\Psi(N,M)\to\infty} \inf_{\theta_k \in \Delta_\delta} \left(\mathcal{L}_k(\theta_k) - \mathcal{L}_k(\theta_k^0) \right) > 0 \quad a.s. \tag{11}$$

Proof: See Appendix A for the proof. ∎

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

$$\Pr\{x_n \leq 0 \ i.o.\} \leq \Pr\{\liminf_{n\to\infty} x_n \leq 0\}, \tag{12}$$

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

Proof: See Appendix B 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 sinusoids.

Theorem 1. *Let Assumptions 1-4 be satisfied. Then, the k-regressor parameter vector $\hat{\theta}_k = (\hat{\rho}_1, \hat{\omega}_1, \hat{v}_1, \hat{\varphi}_1, \ldots, \hat{\rho}_k, \hat{\omega}_k, \hat{v}_k, \hat{\varphi}_k)$ that minimizes (7) is a strongly consistent estimator of $\theta_k^0 = (\rho_1^0, \omega_1^0, v_1^0, \varphi_1^0, \ldots, \rho_k^0, \omega_k^0, v_k^0, \varphi_k^0)$ as $\Psi(N, M) \to \infty$. That is,*

$$\hat{\theta}_k \to \theta_k^0 \ a.s. \ as \ \Psi(N, M) \to \infty. \tag{13}$$

Proof: The proof follows an argument proposed by Wu [15], Lemma 1. Let $\hat{\theta}_k = (\hat{\rho}_1, \hat{\omega}_1, \hat{v}_1, \hat{\varphi}_1, \ldots, \hat{\rho}_k, \hat{\omega}_k, \hat{v}_k, \hat{\varphi}_k)$ be a parameter vector that minimizes (7). Assume that the proposition $\hat{\theta}_k \to \theta_k^0 \ a.s. \ as \ \Psi(N, M) \to \infty$ is not true. Then, there exists some $\delta > 0$, such that ([1], Theorem 4.2.2, p. 69),

$$\Pr(\hat{\theta}_k \in \Delta_\delta \ i.o.) > 0. \tag{14}$$

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

$$\Pr(\inf_{\theta_k \in \Delta_\delta} \left(\mathcal{L}_k(\theta_k) - \mathcal{L}_k(\theta_k^0) \right) \leq 0 \ i.o.) > 0. \tag{15}$$

Using Lemma 2 we obtain

$$\Pr\left(\liminf_{\Psi(N,M)\to\infty}\inf_{\theta_k\in\Delta_\delta}\left(\mathcal{L}_k(\theta_k)-\mathcal{L}_k(\theta_k^0)\right)\le 0\right)\ge\Pr\left(\inf_{\theta_k\in\Delta_\delta}\left(\mathcal{L}_k(\theta_k)-\mathcal{L}_k(\theta_k^0)\right)\le 0 \ i.o.\right)>0, \quad (16)$$

which contradicts (11). Hence,

$$\hat{\theta}_k\to\theta_k^0 \ a.s. \ as \ \Psi(N,M)\to\infty. \tag{17}$$

∎

Remark: Lemma 1 and Theorem 1 remain valid even under less restrictive assumptions regarding the noise field $\{w(n,m)\}$. If the field $\{u(n,m)\}$ is an i.i.d. real valued zero-mean random field with finite variance σ^2, and the sequence $a(i,j)$ is a square summable deterministic sequence, $i.e.$, $\sum_{(r,s)\in D}a^2(r,s)<\infty$, then Lemma 1 and Theorem 1 hold.

3.2 Consistency of the LSE for an Over-Estimated Model Order

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

$$I_w(\omega,\upsilon)=\frac{2}{NM}\left|\sum_{n=0}^{N-1}\sum_{m=0}^{M-1}w(n,m)e^{-j(n\omega+m\upsilon)}\right|^2. \tag{18}$$

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

Theorem 2. *Let Assumptions 1-4 be satisfied. Then, the parameter vector*
$\hat{\theta}_{P+1}=(\hat{\rho}_1,\hat{\omega}_1,\hat{\upsilon}_1,\hat{\varphi}_1,\ldots,\hat{\rho}_P,\hat{\omega}_P,\hat{\upsilon}_P,\hat{\varphi}_P,\hat{\rho}_{P+1},\hat{\omega}_{P+1},\hat{\upsilon}_{P+1},\hat{\varphi}_{P+1})\in\Theta_{P+1}$ *that minimizes (7) with*
$k=P+1$ *regressors as* $\Psi(N,M)\to\infty$ *is composed of the vector* $\hat{\theta}_P=(\hat{\rho}_1,\hat{\omega}_1,\hat{\upsilon}_1,\hat{\varphi}_1,\ldots,\hat{\rho}_P,\hat{\omega}_P,\hat{\upsilon}_P,\hat{\varphi}_P)$
which is a strongly consistent estimator of $\theta_P^0=(\rho_1^0,\omega_1^0,\upsilon_1^0,\varphi_1^0,\ldots,\rho_P^0,\omega_P^0,\upsilon_P^0,\varphi_P^0)$ *as* $\Psi(N,M)\to$
∞; *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* $\{w(n,m)\}$, *i.e.*,

$$(\hat{\omega}_{P+1},\hat{\upsilon}_{P+1})=\underset{(\omega,\upsilon)\in(0,2\pi)^2}{\arg\max}I_w(\omega,\upsilon), \tag{19}$$

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

$$\hat{\rho}_{P+1}^2=\frac{2}{NM}I_w(\hat{\omega}_{P+1},\hat{\upsilon}_{P+1}) . \tag{20}$$

6

Proof: Let $\theta_{P+1} = (\rho_1, \omega_1, \upsilon_1, \varphi_1, \ldots, \rho_P, \omega_P, \upsilon_P, \varphi_P, \rho_{P+1}, \omega_{P+1}, \upsilon_{P+1}, \varphi_{P+1})$, be some vector in the parameter space Θ_{P+1}. We have,

$$
\begin{aligned}
\mathcal{L}_{P+1}(\theta_{P+1}) &= \frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \left(y(n,m) - \sum_{i=1}^{P+1} \rho_i \cos(\omega_i n + \upsilon_i m + \varphi_i) \right)^2 \\
&= \frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \left(y(n,m) - \sum_{i=1}^{P} \rho_i \cos(\omega_i n + \upsilon_i m + \varphi_i) \right)^2 \\
&\quad + \frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \left(\rho_{P+1} \cos(\omega_{P+1} n + \upsilon_{P+1} m + \varphi_{P+1}) \right)^2 \\
&\quad - \frac{2}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \left(y(n,m) - \sum_{i=1}^{P} \rho_i \cos(\omega_i n + \upsilon_i m + \varphi_i) \right) \left(\rho_{P+1} \cos(\omega_{P+1} n + \upsilon_{P+1} m + \varphi_{P+1}) \right) \\
&= \mathcal{L}_P(\theta_P) + \frac{\rho_{P+1}^2}{2} + \frac{1}{2NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \rho_{P+1}^2 \cos(2\omega_{P+1} n + 2\upsilon_{P+1} m + 2\varphi_{P+1}) \\
&\quad - \frac{2}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} w(n,m) \rho_{P+1} \cos(\omega_{P+1} n + \upsilon_{P+1} m + \varphi_{P+1}) \\
&\quad - \frac{2}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \left(\sum_{i=1}^{P} \rho_i^0 \cos(\omega_i^0 n + \upsilon_i^0 m + \varphi_i^0) - \sum_{i=1}^{P} \rho_i \cos(\omega_i n + \upsilon_i m + \varphi_i) \right) \\
&\quad \left(\rho_{P+1} \cos(\omega_{P+1} n + \upsilon_{P+1} m + \varphi_{P+1}) \right) = H_1(\theta_{P+1}) + H_2(\theta_{P+1}) + H_3(\theta_{P+1})
\end{aligned}
$$
$$(21)$$

where, $\theta_P = (\rho_1, \omega_1, \upsilon_1, \varphi_1, \ldots, \rho_P, \omega_P, \upsilon_P, \varphi_P) \in \Theta_P$ and,

$$
H_1(\theta_{P+1}) = \mathcal{L}_P(\rho_1, \omega_1, \upsilon_1, \varphi_1, \ldots, \rho_P, \omega_P, \upsilon_P, \varphi_P) = \mathcal{L}_P(\theta_P), \tag{22}
$$

$$
H_2(\theta_{P+1}) = \frac{\rho_{P+1}^2}{2} - \frac{2}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} w(n,m) \rho_{P+1} \cos(\omega_{P+1} n + \upsilon_{P+1} m + \varphi_{P+1}), \tag{23}
$$

$$
\begin{aligned}
H_3(\theta_{P+1}) &= \frac{1}{2NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \rho_{P+1}^2 \cos(2\omega_{P+1} n + 2\upsilon_{P+1} m + 2\varphi_{P+1}) \\
&\quad - \frac{2}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \left(\sum_{i=1}^{P} \rho_i^0 \cos(\omega_i^0 n + \upsilon_i^0 m + \varphi_i^0) - \sum_{i=1}^{P} \rho_i \cos(\omega_i n + \upsilon_i m + \varphi_i) \right) \\
&\quad \left(\rho_{P+1} \cos(\omega_{P+1} n + \upsilon_{P+1} m + \varphi_{P+1}) \right).
\end{aligned}
$$
$$(24)$$

Let $\hat{\theta}_P = (\hat{\rho}_1, \hat{\omega}_1, \hat{\upsilon}_1, \hat{\varphi}_1, \ldots, \hat{\rho}_P, \hat{\omega}_P, \hat{\upsilon}_P, \hat{\varphi}_P)$ be a vector in Θ_P that minimizes $H_1(\theta_{P+1}) = \mathcal{L}_P(\theta_P)$. From [11] (or using Theorem 1 in the previous section),

$$
\hat{\theta}_P \to \theta_P^0 \quad a.s. \quad \text{as} \quad \Psi(N, M) \to \infty. \tag{25}
$$

The function H_2 is a function of $\rho_{P+1}, \omega_{P+1}, \upsilon_{P+1}, \varphi_{P+1}$ 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

7

extremum points of the periodogram of the realization of the noise field. Moreover, let $\rho^e, \omega^e, \upsilon^e, \varphi^e$ denote an extremum point of H_2. Then at this point

$$H_2(\rho^e, \omega^e, \upsilon^e, \varphi^e) = -\frac{I_w(\omega^e, \upsilon^e)}{NM}. \tag{26}$$

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

$$(\hat{\omega}_{P+1}, \hat{\upsilon}_{P+1}) = \underset{(\omega,\upsilon)\in(0,2\pi)^2}{\arg\min} H_2(\rho_{P+1}, \omega_{P+1}, \upsilon_{P+1}, \varphi_{P+1}) = \underset{(\omega,\upsilon)\in(0,2\pi)^2}{\arg\max} I_w(\omega, \upsilon), \tag{27}$$

and

$$\hat{\rho}_{P+1}^2 = \frac{2}{NM} I_w(\hat{\omega}_{P+1}, \hat{\upsilon}_{P+1}). \tag{28}$$

By Assumption 1, 2 and Theorem 1, [13], we have

$$\sup_{\omega,\upsilon} I_w(\omega, \upsilon) = O(\log NM). \tag{29}$$

Therefore,

$$H_2(\hat{\rho}_{P+1}, \hat{\omega}_{P+1}, \hat{\upsilon}_{P+1}, \hat{\varphi}_{P+1}) = O\left(\frac{\log NM}{NM}\right). \tag{30}$$

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 $\hat{\rho}_{P+1}, \hat{\omega}_{P+1}, \hat{\upsilon}_{P+1}, \hat{\varphi}_{P+1}$, defined above, $i.e.$,

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

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

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

$$\sum_{n=0}^{N-1} \cos(\omega n + \varphi) = \frac{\sin\left([N - \frac{1}{2}]\omega + \varphi\right) + \sin\left(\frac{\omega}{2} - \varphi\right)}{2\sin\left(\frac{\omega}{2}\right)} = O(1). \tag{31}$$

Hence, as $N \to \infty$

$$\frac{1}{\log N} \sum_{n=0}^{N-1} \cos(\omega n + \varphi) = o(1), \tag{32}$$

and consequently

$$\frac{1}{N} \sum_{n=0}^{N-1} \cos(\omega n + \varphi) = o\left(\frac{\log N}{N}\right). \tag{33}$$

8

Next, we evaluate H_3. Consider the first term in (24). By (33) we have

$$\frac{1}{2NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \rho_{P+1}^2 \cos(2\omega_{P+1}n + 2\upsilon_{P+1}m + 2\varphi_{P+1}) = o\left(\frac{\log NM}{NM}\right), \tag{34}$$

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

Consider the second term in (24). By (33) and unless there exists some i, $1 \leq i \leq P$, such that $(\omega_{P+1}, \upsilon_{P+1}) = (\omega_i^0, \upsilon_i^0)$, we have as $\Psi(N, M) \to \infty$,

$$\frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \sum_{i=1}^{P} \rho_i^0 \rho_{P+1} \cos(\omega_i^0 n + \upsilon_i^0 m + \varphi_i^0) \cos(\omega_{P+1}n + \upsilon_{P+1}m + \varphi_{P+1}) = o\left(\frac{\log NM}{NM}\right), \tag{35}$$

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

Assume now that there exists some i, $1 \leq i \leq P$, 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. Hence, 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 (35). Similarly, for the third term in (24): By (33) and unless there exists some i, $1 \leq i \leq P$, such that $(\omega_{P+1}, \upsilon_{P+1}) = (\omega_i, \upsilon_i)$, we have as $\Psi(N, M) \to \infty$,

$$\frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \sum_{i=1}^{P} \rho_i \rho_{P+1} \cos(\omega_i n + \upsilon_i m + \varphi_i) \cos(\omega_{P+1}n + \upsilon_{P+1}m + \varphi_{P+1}) = o\left(\frac{\log NM}{NM}\right). \tag{36}$$

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 $\Psi(N, M) \to \infty$

$$H_3(\theta_{P+1}) = o\left(\frac{\log NM}{NM}\right). \tag{37}$$

On the other hand, the strong consistency (25) of the LSE under the correct model order assumption implies that as $\Psi(N, M) \to \infty$ the minimal value of $\mathcal{L}_P(\theta_P) = \sigma^2 \sum_{(r,s)\in D} a^2(r, s)$ a.s., while from (30) we have for the minimal value of H_2 that $H_2(\theta_{P+1}) = O\left(\frac{\log NM}{NM}\right)$. 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 (21) as $\Psi(N, M) \to \infty$ we have

$$\begin{aligned} \mathcal{L}_{P+1}(\theta_{P+1}) &= \mathcal{L}_P(\theta_P) + H_2(\rho_{P+1}, \omega_{P+1}, \upsilon_{P+1}, \varphi_{P+1}) + H_3(\theta_{P+1}) \\ &= \mathcal{L}_P(\theta_P) + H_2(\rho_{P+1}, \omega_{P+1}, \upsilon_{P+1}, \varphi_{P+1}) + o\left(\frac{\log NM}{NM}\right). \end{aligned} \tag{38}$$

Since $\mathcal{L}_P(\theta_P)$ is a function of the parameter vector θ_P and is independent of $\rho_{P+1}, \omega_{P+1}, \upsilon_{P+1}, \varphi_{P+1}$, while H_2 is a function of $\rho_{P+1}, \omega_{P+1}, \upsilon_{P+1}, \varphi_{P+1}$ and is independent of θ_P, the problem of minimizing $\mathcal{L}_{P+1}(\theta_{P+1})$ becomes *separable* as $\Psi(N, M) \to \infty$. Thus minimizing (38) is equivalent

9

to separately minimizing $\mathcal{L}_P(\theta_P)$ and $H_2(\rho_{P+1}, \omega_{P+1}, v_{P+1}, \varphi_{P+1})$ as $\Psi(N, M) \to \infty$. Using the foregoing conclusions, the theorem follows. ∎

3.3 Discussion

In the above theorems, we have considered the problem of least squares estimation of the parameters of 2-D sinusoidal signals observed in the presence of an additive colored noise field, when the assumed number of sinusoids is incorrect. In the case where the number of sinusoidal signals is under-estimated we have established the almost sure convergence of the least squares estimates to the parameters of the dominant sinusoids. This result can be intuitively explained using the basic principles of least squares estimation: Since the least squares estimate is the set of model parameters that minimizes the ℓ_2 norm of the error between the observations and the assumed model, it follows that in the case where the model order is under-estimated the minimum error norm is achieved when the k most dominant sinusoids are correctly estimated. Similarly, in the case where the number of sinusoidal signals is over-estimated, the estimated parameter vector obtained by the least squares estimator contains a $4P$-dimensional sub-vector that converges almost surely to the correct parameters of the sinusoids, while the remaining $k - P$ components assumed to exist, are assigned to the $k - P$ most dominant spectral peaks of the noise power to further minimize the norm of the estimation error.

4 Strong Consistency of a Family of Model Order Selection Rules

In this section we employ the results derived in the previous section in order to establish the strong consistency of a new model order selection rule.

It is assumed that there are Q competing models, where Q is finite, $Q > P$, and that each competing model $k \in Z_Q = \{0, 1, 2, \ldots, Q - 1\}$ is equiprobable. Following the MDL-MAP template, define the statistic

$$\chi_\xi(k) = NM \log \mathcal{L}_k(\hat{\theta}_k) + \xi k \log NM, \tag{39}$$

where ξ is some finite constant to be specified later, and $\mathcal{L}_k(\hat{\theta}_k)$ is the minimal value of the error variance of the least squares estimator.

The number of 2-D sinusoids is estimated by minimizing $\chi_\xi(k)$ over $k \in Z_Q$, i.e.,

$$\hat{P} = \arg\min_{k \in Z_Q} \left\{ \chi_\xi(k) \right\}. \tag{40}$$

Let

$$\mathcal{A} := \frac{\sum_{(r,s)\in D} \sum_{(q,t)\in D} |a(r,s)a(q,t)|}{\sum_{(r,s)\in D} a^2(r,s)}. \tag{41}$$

The objective of the next theorem is to prove the asymptotic consistency of the model order selection procedure in (40).

Theorem 3. *Let Assumptions 1-4 be satisfied. Let $\hat{P}$ be given by (40) with $\xi > 14\mathcal{A}$. Then as $\Psi(N,M) \to \infty$*

$$\hat{P} \to P \quad a.s. \tag{42}$$

Proof:

For $k \leq P$,

$$\begin{aligned}
&\chi_\xi(k-1) - \chi_\xi(k) \\
&= NM \log \mathcal{L}_{k-1}(\hat{\theta}_{k-1}) + \xi(k-1)\log NM - NM \log \mathcal{L}_k(\hat{\theta}_k) - \xi k \log NM \\
&= NM \log \left(\frac{\mathcal{L}_{k-1}(\hat{\theta}_{k-1})}{\mathcal{L}_k(\hat{\theta}_k)} \right) - \xi \log NM.
\end{aligned} \tag{43}$$

From Theorem 1 as $\Psi(N,M) \to \infty$

$$\hat{\theta}_k \to \theta_k^0 \quad a.s., \tag{44}$$

and

$$\hat{\theta}_{k-1} \to \theta_{k-1}^0 \quad a.s. \tag{45}$$

From the definition of $\mathcal{L}_k(\hat{\theta}_k)$, and (44)

$$\begin{aligned}
\mathcal{L}_k(\hat{\theta}_k) &= \frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \left(y(n,m) - \sum_{i=1}^k \hat{\rho}_i \cos(\hat{\omega}_i n + \hat{v}_i m + \hat{\varphi}_i) \right)^2 \\
&= \frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \left(\sum_{i=1}^P \rho_i^0 \cos(\omega_i^0 n + v_i^0 m + \varphi_i^0) + w(n,m) - \sum_{i=1}^k \hat{\rho}_i \cos(\hat{\omega}_i n + \hat{v}_i m + \hat{\varphi}_i) \right)^2 \\
&\xrightarrow[\Psi(N,M)\to\infty]{} \frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \left(\sum_{i=k+1}^P \rho_i^0 \cos(\omega_i^0 n + v_i^0 m + \varphi_i^0) + w(n,m) \right)^2.
\end{aligned} \tag{46}$$

11

From Lemma 3 in Appendix C we have that as $\Psi(N, M) \to \infty$

$$\sup_{\omega, \upsilon} \left| \frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} w(n, m) \cos(\omega n + \upsilon m) \right| \to 0 \quad \text{a.s.} \tag{47}$$

Hence, from the Assumption 3, (31), (47) and the Strong Law of Large Numbers, we conclude that as $\Psi(N, M) \to \infty$

$$\mathcal{L}_k(\hat{\theta}_k) \to \sigma^2 \sum_{(r,s) \in D} a^2(r, s) + \sum_{i=k+1}^{P} \frac{(\rho_i^0)^2}{2} \quad \text{a.s.} \tag{48}$$

and similarly

$$\mathcal{L}_{k-1}(\hat{\theta}_{k-1}) \to \sigma^2 \sum_{(r,s) \in D} a^2(r, s) + \sum_{i=k}^{P} \frac{(\rho_i^0)^2}{2} \quad \text{a.s.} \tag{49}$$

Since $\frac{\log NM}{NM}$ tends to zero, as $\Psi(N, M) \to \infty$, then as $\Psi(N, M) \to \infty$

$$(NM)^{-1}(\chi_\xi(k-1) - \chi_\xi(k)) \to \log\left(1 + \frac{(\rho_k^0)^2}{2\sigma^2 \sum_{(r,s) \in D} a^2(r, s) + \sum_{i=k+1}^{P} (\rho_i^0)^2}\right) \text{a.s.} \tag{50}$$

Since $\log\left(1 + \frac{(\rho_k^0)^2}{2\sigma^2 \sum_{(r,s) \in D} a^2(r,s) + \sum_{i=k+1}^{P} (\rho_i^0)^2}\right)$ is strictly positive, then $\chi_\xi(k-1) > \chi_\xi(k)$. Hence, for $k \leq P$, the function $\chi_\xi(k)$ is monotonically decreasing with k.

We next consider the case where $k = P + l$ for any integer $l \geq 1$.

Based on [13], Theorem 1 and Assumptions 1, 2 we have that

$$\limsup_{\Psi(N,M) \to \infty} \frac{\sup_{\omega, \upsilon} I_w(\omega, \upsilon)}{\sup_{\omega, \upsilon} f_w(\omega, \upsilon) \log(NM)} \leq 14 \quad \text{a.s.} \tag{51}$$

Based on an extension of Theorem 2 we have that a.s. as $\Psi(N, M) \to \infty$

$$\mathcal{L}_{P+l}(\hat{\theta}_{P+l}) = \mathcal{L}_P(\hat{\theta}_P) - \frac{U_l}{NM} + o\left(\frac{\log NM}{NM}\right), \tag{52}$$

where

$$U_l = \sum_{i=1}^{l} I_w(\omega_i, \upsilon_i), \tag{53}$$

is the sum of the l largest elements of the periodogram of the noise field $\{w(s, t)\}$. Clearly

$$U_l \leq l \sup_{\omega, \upsilon} I_u(\omega, \upsilon). \tag{54}$$

Similarly to (43), a.s. as $\Psi(N, M) \to \infty$,

$$\chi_\xi(P + l) - \chi_\xi(P)$$
$$= NM \log \mathcal{L}_{P+l}(\hat{\theta}_{P+l}) + \xi(P + l) \log NM - NM \log \mathcal{L}_P(\hat{\theta}_P) - \xi P \log NM$$
$$= \xi l \log NM + NM \log \left(1 - \frac{U_l}{NM\mathcal{L}_P(\hat{\theta}_P)} + o\left(\frac{\log NM}{NM} \right) \right)$$
$$= \xi l \log NM - \left(\frac{U_l}{\mathcal{L}_P(\hat{\theta}_P)} + o(\log NM) \right)(1 + o(1))$$
$$= \log NM \left(\xi l - \frac{U_l}{\mathcal{L}_P(\hat{\theta}_P) \log NM} + o(1) \right) \geq \log NM \left(\xi l - \frac{l \sup\limits_{\omega,\upsilon} I_w(\omega, \upsilon)}{\mathcal{L}_P(\hat{\theta}_P) \log NM} + o(1) \right)$$
$$= l \log NM \left(\xi - \frac{\sup\limits_{\omega,\upsilon} I_w(\omega, \upsilon)}{\sup\limits_{\omega,\upsilon} f_w(\omega, \upsilon) \log NM} \frac{\sup\limits_{\omega,\upsilon} f_w(\omega, \upsilon)}{\mathcal{L}_P(\hat{\theta}_P)} + o(1) \right), \tag{55}$$

where the second equality is obtained by substituting $\mathcal{L}_{P+l}(\hat{\theta}_{P+l})$ using the equality (52). The third equality is due to the property that for $x \to 0$, $\log(1 + x) = x(1 + o(1))$, where the observation that the term $\frac{U_l}{NM\mathcal{L}_P(\hat{\theta}_P)}$ tends to zero a.s. as $\Psi(N, M) \to \infty$ is due to (51).

From [11] (or using Theorem 1 in the previous section),

$$\hat{\theta}_P \to \theta_P^0 \quad a.s. \quad \text{as} \quad \Psi(N, M) \to \infty. \tag{56}$$

Hence, the strong consistency (56) of the LSE under the correct model order assumption implies that as $\Psi(N, M) \to \infty$

$$\mathcal{L}_P(\hat{\theta}_P) \to \sigma^2 \sum_{(r,s)\in D} a^2(r, s) \quad \text{a.s.} \tag{57}$$

On the other hand using the triangle inequality

$$\sup_{\omega,\upsilon} f_w(\omega, \upsilon) \leq \sigma^2 \sum_{(r,s)\in D} \sum_{(q,t)\in D} |a(r, s)a(q, t)|. \tag{58}$$

Substituting (51),(57) and (58) into (55) we conclude that

$$\chi_\xi(P + l) - \chi_\xi(P) > 0 \tag{59}$$

for any integer $l \geq 1$. Therefore, a.s. as $\Psi(N, M) \to \infty$, the function $\chi_\xi(k)$ has a **global minimum** for $k = P$. ∎

5 Special Case

Introducing some additional restrictions on the structure of the noise field, we can establish a tighter (in terms of ξ) model order selection rule. We thus modify our earlier Assumption 1, 2 regarding the noise field as follows:

Assumption 1' The noise field $\{w(n,m)\}$ is an infinite order *quarter-plane* MA field, *i.e.*,

$$w(n,m) = \sum_{r,s=0}^{\infty} a(r,s)u(n-r, m-s) \tag{60}$$

where the field $\{u(n,m)\}$ is an i.i.d. real valued zero-mean random field with finite variance σ^2, such that $E[u(n,m)^2 \log |u(n,m)|] < \infty$.

Assumption 2' The sequence $a(i,j)$ is a deterministic sequence which satisfied the condition

$$\sum_{r,s=0}^{\infty} (r+s)|a(r,s)| < \infty. \tag{61}$$

In this case, based on [7], Theorem 3.2 and Assumption 1', 2' we have that

$$\limsup_{\Psi(N,M)\to\infty} \frac{\sup_{\omega,\upsilon} I_w(\omega,\upsilon)}{\sup_{\omega,\upsilon} f_w(\omega,\upsilon)\log(NM)} \leq 8 \quad \text{a.s.} \tag{62}$$

The results of Theorem 1 and 2 are not affected by this assumption. The only change is in Theorem 3. Therefore we can formulate the next theorem:

Theorem 4. *Let Assumptions 1', 2', 3 and 4 be satisfied. Let $\hat{P}$ be given by (40) with $\xi > 8\mathcal{A}$. Then as $\Psi(N,M) \to \infty$*

$$\hat{P} \to P \quad a.s. \tag{63}$$

The proof of this theorem is identical to the proof of Theorem 3, where instead of (51) we employ the inequality in (62).

6 Conclusions

We have considered the problem of jointly estimating the number as well as the parameters of two-dimensional sinusoidal signals, observed in the presence of an additive colored noise field. We have established the strong consistency of the LSE when the number of sinusoidal signals is

under-estimated, or over-estimated. Based on these results, we have proved the strong consistency of a new model order selection rule for the number of sinusoidal components.

Appendix A

Lemma 1.

$$\liminf_{\Psi(N,M)\to\infty} \inf_{\theta_k\in\Delta_\delta} \left(\mathcal{L}_k(\theta_k) - \mathcal{L}_k(\theta_k^0)\right) > 0 \quad a.s. \tag{64}$$

Proof:

In the following we first show that on Δ_δ the sequence $\mathcal{L}_k(\theta_k) - \mathcal{L}_k(\theta_k^0)$ (indexed in N, M) is uniformly lower bounded by a strictly positive constant as $\Psi(N, M) \to \infty$. Since the sequence elements are uniformly lower bounded by a strictly positive constant the sequence of infimums, $\inf_{\theta_k\in\Delta_\delta} (\mathcal{L}_k(\theta_k) - \mathcal{L}_k(\theta_k^0))$, is uniformly lower bounded by the same strictly positive constant as $\Psi(N, M) \to \infty$, and hence, $\liminf_{\Psi(N,M)\to\infty} \inf_{\theta_k\in\Delta_\delta} (\mathcal{L}_k(\theta_k) - \mathcal{L}_k(\theta_k^0))$.

Thus, we first prove that the sequence $\mathcal{L}_k(\theta_k) - \mathcal{L}_k(\theta_k^0)$ is uniformly lower bounded away from zero on Δ_δ as $\Psi(N, M) \to \infty$.

$$
\begin{aligned}
&\mathcal{L}_k(\theta_k) - \mathcal{L}_k(\theta_k^0) \\
&= \frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \left(y(n,m) - \sum_{i=1}^{k} \rho_i \cos(\omega_i n + \upsilon_i m + \varphi_i) \right)^2 \\
&\quad - \frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \left(y(n,m) - \sum_{i=1}^{k} \rho_i^0 \cos(\omega_i^0 n + \upsilon_i^0 m + \varphi_i^0) \right)^2 \\
&= \frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \left(\sum_{i=1}^{P} \rho_i^0 \cos(\omega_i^0 n + \upsilon_i^0 m + \varphi_i^0) + w(n,m) - \sum_{i=1}^{k} \rho_i \cos(\omega_i n + \upsilon_i m + \varphi_i) \right)^2 \\
&\quad - \frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \left(\sum_{i=k+1}^{P} \rho_i^0 \cos(\omega_i^0 n + \upsilon_i^0 m + \varphi_i^0) + w(n,m) \right)^2 \\
&= \frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \left(\sum_{i=1}^{k} \rho_i^0 \cos(\omega_i^0 n + \upsilon_i^0 m + \varphi_i^0) - \sum_{i=1}^{k} \rho_i \cos(\omega_i n + \upsilon_i m + \varphi_i) \right)^2 \\
&\quad + \frac{2}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \left(\sum_{i=k+1}^{P} \rho_i^0 \cos(\omega_i^0 n + \upsilon_i^0 m + \varphi_i^0) \right) \\
&\qquad \left(\sum_{i=1}^{k} \rho_i^0 \cos(\omega_i^0 n + \upsilon_i^0 m + \varphi_i^0) - \sum_{i=1}^{k} \rho_i \cos(\omega_i n + \upsilon_i m + \varphi_i) \right) \\
&\quad + \frac{2}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} w(n,m) \left(\sum_{i=1}^{k} \rho_i^0 \cos(\omega_i^0 n + \upsilon_i^0 m + \varphi_i^0) - \sum_{i=1}^{k} \rho_i \cos(\omega_i n + \upsilon_i m + \varphi_i) \right) \\
&= I_1 + I_2 + I_3.
\end{aligned}
\tag{65}
$$

Thus, to check the asymptotic behavior of L.H.S. of (65) we have to evaluate $\lim_{\Psi(N,M)\to\infty} (I_1 + I_2 + I_3)$

15

for all vectors $\theta_k \in \Delta_\delta$:

$$
\begin{aligned}
\lim_{\Psi(N,M)\to\infty} I_1 &= \lim_{\Psi(N,M)\to\infty} \frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \left(\sum_{i=1}^{k} \rho_i^0 \cos(\omega_i^0 n + \upsilon_i^0 m + \varphi_i^0) \right)^2 \\
&- \lim_{\Psi(N,M)\to\infty} \left[2 \frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \sum_{i=1}^{k} \sum_{j=1}^{k} \rho_i \rho_j^0 \cos(\omega_i n + \upsilon_i m + \varphi_i) \cos(\omega_j^0 n + \upsilon_j^0 m + \varphi_j^0) \right] \\
&+ \lim_{\Psi(N,M)\to\infty} \frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \left(\sum_{i=1}^{k} \rho_i \cos(\omega_i n + \upsilon_i m + \varphi_i) \right)^2 = T_1 + T_2 + T_3.
\end{aligned}
\tag{66}
$$

Recall that for $|\rho| < \infty$ and $\varphi \in [0, 2\pi)$

$$
\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \rho \cos(\omega n + \varphi) = 0,
\tag{67}
$$

uniformly in ω on any closed interval in $(0, 2\pi)$. The same equality is hold for the sine function. Hence, due to Assumption 3 and (67), we have

$$
T_1 = \lim_{\Psi(N,M)\to\infty} \frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \left(\sum_{i=1}^{k} \rho_i^0 \cos(\omega_i^0 n + \upsilon_i^0 m + \varphi_i^0) \right)^2 = \sum_{i=1}^{k} \frac{(\rho_i^0)^2}{2},
\tag{68}
$$

independently of θ_k.

Also,

$$
\begin{aligned}
T_3 &= \lim_{\Psi(N,M)\to\infty} \frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \left(\sum_{i=1}^{k} \rho_i \cos(\omega_i n + \upsilon_i m + \varphi_i) \right)^2 = \sum_{i=1}^{k} \frac{(\rho_i)^2}{2} \\
&+ \lim_{\Psi(N,M)\to\infty} \frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \sum_{\substack{i=1 \\ i\neq j}}^{k} \sum_{j=1}^{k} \rho_i \rho_j \cos(\omega_i n + \upsilon_i m + \varphi_i) \cos(\omega_j n + \upsilon_j m + \varphi_j).
\end{aligned}
\tag{69}
$$

Since the pairs (ω_i, υ_i) are pairwise different, then on any closed interval in $(0, 2\pi)$ the sequence of partial sums $\frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \sum_{\substack{i=1 \\ i\neq j}}^{k} \sum_{j=1}^{k} \rho_i \rho_j \cos(\omega_i n + \upsilon_i m + \varphi_i) \cos(\omega_j n + \upsilon_j m + \varphi_j)$ converges uniformly to zero as $\Psi(N, M) \to \infty$.

Hence,

$$
T_3 = \sum_{i=1}^{k} \frac{(\rho_i)^2}{2},
\tag{70}
$$

as $\Psi(N, M) \to \infty$ uniformly on Δ_δ.

Leaving T_2 unchanged we obtain

$$
\begin{aligned}
\lim_{\Psi(N,M)\to\infty} I_1 &= \sum_{i=1}^{k} \left(\frac{(\rho_i^0)^2}{2} + \frac{(\rho_i)^2}{2} \right) \\
&- \lim_{\Psi(N,M)\to\infty} \frac{2}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \sum_{i=1}^{k} \sum_{j=1}^{k} \rho_i \rho_j^0 \cos(\omega_i n + \upsilon_i m + \varphi_i) \cos(\omega_j^0 n + \upsilon_j^0 m + \varphi_j^0),
\end{aligned}
\tag{71}
$$

16

uniformly on Δ_δ.

Using the similar considerations to those employed in the evaluation of (68) we obtain

$$
\lim_{\Psi(N,M)\to\infty} I_2 = \lim_{\Psi(N,M)\to\infty} \left[\frac{2}{NM} \sum_{n=0}^{N-1}\sum_{m=0}^{M-1} \left(\sum_{i=k+1}^{P} \rho_i^0 \cos(\omega_i^0 n + \upsilon_i^0 m + \varphi_i^0) \right) \right.
$$
$$
\left(\sum_{i=1}^{k} \rho_i^0 \cos(\omega_i^0 n + \upsilon_i^0 m + \varphi_i^0) - \sum_{i=1}^{k} \rho_i \cos(\omega_i n + \upsilon_i m + \varphi_i) \right) \right] \tag{72}
$$
$$
= - \lim_{\Psi(N,M)\to\infty} \left[\frac{2}{NM} \sum_{n=0}^{N-1}\sum_{m=0}^{M-1}\sum_{i=1}^{k}\sum_{j=k+1}^{P} \rho_i \rho_j^0 \cos(\omega_i n + \upsilon_i m + \varphi_i) \cos(\omega_j^0 n + \upsilon_j^0 m + \varphi_j^0) \right].
$$

By Lemma 3 in Appendix C, we have that a.s. as $\Psi(N,M) \to \infty$:

$$
\sup_{\theta_k \in \Delta_\delta} \left| \frac{2}{NM} \sum_{n=0}^{N-1}\sum_{m=0}^{M-1} w(n,m) \left(\sum_{i=1}^{k} \rho_i^0 \cos(\omega_i^0 n + \upsilon_i^0 m + \varphi_i^0) - \sum_{i=1}^{k} \rho_i \cos(\omega_i n + \upsilon_i m + \varphi_i) \right) \right| \to 0. \tag{73}
$$

Hence $I_3 \to 0$ *a.s.* as $\Psi(N,M) \to \infty$ uniformly on Δ_δ. Using (71), (72) and (73) we conclude that a.s.

$$
\lim_{\Psi(N,M)\to\infty} (\mathcal{L}_k(\theta_k) - \mathcal{L}_k(\theta_k^0)) = \sum_{i=1}^{k} \left(\frac{(\rho_i^0)^2}{2} + \frac{(\rho_i)^2}{2} \right)
$$
$$
- \lim_{\Psi(N,M)\to\infty} \frac{2}{NM} \sum_{n=0}^{N-1}\sum_{m=0}^{M-1}\sum_{i=1}^{k}\sum_{j=1}^{P} \rho_i \rho_j^0 \cos(\omega_i n + \upsilon_i m + \varphi_i) \cos(\omega_j^0 n + \upsilon_j^0 m + \varphi_j^0). \tag{74}
$$

To complete the evaluation of (74) we consider the vectors $\theta_k \in \Delta_\delta$. Let us first assume that $\Delta_\delta \equiv \mathcal{R}_{q\delta}$ for some q, $1 \le q \le k$. Thus, the coordinate ρ_q of each vector in this subset is different from the corresponding coordinate ρ_q^0 by at least $\delta > 0$. Consider first the case where all the other elements of the vector $\theta_k \in \mathcal{R}_{q\delta}$ are identical to the corresponding elements of θ_k^0. Since by this assumption $\omega_j = \omega_j^0$, $\upsilon_j = \upsilon_j^0$, $\varphi_j = \varphi_j^0$ for $1 \le j \le k$, and $\rho_j = \rho_j^0$ for $1 \le j \le k$, $j \ne q$, on this set we have

$$
\lim_{\Psi(N,M)\to\infty} (\mathcal{L}_k(\theta_k) - \mathcal{L}_k(\theta_k^0)) = \left(\frac{\rho_q^0}{\sqrt{2}} - \frac{\rho_q}{\sqrt{2}} \right)^2
$$
$$
- \lim_{\Psi(N,M)\to\infty} \frac{2}{NM} \sum_{n=0}^{N-1}\sum_{m=0}^{M-1}\sum_{\substack{i=1 \\ i \ne j}}^{k}\sum_{j=1}^{P} \rho_i \rho_j^0 \cos(\omega_i n + \upsilon_i m + \varphi_i) \cos(\omega_j^0 n + \upsilon_j^0 m + \varphi_j^0)
$$
$$
= \left(\frac{\rho_q^0}{\sqrt{2}} - \frac{\rho_q}{\sqrt{2}} \right)^2 \ge \frac{\delta^2}{2} > 0, \tag{75}
$$

uniformly in ρ_q, where the second equality is due to Assumption 3 and following the arguments employed to obtain (70).

Assume next that $\theta_k \in \mathcal{R}_{q\delta}$ (*i.e.*, the coordinate ρ_q is different from the corresponding coordinate ρ_q^0 by at least $\delta > 0$) and that in addition, there exists an element ρ_t of θ_k, such that

$1 \leq t \leq k$, $t \neq q$ and $|\rho_t - \rho_t^0| \geq \lambda, \lambda > 0$ while all the other elements of the vector θ_k are identical to the corresponding elements of θ_k^0. Following a similar derivation to the one in (75) we conclude that

$$\lim_{\Psi(N,M)\to\infty} \left(\mathcal{L}_k(\theta_k) - \mathcal{L}_k(\theta_k^0)\right) = \left(\frac{\rho_q^0}{\sqrt{2}} - \frac{\rho_q}{\sqrt{2}}\right)^2 + \left(\frac{\rho_t^0}{\sqrt{2}} - \frac{\rho_t}{\sqrt{2}}\right)^2 \geq \frac{\delta^2}{2} + \frac{\lambda^2}{2} > \frac{\delta^2}{2}, \quad (76)$$

uniformly in ρ_q and ρ_t.

Consider the case where $\theta_k \in \mathcal{R}_{q\delta}$ while there exists an element φ_l of $\theta_k \in \mathcal{R}_{q\delta}$, such that $|\varphi_l - \varphi_l^0| \geq \eta, \eta > 0$ and all the other elements of the vector θ_k are identical to the corresponding elements of θ_k^0. Following a similar derivation to the one in (75) we conclude that

$$\lim_{\Psi(N,M)\to\infty} \left(\mathcal{L}_k(\theta_k) - \mathcal{L}_k(\theta_k^0)\right) = \begin{cases} \left(\frac{\rho_q^0}{\sqrt{2}} - \frac{\rho_q}{\sqrt{2}}\right)^2 + (\rho_l^0)^2 - (\rho_l^0)^2 \cos(\varphi_l - \varphi_l^0), & l \neq q \\ \frac{(\rho_q^0)^2}{2} + \frac{(\rho_q)^2}{2} - \rho_q^0 \rho_q \cos(\varphi_q - \varphi_q^0), & l = q \end{cases} > \frac{\delta^2}{2}, \quad (77)$$

uniformly in ρ_q and φ_l.

Finally, consider the case where $\theta_k \in \mathcal{R}_{q\delta}$ while there exists an element ω_l of $\theta_k \in \mathcal{R}_{q\delta}$, such that $|\omega_l - \omega_l^0| \geq \eta, \eta > 0$ and all the other elements of the vector θ_k are identical to the corresponding elements of θ_k^0. Following a similar derivation to the one in (75) we conclude that

$$\liminf_{\Psi(N,M)\to\infty} \left(\mathcal{L}_k(\theta_k) - \mathcal{L}_k(\theta_k^0)\right) = \begin{cases} \left(\frac{\rho_q^0}{\sqrt{2}} - \frac{\rho_q}{\sqrt{2}}\right)^2 + (\rho_l^0)^2, & l \neq q \\ \frac{(\rho_q^0)^2}{2} + \frac{(\rho_q)^2}{2}, & l = q \end{cases} > \frac{\delta^2}{2}, \quad (78)$$

uniformly in ρ_q and ω_l.

From the above analysis it is clear that $\lim_{\Psi(N,M)\to\infty} \left(\mathcal{L}_k(\theta_k) - \mathcal{L}_k(\theta_k^0)\right)$ is lower bounded by $\frac{\delta^2}{2}$ uniformly in $\mathcal{R}_{q\delta}$.

Following similar reasoning, the next subset we consider is $W_{q\delta} \cup V_{q\delta}$. We first consider a subset of this set:

$$\Lambda = \left\{\theta_k \in W_{q\delta} \cup V_{q\delta} : \exists p, \, k+1 \leq p \leq P, \, (\omega_q, \upsilon_q) = (\omega_p^0, \upsilon_p^0) \right\} \subset W_{q\delta} \cup V_{q\delta} \quad (79)$$

This subset includes vectors in Θ_k, such that their coordinate pairs (ω_q, υ_q) are different from the corresponding pairs of θ_k^0 and equal to some pair $(\omega_p^0, \upsilon_p^0)$ where $p \geq k+1$. As above, the minimum is obtained when all the other elements of θ_k are identical to the corresponding elements of θ_k^0. Hence, uniformly on Λ, we have

$$\lim_{\Psi(N,M)\to\infty} \left(\mathcal{L}_k(\theta_k) - \mathcal{L}_k(\theta_k^0)\right) \geq \frac{(\rho_q^0)^2}{2} + \frac{(\rho_q)^2}{2} - \rho_p^0 \rho_q$$

$$= \frac{(\rho_q^0)^2}{2} - \frac{(\rho_p^0)^2}{2} + \left(\frac{\rho_p^0}{\sqrt{2}} - \frac{\rho_q}{\sqrt{2}}\right)^2 \geq \frac{(\rho_q^0)^2}{2} - \frac{(\rho_p^0)^2}{2} = \epsilon_\Lambda > 0, \quad (80)$$

18

where the last inequality is due to Assumption 4.

On the complementary set:

$$\Lambda^c = (W_{q\delta} \cup V_{q\delta}) \setminus \Lambda = \left\{ \theta_k \in W_{q\delta} \cup V_{q\delta} : (\omega_q, \upsilon_q) \neq (\omega_p^0, \upsilon_p^0), \forall p, \ k+1 \leq p \leq P \right\} \quad (81)$$

we have

$$\lim_{\Psi(N,M) \to \infty} (\mathcal{L}_k(\theta_k) - \mathcal{L}_k(\theta_k^0)) \geq \frac{(\rho_q^0)^2}{2} + \frac{(\rho_q)^2}{2} \geq \frac{(\rho_q^0)^2}{2} = \epsilon_{\Lambda^c} > 0. \quad (82)$$

Finally, on the set $\Phi_{q\delta}$ the coordinate φ_q of the each vector in this subset is different from the corresponding coordinate φ_q^0 by at least $\delta > 0$. As in previous cases , the minimum is obtained when all the other elements of $\theta_k \in \Phi_{q\delta}$ are identical to the corresponding elements of θ_k^0. Hence, uniformly on $\Phi_{q\delta}$, we have

$$\lim_{\Psi(N,M) \to \infty} (\mathcal{L}_k(\theta_k) - \mathcal{L}_k(\theta_k^0)) \geq (\rho_q^0)^2 - (\rho_q^0)^2 \cos(\varphi_q - \varphi_q^0) \geq (\rho_q^0)^2(1 - \cos\delta) = \epsilon_{\Phi_{q\delta}} > 0. \quad (83)$$

Let $\epsilon_q = \min(\frac{\delta^2}{2}, \epsilon_\Lambda, \epsilon_{\Lambda^c}, \epsilon_{\Phi_{q\delta}})$. Collecting (75),(80), (82) and (83) together we conclude that the sequence $\mathcal{L}_k(\theta_k) - \mathcal{L}_k(\theta_k^0)$ is lower bounded by $\epsilon_q > 0$ uniformly on $\mathcal{R}_{q\delta} \cup \Phi_{q\delta} \cup W_{q\delta} \cup V_{q\delta}$ as $\Psi(N, M) \to \infty$.

By repeating the same arguments for every q, $1 \leq q \leq k$, and by letting $\epsilon = \min(\epsilon_1, \ldots, \epsilon_k)$, we conclude that the sequence $\mathcal{L}_k(\theta_k) - \mathcal{L}_k(\theta_k^0)$ (indexed in N, M) is lower bounded by $\epsilon > 0$ uniformly on Δ_δ as $\Psi(N, M) \to \infty$.

Hence, it follows that sequence $\inf_{\theta_k \in \Delta_\delta} (\mathcal{L}_k(\theta_k) - \mathcal{L}_k(\theta_k^0))$ (indexed in N, M) is also asymptotically lower bounded by $\epsilon > 0$, *i.e.*,

$$\inf_{\theta_k \in \Delta_\delta} \left(\mathcal{L}_k(\theta_k) - \mathcal{L}_k(\theta_k^0) \right) \geq \epsilon, \quad (84)$$

as $\Psi(N, M) \to \infty$.

Hence, by the definition of $\liminf$

$$\liminf_{\Psi(N,M) \to \infty} \inf_{\theta_k \in \Delta_\delta} \left(\mathcal{L}_k(\theta_k) - \mathcal{L}_k(\theta_k^0) \right) \geq \epsilon > 0. \quad (85)$$

■

Appendix B

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

$$\Pr\{x_n \leq 0 \ i.o.\} \leq \Pr\{\liminf_{n \to \infty} x_n \leq 0\} \quad (86)$$

19

Proof: Let (Ω, Σ, p) be some probability space. Let $\{x_n(\omega), n \geq 1\}$ be a sequence of random variables. Let $\{A_n \in \Sigma, n \geq 1\}$ be a sequence of subsets of Ω, such that $A_n = \{\omega \in \Omega : x_n(\omega) \leq 0\}$. Define

$$A_n^m = \bigcup_{n=m}^{\infty} \{\omega : x_n \leq 0\} . \tag{87}$$

Then

$$A_n^m \subseteq \{\omega : \inf_{n \geq m} x_n \leq 0\} . \tag{88}$$

Hence

$$\bigcap_{m1}^{\infty} A_n^m \subseteq \bigcap_{m1}^{\infty} \{\omega : \inf_{n \geq m} x_n \leq 0\} . \tag{89}$$

Consider the R.H.S. of (89), and let $y_m(\omega) = \inf_{n \geq m} x_n$. Since for all $\omega \in \bigcap_{m1}^{\infty} \{\omega : \inf_{n \geq m} x_n \leq 0\}$, $y_m(\omega) \leq 0$ for all m, then by definition $\sup_m y_m(\omega) \leq 0$ as well. On the other hand if $\sup_m y_m(\omega) \leq 0$, then for all m, $y_m(\omega) \leq 0$. Hence we have the following set equality

$$\bigcap_{m1}^{\infty} \{\omega : \inf_{n \geq m} x_n \leq 0\} = \{\omega : \sup_m \inf_{n \geq m} x_n \leq 0\}. \tag{90}$$

Rewriting (89) we have

$$\bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} A_n \subseteq \{\omega : \sup_m \inf_{n \geq m} x_n \leq 0\} = \{\omega : \liminf_{n \to \infty} x_n(\omega) \leq 0\}, \tag{91}$$

where the equality on the R.H.S. of (91) follows from the definition of $\liminf_{n \to \infty}(\cdot)$ of a sequence x_n.

Also by definition, $\bigcap_{m1}^{\infty} \bigcup_{n=m}^{\infty} A_n = \limsup_{n \to \infty} A_n$. Hence, (see, *e.g.*, [1], p. 67)

$$\limsup_{n \to \infty} A_n = \{\omega : x_n(\omega) \leq 0 \ i.o.\} \subseteq \{\omega : \liminf_{n \to \infty} x_n(\omega) \leq 0\}. \tag{92}$$

Due to the monotonicity of the probability measure, the lemma follows. ∎

Appendix C

Let D be an *infinite* order non-symmetrical half-plane support defined as in (3) and let $D(k,l)$ be a *finite* order non-symmetrical half-plane support, defined by

$$D(k,l) = \left\{(i,j) \in \mathbb{Z}^2 : i = 0, 0 \leq j \leq l\right\} \cup \left\{(i,j) \in \mathbb{Z}^2 : 0 < i \leq k, -l \leq j \leq l\right\} \tag{93}$$

Let the field $\{w(n,m)\}$ be defined as in (4), and the field $\{u(n,m)\}$ is an i.i.d. real valued zero-mean random field with finite second order moment, σ^2. The sequence $a(i,j)$ is a square summable deterministic sequence,

$$\sum_{(r,s)\in D} a^2(r,s) < \infty. \tag{94}$$

The next lemma is an extension of a lemma originally proposed by Hannan, [6] for the case of 1-D signals. Similar result can be found in [11], Lemma 2, but with only a partial proof. Since this lemma is crucial for our work we will prove it here.

Lemma 3.

$$\sup_{\omega,\upsilon} \left| \frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} w(n,m) \cos(\omega n + \nu m) \right| \to 0 \ a.s. \ as \ \Psi(N,M) \to \infty \tag{95}$$

Proof:

First, it is easy to see that,

$$\sup_{\omega,\upsilon} \left| \frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} w(n,m) \cos(\omega n + \nu m) \right| \leq$$

$$\sup_{\omega,\upsilon} \left| \frac{1}{2NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} w(n,m) e^{j(\omega n + \nu m)} \right| + \sup_{\omega,\upsilon} \left| \frac{1}{2NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} w(n,m) e^{-j(\omega n + \nu m)} \right|. \tag{96}$$

Hence it is sufficient to prove the lemma for exponentials, *i.e.*, we wish to prove that

$$\sup_{\omega,\upsilon} \left| \frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} w(n,m) e^{j(\omega n + \nu m)} \right| \to 0 \ a.s. \ as \ \Psi(N,M) \to \infty \tag{97}$$

Define the set $D(k,l)^C = D \setminus D(k,l)$. Then,

$$w(n,m) = \sum_{D(k,l)} a(r,s) u(n-r,m-s) + \sum_{D(k,l)^C} a(r,s) u(n-r,m-s) = v(n,m) + z(n,m). \tag{98}$$

Then,

$$\sup_{\omega,\upsilon} \left| \frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} z(n,m) e^{j(\omega n + \nu m)} \right| \leq \frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} |z(n,m)| \leq \left\{ \frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} z^2(n,m) \right\}^{\frac{1}{2}}. \tag{99}$$

By the SLLN, the R.H.S. of the last inequality convergence, almost surely, to

$$E[z(0,0)^2]^{\frac{1}{2}} = \left(\sigma^2 \sum_{D(k,l)^C} a(r,s)^2 \right)^{\frac{1}{2}}, \tag{100}$$

21

which due to (94) may be made arbitrary small by taking k and l sufficiently large.

Hence it is sufficient to prove the lemma with $w(n,m)$ replaced by $v(n,m)$.

$$\sup_{\omega,v}\left|\frac{1}{NM}\sum_{n=0}^{N-1}\sum_{m=0}^{M-1}v(n,m)e^{j(\omega n+vm)}\right| \leq \sum_{D(k,l)}|a(r,s)|\sup_{\omega,v}\left|\frac{1}{NM}\sum_{n=0}^{N-1}\sum_{m=0}^{M-1}u(n-r,m-s)e^{j(\omega n+vm)}\right| \quad (101)$$

Since the summation is finite and $\{u(n,m)\}$ is i.i.d., it is sufficient to prove the lemma with $w(n,m)$ replaced by $u(n,m)$. Thus, we consider the mean square of the discussed supremum

$$E\left[\sup_{\omega,v}\left|\frac{1}{NM}\sum_{n=0}^{N-1}\sum_{m=0}^{M-1}u(n,m)e^{j(\omega n+vm)}\right|^2\right]$$

$$= E\left[\sup_{\omega,v}\frac{1}{(NM)^2}\sum_{n=0}^{N-1}\sum_{m=0}^{M-1}\sum_{k=0}^{N-1}\sum_{l=0}^{M-1}u(n,m)u(k,l)e^{j(\omega(n-k)+v(m-l))}\right]. \quad (102)$$

By letting,

$$n - k = p,$$
$$m - l = r, \quad (103)$$

substitute,

$$\sum_{n=0}^{N-1}\sum_{k=0}^{N-1} = \sum_{|p|<N}\sum_{n\in S_N},$$
$$\sum_{m=0}^{M-1}\sum_{l=0}^{M-1} = \sum_{|r|<M}\sum_{m\in S_M}, \quad (104)$$

where,

$$S_N = \{n \in \mathbb{Z} : \max(0,p) \leq n \leq \min(N-1, p+N-1)\},$$
$$S_M = \{m \in \mathbb{Z} : \max(0,r) \leq m \leq \min(M-1, r+M-1)\}, \quad (105)$$

and,

$$\sum_{n\in S_N}1 = \begin{cases} N-p, & p \geq 0 \\ N+p, & p < 0 \end{cases} = N - |p|,$$
$$\sum_{m\in S_M}1 = \begin{cases} M-r, & r \geq 0 \\ M+r, & r < 0 \end{cases} = M - |r|. \quad (106)$$

Hence, rewriting (102) we have

$$\frac{1}{(NM)^2} E\left[\sup_{\omega,v} \sum_{|p|<N}\sum_{|r|<M}\sum_{n\in S_N}\sum_{m\in S_M} u(n,m)u(n-p,m-r)e^{j(\omega p+\nu r)}\right]$$

$$= \frac{1}{(NM)^2} E\left[\sum_{n=0}^{N-1}\sum_{m=0}^{M-1} u(n,m)^2 + \sup_{\omega,v} \sum_{\substack{|p|<N\\p\neq 0}}\sum_{\substack{|r|<M\\r\neq 0}}\sum_{n\in S_N}\sum_{m\in S_M} u(n,m)u(n-p,m-r)e^{j(\omega p+\nu r)}\right]$$

$$\leq \frac{1}{(NM)^2}\left\{NM\sigma^2 + \sum_{\substack{|p|<N\\p\neq 0}}\sum_{\substack{|r|<M\\r\neq 0}} E\left[\left|\sum_{n\in S_N}\sum_{m\in S_M} u(n,m)u(n-p,m-r)\right|\right]\right\}, \qquad (107)$$

where in the first equality we split up the sum into the squared term and the remainder, and then employ the triangular inequality.

Let us investigate the second term on the R.H.S. of (107). From the Cauchy-Schwartz inequality, for any r.v. x, $E[|x|] \leq E[|x|^2]^{\frac{1}{2}}$, hence

$$E\left[\left|\sum_{n\in S_N}\sum_{m\in S_M} u(n,m)u(n-p,m-r)\right|\right] \leq E\left[\left|\sum_{n\in S_N}\sum_{m\in S_M} u(n,m)u(n-p,m-r)\right|^2\right]^{\frac{1}{2}}$$

$$= \left(\sum_{n\in S_N}\sum_{m\in S_M}\sum_{n'\in S_N}\sum_{m'\in S_M} E[u(n,m)u(n-p,m-r)u(n',m')u(n'-p,m'-r)]\right)^{\frac{1}{2}}$$

$$= \left(\sum_{n\in S_N}\sum_{m\in S_M}\sigma^4\right)^{\frac{1}{2}} = \sigma^2(N-|p|)^{\frac{1}{2}}(M-|r|)^{\frac{1}{2}}. \qquad (108)$$

which follows from the observation that for $p, r \neq 0$, the fourth order moment of the field $\{u(n,m)\}$ equals zero for all $n \neq n'$ or $m \neq m'$.

Hence we can finally write

$$E\left[\sup_{\omega,\upsilon}\left|\frac{1}{NM}\sum_{n=0}^{N-1}\sum_{m=0}^{M-1}u(n,m)e^{j(\omega n+\nu m)}\right|^2\right]$$

$$\leq \frac{1}{(NM)^2}\left\{NM\sigma^2 + \sum_{\substack{|p|<N\\p\neq0}}\sum_{\substack{|r|<M\\r\neq0}}\sigma^2(N-|p|)^{\frac{1}{2}}(M-|r|)^{\frac{1}{2}}\right\}$$

$$\leq \frac{\sigma^2}{(NM)^2}\{NM + 4(NM)^{\frac{3}{2}}\} \leq \frac{K}{(NM)^{\frac{1}{2}}} = O(N^{-\frac{1}{2}}M^{-\frac{1}{2}}). \tag{109}$$

where K some finite positive constant.

Now following the ideas of Doob, [2](ch. X, 6), let R and S be some positive integers such that $N > R^\delta$, and $M > S^\delta$, for $\delta > 2$. Hence, for any such choice of N and M, from (109),

$$E\left[\sup_{\omega,\upsilon}\left|\frac{1}{NM}\sum_{n=0}^{N-1}\sum_{m=0}^{M-1}u(n,m)e^{j(\omega n+\nu m)}\right|^2\right] \leq \frac{K}{(RS)^{\frac{\delta}{2}}}. \tag{110}$$

Hence, if we take $N = N(R)$ and $M = M(S)$ to be the smallest integers not smaller then R^δ and S^δ, respectively, then (110) still holds.

Hence, by Chebyshev inequality for every $\epsilon > 0$

$$P\left(\sup_{\omega,\upsilon}\left|\frac{1}{N(R)M(S)}\sum_{n=0}^{N(R)-1}\sum_{m=0}^{M(S)-1}u(n,m)e^{j(\omega n+\nu m)}\right| \geq \epsilon\right)$$

$$\leq \frac{E\left[\sup_{\omega,\upsilon}\left|\frac{1}{N(R)M(S)}\sum_{n=0}^{N(R)-1}\sum_{m=0}^{M(S)-1}u(n,m)e^{j(\omega n+\nu m)}\right|^2\right]}{\epsilon^2} \leq \frac{K}{\epsilon^2(RS)^{\frac{\delta}{2}}} \tag{111}$$

and then since $\delta > 2$

$$\sum_{R=1}^{\infty}\sum_{S=1}^{\infty}P\left(\sup_{\omega,\upsilon}\left|\frac{1}{N(R)M(S)}\sum_{n=0}^{N(R)-1}\sum_{m=0}^{M(S)-1}u(n,m)e^{j(\omega n+\nu m)}\right| > \epsilon\right) \leq \sum_{R=1}^{\infty}\sum_{S=1}^{\infty}\frac{K}{\epsilon^2(RS)^{\frac{\delta}{2}}} < \infty. \tag{112}$$

Hence, by the Borel-Cantelly lemma,

$$\sup_{\omega,\upsilon}\left|\frac{1}{N(R)M(S)}\sum_{n=0}^{N(R)-1}\sum_{m=0}^{M(S)-1}u(n,m)e^{j(\omega n+\nu m)}\right| \to 0 \text{ a.s. as } \Psi(R,S) \to \infty. \tag{113}$$

Now,

$$\sup_{\substack{N(R)\leq N\leq N(R+1)\\M(S)\leq M\leq M(S+1)}} \sup_{\omega,\upsilon} \left| \frac{1}{NM}\sum_{n=0}^{N-1}\sum_{m=0}^{M-1} u(n,m)e^{j(\omega n+\nu m)} - \frac{1}{NM}\sum_{n=0}^{N(R)-1}\sum_{m=0}^{M(S)-1} u(n,m)e^{j(\omega n+\nu m)} \right|$$

$$\leq \sup_{\substack{N(R)\leq N\leq N(R+1)\\M(S)\leq M\leq M(S+1)}} \sup_{\omega,\upsilon} \frac{1}{NM} \left| \sum_{n=0}^{N(R)-1}\sum_{m=M(S)}^{M-1} u(n,m)e^{j(\omega n+\nu m)} \right|$$

$$+ \sup_{\substack{N(R)\leq N\leq N(R+1)\\M(S)\leq M\leq M(S+1)}} \sup_{\omega,\upsilon} \frac{1}{NM} \left| \sum_{n=N(R)}^{N-1}\sum_{m=0}^{M(S)-1} u(n,m)e^{j(\omega n+\nu m)} \right|$$

$$+ \sup_{\substack{N(R)\leq N\leq N(R+1)\\M(S)\leq M\leq M(S+1)}} \sup_{\omega,\upsilon} \frac{1}{NM} \left| \sum_{n=N(R)}^{N-1}\sum_{m=M(S)}^{M-1} u(n,m)e^{j(\omega n+\nu m)} \right| = I_1 + I_2 + I_3. \tag{114}$$

Consider the first term in the previous equation. Using the triangular inequality

$$I_1 \leq \frac{1}{M(S)} \sum_{m=M(S)}^{M(S+1)-1} \left(\sup_\omega \frac{1}{N(R)} \left| \sum_{n=0}^{N(R)-1} u(n,m)e^{j\omega n} \right| \right). \tag{115}$$

Let

$$\tilde{u}(m) = \sup_\omega \frac{1}{N(R)} \left| \sum_{n=0}^{N(R)-1} u(n,m)e^{j\omega n} \right|. \tag{116}$$

Since $\{u(n,m)\}$ is i.i.d., it is clear that $\{\tilde{u}(m)\}$ is an i.i.d. sequence of random variables. Moreover, from [6] (or by repeating the derivation in (98)-(110) for the process $u(n,m)$ with a fixed m) we have

$$E\left[\tilde{u}(m)^2\right] = E\left[\sup_\omega \frac{1}{N(R)} \left| \sum_{n=0}^{N(R)-1} u(n,m)e^{j\omega n} \right|^2 \right] \leq \frac{K_1}{R^{\frac{\delta}{2}}}. \tag{117}$$

Taking the mean of the square of the I_1 we have

$$E\left[|I_1|^2\right] \leq \frac{1}{M(S)^2} \sum_{m=M(S)}^{M(S+1)-1}\sum_{m'=M(S)}^{M(S+1)-1} E\left[\tilde{u}(m)\tilde{u}(m')\right]$$

$$\leq \frac{1}{M(S)^2} \sum_{m=M(S)}^{M(S+1)-1}\sum_{m'=M(S)}^{M(S+1)-1} E\left[\tilde{u}(m)^2\right]^{\frac{1}{2}} E\left[\tilde{u}(m')^2\right]^{\frac{1}{2}}$$

$$\leq \frac{K_1(M(S+1)-1-M(S))^2}{R^{\frac{\delta}{2}}M(S)^2} \leq \frac{K}{R^{\frac{\delta}{2}}S^2}. \tag{118}$$

Using once again the Chebyshev inequality and the Borel-Cantelli lemma we have that $I_1 \to 0$ a.s. as $\Psi(R, S) \to \infty$. Repeating the same consideration for I_2 we have that $I_2 \to 0$ a.s. as $\Psi(R, S) \to \infty$. Finally, for I_3 we have

$$
\begin{aligned}
E[|I_3|^2] &\leq E\left[\left|\frac{1}{N(R)M(S)} \sum_{n=N(R)}^{N(R+1)-1} \sum_{m=M(S)}^{M(R+1)-1} |u(n,m)|\right|^2\right] \\
&= \frac{1}{(N(R)M(S))^2} \sum_{n=N(R)}^{N(R+1)-1} \sum_{m=M(S)}^{M(S+1)-1} \sum_{n'=N(R)}^{N(R+1)-1} \sum_{m'=M(S)}^{M(S+1)-1} E[|u(n,m)u(n',m')|] \\
&\leq \sigma^2 \frac{(N(R+1)-1-N(R))^2(M(S+1)-1-M(S))^2}{(N(R)M(S))^2} \leq \frac{K}{(RS)^2}.
\end{aligned} \tag{119}
$$

Using again the Chebyshev inequality and the Borel-Cantelli lemma we have that $I_3 \to 0$ a.s. as $\Psi(R, S) \to \infty$.

Finally, we have that

$$
\sup_{\omega,\nu} \left| \frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} u(n,m)e^{j(\omega n+\nu m)} - \frac{1}{NM} \sum_{n=0}^{N(R)-1} \sum_{m=0}^{M(S)-1} u(n,m)e^{j(\omega n+\nu m)} \right| \to 0 \text{ a.s.} \tag{120}
$$

for all $N(R) \leq N < N(R+1)$ and $M(S) \leq M < M(S+1)$, as $\Psi(R, S) \to \infty$, and hence as $\Psi(N, M) \to \infty$,

Since $\frac{N(R)}{N(R+1)} \to 1$ and $\frac{M(S)}{M(S+1)} \to 1$ as $\Psi(R, S) \to \infty$ we can replace $\frac{1}{NM}$ in the second term by $\frac{1}{N(R)M(S)}$. Therefore, we have

$$
\sup_{\omega,\nu} \left| \frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} u(n,m)e^{j(\omega n+\nu m)} - \frac{1}{N(R)M(S)} \sum_{n=0}^{N(R)-1} \sum_{m=0}^{M(S)-1} u(n,m)e^{j(\omega n+\nu m)} \right| \to 0 \text{ a.s.} \tag{121}
$$

From (121) and (113) the lemma follows.

∎

References

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

[2] J. L. Doob, *Stochastic Processes*, John Wiley, 1953.

[3] 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 Process.*, 44 (1996), 1242-1255.

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

[5] J. M. Francos, "Bounds on the Accuracy of Estimating the Parameters of Discrete Homogeneous Random Fields with Mixed Spectral Distributions," *IEEE Trans. Info. Theory*, 43 (1997), 908-922.

[6] E. J. Hannan, "The Estimation of Frequency," *J. Appl. Prob.*,10 (1973),510-519.

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

[8] S. He, "Parameter Estimation of Hidden Periodic Model in Random Fields," *Science in China (Ser. A)*, 42 (1998), 238-245.

[9] M. Kliger and J. M. Francos, "MAP Model Order Selection Rule for 2-D Sinusoids in White Noise," *IEEE Trans. Signal Process.*, 53 (2005), 2563-2575.

[10] M. Kliger and J. M. Francos, "Strong Consistency of a Family of Model Order Selection Rules for Estimating the Parameters of 2-D Sinusoids in White Noise," *Proc. Int. Conf. Acoust., Speech, Signal Processing*, Philadelphia, 2005.

[11] D. Kundu and S. Nandi, "Determination of Discrete Spectrum in a Random Field," *Statistica Neerlandica*, 57 (2003), 258-283.

[12] J. Li and P. Stoica, "Efficient Mixed-Spectrum Estimation with Applications to Target Feature Extraction, *IEEE Transactions on Signal Processing*," 44 (1996), 281-295.

[13] H. Zhang and V. Mandrekar, "Estimation of Hidden Frequencies for 2D Stationary Processes," *J. Time Ser. Anal.*, 22 (2001), 613-629.

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

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