ESTIMATION OF HIDDEN FREQUENCIES FOR 2D STATIONARY PROCESSES

By Hao Zhang and V. Mandrekar

Washington State University and Michigan State University

First Version received January 1995

Abstract. We study a stationary random field model that is composed of a signal of an unknown number of sine and cosine functions, and a coloured noise. This model has been used in image analysis and modelling spatial data, and is useful for signal extraction in the presence of coloured noise. The problem is to estimate the number of unknown frequencies and the unknown frequencies. The analogous time series model and related problems have been extensively studied. Our approach is based on some analytic properties of periodograms of stationary random fields that we establish in the paper. In particular, we show that the periodogram of a stationary random field of a moving average has a uniform upper bound of $O(\ln(N^2))$ where N^2 is the sample size, and that the periodogram of the observed process has a magnitude of the order N^2 uniformly in a neighbourhood of any hidden frequency, and much smaller outside.

Keywords. Mixed spectra; periodogram; random field; spatial data; uniform upper bound.

1. INTRODUCTION

We consider the following second-order stationary random field model

$$y(m, n) = \sum_{k=1}^{p} \{C_k \cos(m\lambda_k + n\mu_k) + D_k \sin(m\lambda_k + n\lambda_k)\} + x(m, n)$$

$$m, n = 0, \pm 1, \pm 2, \ldots \qquad (1)$$

where $\{C_k, D_k\}$ is a set of uncorrelated random variables and uncorrelated with $\{x(m, n)\}$, $\mathrm{Var}(C_k) = \mathrm{Var}(D_k)$, and $\{x(m, n), (m, n) \in \mathbb{Z}^2\}$ is a stationary random field with an absolutely continuous spectral distribution. The problem is to estimate the number of frequencies p, and the unknown frequencies (λ_k, μ_k), $k = 1, 2, \ldots, p$, given observations $y(m, n)$, $m, n = 1, 2, \ldots, N$. When p is known, estimation of the unknown frequencies can be given via the least squares method – see, for example, Kundu and Mitra (1996) and Bansal *et al.* (1999) – or the maximum likelihood method (Rao *et al.*, 1994). However, determination of p is a difficult problem.

The first term on the right-hand side of (1) is usually called a signal (also called a harmonic random field), and the detection of signal is an important problem in signal processing. The first term corresponds to the purely deterministic part in the Wold decomposition of $\{y(m, n)\}$ and has a spectral measure that concentrates on (λ_k, μ_k), $k = 1, \ldots, p$, while the second term corresponds to the regular part in the Wold decomposition (Helson and Lowdenslager, 1958, 1962). Therefore, the spectrum of $y(m, n)$ has both discrete and continuous components. This observation leads to the detection of signal and estimation of unknown frequencies through the analysis of periodogram, an approach first employed by Schuster (1898) who introduced the periodogram to detect the hidden periodicities in a time series. Priestley (1964) considered model (1) of mixed spectra and proposed a method for detecting the signal under a more general scheme where the purely deterministic part could have a spectral distribution concentrated on a set of zero Lebesgue measure. Francos *et al.* (1993) used (1) to model some texture images and estimated the unknown frequencies by choosing 'the largest and sharpest isolated peaks' of the periodogram of $y(m, n)$. Our results in this paper will help to consolidate the approach of Francos *et al.* and others on the analysis of periodogram; see, for example, Ripley (1981, ch. 5) for an analysis on agricultural uniformity trial data.

The analogous time series model has been extensively studied. Early works include Schuster (1898), Fisher (1929) and Hartley (1949) for testing the presence of hidden periodicities in which the noise is a Gaussian white noise. Later on, research has been focused more on the detection of signal in the presence of a coloured noise. In this case, the fundamental work was done in Whittle (1952, 1954) and Priestley (1962a, b). We refer to Priestley (1997) for a review of the history and different tests for the detection of signal, and Brillinger (1987) for different procedures for estimating the discrete frequencies of the signal. One of the open problems listed in Brillinger (1987) was the estimation of the number of frequencies in the signal, p. Since then much work has been done on the estimation of p. Some of the methods employ the uniform upper bound of the periodogram of a stationary time series established by Brillinger (1981, Thm 5.3.2) and An *et al.* (1983). For example, Quinn (1989) and Wang (1993) proposed AIC type procedures by adding a penalty term to the logarithm of the residual mean square of the least squares estimation and choosing the penalty term to make the estimators consistent. In a more direct way, Chen (1988) applied the uniform upper bound to study the behaviour of the periodogram in and out of a small neighbourhood of a hidden frequency, and proposed a method to obtain a consistent estimator of p.

Since analogous results on the periodogram of a stationary random field can be established, some methods on estimating the frequencies and the number of frequencies for a time series can be extended. In the present work, we propose a method for estimating p and the hidden frequencies that is based on the properties of the periodogram. As in the time series case, the periodogram is important in the spectral analysis of random fields; see, for example, Yuan and Subba Rao (1992, 1993). The periodogram of a stationary random field $y(m, n)$ with observations $y(m, n)$, $m, n = 1, 2, \ldots, N$ is defined as

$$I_N(\lambda, \mu; y) = \frac{1}{(2\pi N)^2} \left| \sum_{m=1}^{N} \sum_{n=1}^{N} y(m, n) e^{-i(m\lambda + n\mu)} \right|^2$$

We will show in Theorem 1 that the periodogram of the regular random field $x(m, n)$ has a uniform upper bound of $O(\ln(N^2))$ under some regularity conditions. Since the periodogram of the harmonic field has a magnitude of the order N^2 in a small neighbourhood of (λ_k, μ_k) for any k, and substantially smaller outside these neighbourhoods, $I_N(\lambda, \mu; y)$ is substantially larger in these neighbourhoods. The behaviour of $I_N(\lambda, \mu; y)$ is given in Theorems 2 and 3. Based on these, consistent estimators of p and (λ_k, μ_k), $k = 1, 2, \ldots, p$ are constructed in Theorem 4. Although this approach looks similar to that of Chen (1988), it is not a direct extension of Chen. Here we have to deal with the fact that when (λ, μ) is not close to a frequency (λ_k, μ_k), it is possible that λ is close to λ_k or μ is close to μ_k and consequently $I_N(\lambda, \mu; y)$ can be greater that an $O(\ln N)$ outside the small neighbourhood of (λ_k, μ_k).

Main results are presented in Section 2. An algorithm for estimation and simulation results are presented in Section 3. Proofs of the theorems are provided in the Appendix. Since each term in the harmonic field can be written as

$$C_k \cos(-m\lambda_k - n\mu_k) - D_k \sin(-m\lambda_k - n\mu_k)$$

we assume that $\mu_k \geqslant 0$ for $k = 1, 2, \ldots, p$ to make the model identifiable, and this assumption does not put any constraint on the model. Since the periodogram is symmetric, i.e., $I_N(\lambda, \mu; y) = I_N(-\lambda, -\mu; y)$, we restrict $I_N(\lambda, \mu)$ to the set $\Pi = (-\pi, \pi] \times [0, \pi]$.

2. PROPERTIES OF PERIODOGRAM AND ESTIMATION OF THE HARMONIC FIELD

As we noted, the random field $\{x(m, n)\}$ is the regular component in the Wold decomposition and therefore possesses a non-symmetric half-plane moving average representation,

$$x(m, n) = \sum_{j=-\infty}^{\infty} \sum_{k=0}^{\infty} c(j, k)\epsilon(m - j, n - k)$$

where $\{\epsilon(m, n)\}$ is the innovation process and the coefficients $c(j, k)$ are square summable. To study the upper bound of the periodogram, the absolute summability or stronger condition on $c(i, j)$ is usually needed; see Brillinger (1981) and An *et al.* (1983) for the time series case. He (1995) obtained the $O(\ln(N^2))$ uniform upper bound for the periodogram of a stationary random field that can be represented as a quarter-plane moving average:

$$x(m, n) = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} c(j, k)\epsilon(m - j, n - k)$$

where the coefficients satisfy $\sum_j \sum_k (j+k)|c(j, k)| < \infty$.

Since we need an upper bound for the regular process $\{x(m, n)\}$ represented as a non-symmetric half-plane moving average, we cannot use the results of He. We will first obtain upper bounds for a stationary field representable as a moving average (of any kind) of a white noise with absolutely summable coefficients. We make the following assumption throughout this section.

ASSUMPTION 1. *$\{x(m, n): m, n \in \mathbb{Z}\}$ is a stationary random field and can be represented as*

$$x(m, n) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} c(j, k)\epsilon(m-j, n-k) \tag{2}$$

where $c(j, k)$ are constants and absolutely summable, i.e.,

$$\sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} |c(j, k)| < \infty \tag{3}$$

and $\{\epsilon(m, n), m, n \in \mathbb{Z}\}$ is a double array of independent random variables with

$$\mathrm{E}\epsilon(m, n) = 0 \qquad \mathrm{E}\epsilon^2(m, n) = 1 \qquad \sup_{m,n} \mathrm{E}|\epsilon(m, n)|^r < \infty$$

for some $r > 2$.

Theorem 1 establishes upper bounds for the periodogram:

THEOREM 1. *Let $I_N(\lambda, \mu)$ be the periodogram of $x(m, n)$ satisfying Assumption 1, then*

$$\limsup_{N\to\infty} \frac{I_N(\lambda, \mu)}{\ln \ln N} \leqslant 2f(\lambda, \mu) \qquad \text{for any } (\lambda, \mu)$$

Furthermore, if $x(m, n)$ satisfies Assumption 1 with some $r > 3$,

$$\limsup_{N\to\infty} \frac{\sup_{\lambda,\mu} I_N(\lambda, \mu)}{\ln(N^2)} \leqslant 7\|f\|$$

where f is the spectral density function of $x(m, n)$ and $\|f\| = \sup f(\lambda, \mu)$.

Using Theorem 1, we can easily obtain Theorem 2 which says $I_N(\lambda, \mu; y)$ has a magnitude of the order N^2 at any frequency (λ_j, μ_j), $j = 1, 2, \ldots, p$, and $O(\ln \ln N)$ elsewhere:

THEOREM 2. *If $y(m, n)$ satisfies (1) with $x(m, n)$ satisfying Assumption 1, then for any fixed $(\lambda, \mu) \in \Pi$*

$$I_N(\lambda, \mu; y) = \begin{cases} (A_j N/(4\pi))^2 + o(N^2) & \textit{if } (\lambda, \mu) = (\lambda_j, \mu_j) \textit{ for some } j \\ O(\ln \ln N) & \textit{otherwise} \end{cases}$$

The next theorem states the uniform behaviour of $I_N(\lambda, \mu; y)$ in and out of neighbourhoods of (λ_j, μ_j). For $\alpha > 0$, define

$$\Delta_{j,\alpha} = \{(\lambda, \mu) \in \Pi: |\lambda - \lambda_j| \leqslant \pi/N^\alpha, |\mu - \mu_j| \leqslant \pi/N^\alpha\} \qquad j = 1, \ldots, p$$

$$\Delta_\alpha = \bigcup_{j=1}^{p} \Delta_{j,\alpha}$$

THEOREM 3. *Suppose that $y(m, n)$ satisfies (1) with $x(m, n)$ satisfying Assumption 1 with some $r > 3$. Then, with probability 1, for sufficiently large N*

$$\inf_{(\lambda,\mu)\in\Delta_{j,1}} I_N(\lambda, \mu; y) > \left(\frac{A_j N}{\pi^3}\right)^2 \qquad \textit{and} \qquad \sup_{(\lambda,\mu)\in\Delta_\alpha^c} I_N(\lambda, \mu; y) = O(N^{2\alpha})$$

where $\Delta_\alpha^c = \Pi \backslash \Delta_\alpha$.

Theorems 2 and 3 provide ways to estimate p and the hidden frequencies in the harmonic field. We consider one approach here which can be easily implemented on computers. Let α, β and c be constants such that

$$0 < \alpha < 1 \qquad 2\alpha < \beta < 2 \qquad c > 0$$

Let

$$\Omega = \{(\lambda, \mu) \in \Pi: I_N(\lambda, \mu; y) > cN^\beta\}.$$

Theorem 3 clearly indicates that when N is sufficiently large, Ω is contained in Δ_α and Ω contains some 'clusters' which occur around the frequencies (λ_j, μ_j). To state it more mathematically, we define a cluster as a subset S of Ω such that

1 the diameter of S is no greater than $2\sqrt{2}\pi/N^\alpha$, and
2 for any $(\lambda, \mu) \in \Omega \backslash S$, the diameter of $\{(\lambda, \mu)\} \cup S$ is greater than $2\sqrt{2}\pi/N^\alpha$.

The diameter of a set S is defined as

$$\gamma(S) = \sup\{\rho(x, y), x \in S, y \in S\}$$

where $\rho(x, y)$ is the Euclidean distance between x and y.

Let p_N be the number of clusters. The next theorem says p_N is a consistent estimator of p, the number of frequencies.

THEOREM 4. *Under conditions in Theorem 3, with probability one, there will be exactly $p_N = p$ clusters for large N each of which is of the form $\Delta_{j,\alpha} \cap \Omega$ and consequently for any (λ, μ) in a cluster, $\rho((\lambda, \mu), (\lambda_j, \mu_j)) < 2\sqrt{2}\pi/N^\alpha$ for some j.*

3. AN ALGORITHM AND SIMULATION RESULTS

We present an algorithm and some numerical results in this section. For an integer $d > 0$, let

$$\omega_d = \left\{ \left(\frac{j\pi}{dN}, \frac{k\pi}{dN} \right) : j = 0, \pm 1, \ldots, \pm dN, \ k = 0, 1, \ldots, N \right\}$$

$$\Omega_d = \Omega \cap \omega_d$$

Our algorithm consists of the following steps:

1 Calculate $I_N(\lambda, \mu; y)$ for $(\lambda, \mu) \in \omega_d$.
2 Identify the points in ω_d where $I_N(\lambda, \mu; y) > cN^\beta$.
3 Identify the clusters through the definition.
4 Also, find a value (λ, μ) that maximizes the periodogram in each of the clusters and take this frequency as an estimator of a hidden frequency.

The results in the previous section are asymptotic so that the choices of the constants α, β, c and d do not affect the results when the sample size is sufficiently large. For a fixed sample size, the choices might be empirical. For example, we should choose a small c if one of the amplitudes is deemed small, but a too small value of c might lead to overestimating p. α should be chosen closer to 1 to differentiate two frequencies which are close to each other. Simulation results can help us gain insights into the choices of α, β, c and d for finite sample sizes. We used

$$\alpha = 3/4 \qquad \beta = 1.75 \qquad c = 6/\pi^6 \qquad d = 1 \text{ or } 4$$

For the signal in model (1), we chose

$$p = 2 \qquad C_1 = D_1 = 1/\sqrt{2} \qquad C_2 = \sqrt{3} \qquad D_2 = 1$$
$$(\lambda_1, \mu_1) = (-0.1\pi, 0.3\pi) \qquad (\lambda_2, \mu_2) = (-0.5\pi, 0.7\pi),$$

and used four different noises:

1 i.i.d. standard normal
2 i.i.d. normal random field with mean 0 and standard deviation 2
3 A moving average

$$x(m, n) = 1/\sqrt{6} \sum_{j=-\infty}^{\infty} \sum_{k=0}^{\infty} 2^{-|j|-k} \epsilon(m - j, n - k)$$

where $\epsilon(m, n)$ are i.i.d. N(0, 1). We denote this noise by MA_1.

4 A moving average

$$x(m, n) = 2/\sqrt{6} \sum_{j=-\infty}^{\infty} \sum_{k=0}^{\infty} 2^{-|j|-k} \epsilon(m - j, n - k)$$

where $\epsilon(m, n)$ are i.i.d. N(0, 1). We denote this noise by MA_2.

Then $\mathrm{E}x(m, n)^2 = 1$ and 4 for the MA_1 noise and the MA_2 noise respectively and, therefore, we can compare the estimators from the coloured noises to those that are from the white noises. For each of the four noises, we simulated observations $y(m, n)$, for $m, n = 1, 2 \ldots, 19$ according to model (1) and the above settings, and used the algorithm outlined in this section to estimate p and the frequencies using $d = 1$ and $d = 4$. This procedure was repeated 100 times. We intentionally chose a small value $N = 19$ to evaluate the performance of the estimation for small samples. We expect larger sample sizes only yield better estimators with the same choice of α, β, c and d.

The points where $I_N(\lambda, \mu; y) > cN^{\beta}$ naturally occur in 'clustered sets'. However, we need to check each such 'clustered set' has a diameter no larger than $2\sqrt{2}\pi/N^{\alpha}$, and adding an additional point will increase the diameter to more than $2\sqrt{2}\pi/N^{\alpha}$.

When we use $d = 1$, p is correctly estimated 100% of the time for i.i.d. N(0, 1) and MA_1 noises, and overestimated once ($\hat{p} = 3$) for the i.i.d. N(0, 4) noise, and overestimated twice ($\hat{p} = 3$) and underestimated once ($\hat{p} = 1$) for the MA_2 noise. The average of the estimated frequencies when p is correctly estimated is shown in Table I. We see that a higher signal–noise ratio (hence a smaller variance of the noise) generally improves the estimation of p, but does not greatly affect the precision of the estimation of frequencies. Since the coloured noise does not have a spectral density as flat as that of a white noise, estimation of p may require larger sample sizes in the presence of coloured noise. However, as seen from Table I, coloured noises do not greatly affect the estimation of the frequencies if p is estimated correctly. Since the estimators for the frequencies may be biased, the mean squared errors are reported.

Similar conclusions can be made with $d = 4$. When d is increased to 4, estimation for p and the frequencies generally become better. The number p is correctly estimated 100% of the time for all noises except the MA_2 noise, for

TABLE I

AVERAGES OF ESTIMATES OF FREQUENCIES (IN MULTIPLES OF π) WITH $d = 1$

Noise	λ_1/π	μ_1/π	λ_2/π	μ_2/π
i.i.d. N(0, 1)	−0.1052632	0.3157895	−0.4978947	0.6842105
	(0.0000277)	(0.0002493)	(0.0006925)	(0.0002493)
i.i.d. N(0, 4)	−0.1052632	0.3135965	−0.5010965	0.6842105
	(0.0000277)	(0.0002955)	(0.0006925)	(0.0002493)
MA_1	−0.1052632	0.3157895	−0.4821053	0.6842105
	(0.0000277)	(0.0002493)	(0.0006925)	(0.0002493)
MA_2	−0.1030235	0.3101904	−0.4837626	0.6842105
	(0.0001220)	(0.0003672)	(0.0006925)	(0.0002493)

Notes: Values in parentheses are the mean squared errors of the estimates. Results are based on 100 simulations.

TABLE II

AVERAGES OF ESTIMATES OF FREQUENCIES (IN MULTIPLES OF π) WITH $d = 4$

Noise	λ_1/π	μ_1/π	λ_2/π	μ_2/π
i.i.d. N(0, 1)	−0.0994737 (0.0000429)	0.3018421 (0.0000132)	−0.5 (0.0)	0.6981579 (0.0000132)
i.i.d. N(0, 4)	−0.0993555 (0.0000432)	0.3018260 (0.0000133)	−0.5 (0.0)	0.6981740 (0.0000133)
MA_1	−0.0993555 (0.0000432)	0.3018260 (0.0000133)	−0.5 (0.0)	0.6981740 (0.0000133)
MA_2	−0.0983852 (0.0001166)	0.3002392 (0.0000573)	−0.5 (0.0)	0.6973684 (0.0000069)

Notes: Values in parentheses are the mean squared errors of the estimates. Results are based on 100 simulations.

TABLE III

AVERAGES OF ESTIMATES OF FREQUENCIES (IN MULTIPLES OF π) WITH $d = 4,\ N = 39$

Noise	λ_1/π	μ_1/π	λ_2/π	μ_2/π
MA_2	−0.1003205 (0.0000095)	0.03 (0.0000066)	−0.5 (0.0)	0.6987179 (0.0000016)

Notes: Values in parentheses are the mean squared errors of the estimates. Results are based on 100 simulations.

which p is correctly estimated 94 times and overestimated 6 times. The average of the estimated frequencies when p is correctly estimated is shown in Table II.

We note that $N = 19$ is a small number. Even so, our choices of α, β, c and d yielded fairly satisfactory results. We expect the estimation for p and the frequencies become better with a larger N since the peak near a hidden frequency will become stronger. In fact, we used $N = 39$ for the MA_2 noise and $d = 4$, which gave the worst estimation results for $N = 19$. Again, we ran 100 simulations. The number p was correctly estimated 100 times and the averages of estimated frequencies and the mean squared of errors are provided in Table III. The results become better.

APPENDIX

We provide proofs of the theorems in this appendix. Proof of Theorem 1 is the longest one and is therefore postponed. Let us first use Theorem 1 to prove Theorems 2, 3 and 4.

Let $A_j > 0$, $\phi_j \in [0, 2\pi)$ be such that $C_j = A_j \cos\phi_j$, $D_j = A_j \sin\phi_j$ for $j = 1, 2, \ldots, p$, and extend λ_j, μ_j and ϕ_j to $j = -1, -2, \ldots, -p$ by $\lambda_j = -\lambda_{-j}$, etc. Define for $j = 0, \pm1, \pm2, \ldots, \pm p$, $B_0 = 0$, $B_j = A_{|j|}e^{i\phi_j}$. We rewrite $y(m, n)$ in model (1) as

$$y(m, n) = \frac{1}{2} \sum_{j=-p}^{p} B_j e^{i(\lambda_j m + \mu_j n)} + x(m, n)$$

It is obvious that the periodogram of $y(m, n)$ can be decomposed into three parts

$$I_N(\lambda, \mu; y) = I_N(\lambda, \mu; x) + \frac{1}{(4\pi N)^2} \left| \sum_{k=-p}^{p} B_k H_N(\lambda_k - \lambda) H_N(\mu_k - \mu) \right|^2 + \frac{1}{(2\pi N)^2} \mathrm{Re} \left(\sum_{m=1}^{N} \sum_{n=1}^{N} x(m, n) e^{-i(m\lambda + n\mu)} \sum_{k=-p}^{p} B_k H_N(\mu_k - \lambda) H_N(\lambda_k - \mu) \right) \tag{4}$$

where $\mathrm{Re}(\cdot)$ represents the real part of a complex number and

$$H_N(x) = \sum_{n=1}^{N} e^{inx}$$

$H_N(x)$ has the following property

$$|H_N(x)| = \left| \frac{\sin(Nx/2)}{\sin(x/2)} \right| \begin{array}{ll} = N & \text{if } x = 0, \bmod 2\pi \\ \leqslant 1/|\sin(x/2)| & \text{for } x \neq 0, \bmod 2\pi \end{array} \tag{5}$$

PROOF OF THEOREM 2. For any fixed pair (λ, μ), let us first consider $(\lambda, \mu) \neq (\lambda_j, \mu_j)$, $\forall j$. Then from (5),

$$\left| \sum_{k=-p}^{p} B_k H_N(\lambda_k - \lambda) H_N(\mu_k - \mu) \right| = O(N)$$

which, together with (4) and Theorem 1, implies $I_N(\lambda, \mu; y) = O(\ln \ln N)$. Now, suppose $(\lambda, \mu) = (\lambda_j, \mu_j)$ for some $j = 1, \ldots, p$. Then

$$\frac{1}{N^2} \left| \sum_{k=-p}^{p} B_k H_N(\lambda_k - \lambda) H_N(\mu_k - \mu) \right| \to |B_j| = A_j$$

We see (4) is dominated by the second term and

$$I_N(\lambda, \mu; y) = \left(\frac{A_j N}{4\pi} \right)^2 + o(N^2)$$

PROOF OF THEOREM 3. It is well known that $|H_N(x)|$ is symmetric in $(-\infty, \infty)$ and decreasing in $[0, \pi/N]$. Thus

$$\inf_{|x| \leqslant \pi/N} |H_N(x)| = |H_N(\pi/N)| = \frac{1}{\sin(\pi/2N)} > 2N/\pi$$

If follows, for any j,

$$\inf_{(\lambda,\mu) \in \Delta_{j,1}} |B_j H_N(\lambda_j - \lambda) H_N(\mu_j - \mu)| > \frac{4|B_j| N^2}{\pi^2}.$$

For any $k \neq j$, we get from (5) that

$$\sup_{(\lambda,\mu) \in \Delta_{j,1}} |B_k H_N(\lambda_k - \lambda) H_N(\mu_k - \mu)| = O(N)$$

Therefore, for sufficiently large N,

$$\inf_{(\lambda,\mu)\in\Delta_{j,1}} \left| \sum_{k=-p}^{p} B_k H_N(\lambda_k - \lambda) H_N(\mu_k - \mu) \right| > \frac{4|B_j|N^2}{\pi^2}$$

Applying this inequality, (4) and (5) gives

$$\begin{aligned}\inf_{(\lambda,\mu)\in\Delta_{j,1}} I_N(\lambda, \mu; y) \geqslant & \inf_{(\lambda,\mu)\in\Delta_{j,1}} (4\pi N)^{-2} \left| \sum_{k=-p}^{p} B_k H_N(\lambda_k - \lambda) H_N(\mu_k - \mu) \right|^2 \\ & - \sup_{(\lambda,\mu)\in\Delta_{j,1}} (\pi N)^{-1} \sqrt{I_N(\lambda, \mu; x)} \left| \sum_{k=-p}^{p} B_k H_N(\lambda_k - \lambda) H_N(\mu_k - \mu) \right| \\ & > A_j^2 N^2/\pi^6.\end{aligned}$$

For any $(\lambda, \mu) \in \Delta_\alpha^c$, and for any $j = 1, 2, \ldots, p$, at least one of $|\lambda_j - \lambda|$, $|\mu_j - \mu|$ is greater than π/N^α. Since

$$\sup_{\pi N^{-\alpha}\leqslant|x|\leqslant 2\pi-\pi N^{-\alpha}} |H_N(x)| \leqslant \frac{1}{\sin(\pi/2N^\alpha)} \leqslant \frac{4N^\alpha}{\pi}$$

we get

$$\sup_{(\lambda,\mu)\in\Delta_\alpha^c} |H_N(\lambda - \lambda_j) H_N(\mu - \mu_j)| \leqslant N \cdot \frac{4N^\alpha}{\pi}$$

Consequently, $I_N(\lambda, \mu; y)$ is dominated by

$$\frac{1}{(4\pi N)^2} \left(\sum_{j=-p}^{p} |B_j| N \frac{4N^\alpha}{\pi} \right)^2 = O\left(N^{2\alpha}\right)$$

The second assertion follows.

Proof of Theorem 4. It follows Theorem 3 that when N is sufficiently large Ω is a subset of Δ_α, and hence

$$\Omega = \bigcup_{j=1}^{p} (\Delta_{j,\alpha} \cap \Omega)$$

Write $S_j = \Delta_{j,\alpha} \cap \Omega$, $j = 1, \ldots, p$. Then each S_j contains $\Delta_{j,1}$ and is not empty. The diameter of S_j is bounded by that of $\Delta_{j,\alpha}$ which is $2\sqrt{2}\pi/N^\alpha$. For any $(\lambda, \mu) \in \Omega\backslash S_j$, it belongs to some $(\lambda, \mu) \in S_k$ for a $k \neq j$. For any $(\lambda', \mu') \in S_j$, using the Triangle Inequality twice, gives

$$\begin{aligned}&\rho((\lambda, \mu), (\lambda', \mu')) \\ &\quad \geqslant \rho((\lambda', \mu'), (\lambda_k, \mu_k)) - \rho((\lambda_k, \mu_k), (\lambda, \mu)) \\ &\quad \geqslant \rho((\lambda_j, \mu_j), (\lambda_k, \mu_k)) - \rho((\lambda_j, \mu_j), \lambda', \mu')) - \rho((\lambda_k, \mu_k), (\lambda, \mu))\end{aligned}$$

We see that the diameter of the set $\{(\lambda, \mu)\} \cup S_j$ is greater than or equal to

$$\rho((\lambda_j, \mu_j), (\lambda_k, \mu_k)) - 2\frac{2\sqrt{2}\pi}{N^\alpha}$$

which exceeds $2\sqrt{2}\pi/N^\alpha$ when N is sufficiently large. Thus for large N, each S_j is a cluster by definition and there are exactly p clusters. For any (λ, μ) in a cluster, say S_j, the distance between (λ, μ) and (λ_j, μ_j) is no larger than the diameter of S_j. The theorem is proved. ■

We need the following lemma for the proof of Theorem 1, which is a two-dimensional version of Theorem 8(i) of Lai and Wei (1982). The proof being the same as in Lai and Wei (1982) is omitted.

LEMMA 1. *Let* $\{\epsilon(m, n): m, n \in \mathbb{Z}\}$ *be a double array of independent random variables such that*

$$\mathrm{E}\epsilon(m, n) = 0,\ \mathrm{E}\epsilon^2(m, n) = 1. \sup_{m,n} \mathrm{E}|\epsilon(m, n)|^r < \infty$$

for some constant $r > 2$.

For $N \geqslant 1$, *let* $\{a_N(m, n): m, n \in \mathbb{Z}\}$ *be a double array of constants such that*

$$A_N = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} a_N^2(m, n) < \infty \qquad \lim_{N\to\infty} A_N = \infty$$

and

$$\sup_{m,n} a_N^2(m, n) = o\,(A_N(\ln A_N)^{-\rho}) \qquad \forall \rho > 0$$

Suppose also there exist constants $\alpha_i > 0$, $d > 2/r$, *such that for some* $M_0 > 0$ *and all* $N > M \geqslant M_0$,

$$\sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} (a_N(m, n) - a_M(m, n))^2 \leqslant \left(\sum_{i=M+1}^{N} \alpha_i\right)^d \tag{6}$$

and as $N \to \infty$,

$$\left(\sum_{i=M_0}^{N} \alpha_i\right)^d = O\,(A_N)$$

Let

$$S_N = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} a_N(m, n)\epsilon(m, n). \qquad \sigma_N^2 = \mathrm{Var}(S_N) = A_N$$

Then

$$\limsup_{N\to\infty} \frac{|S_N|}{(2\sigma_N^2 \ln\ln \sigma_N^2)^{1/2}} \leqslant 1$$

PROOF OF THEOREM 1. Define

$$S_N^+(\theta) = \sum_{m=1}^{N} \sum_{n=1}^{N} x(m, n)\{\cos(m\lambda + n\mu) + \sin(m\lambda + n\mu)\}$$

$$S_N^-(\theta) = \sum_{m=1}^{N} \sum_{n=1}^{N} x(m, n)\{\cos(m\lambda + n\mu) - \sin(m\lambda + n\mu)\}$$

where $\theta = (\lambda, \mu)$. Then

$$\mathrm{Var}(S_N^{\pm}) = (2\pi N)^2 \mathrm{E}(I_N(\theta))$$
$$\pm\, 2\,\mathrm{Cov}\left(\sum_{m=1}^{N}\sum_{n=1}^{N} x(m,\, n)\cos(m\lambda + n\mu),\ \sum_{m=1}^{N}\sum_{n=1}^{N} x(m,\, n)\sin(m\lambda + n\mu)\right)$$

Analogous to the time series case, the periodogram is asymptotically unbiased, and the real part and the imaginary part of the finite Fourier transformation of $x(m, n)$ are asymptotically uncorrelated. Therefore,

$$(2\pi N)^{-2}\,\mathrm{Var}(S_N^{\pm}) \to f(\lambda,\, \mu)$$

Since

$$I_N(\theta) = \frac{1}{2(2\pi N)^2}((S_N^{+})^2 + (S_N^{-})^2)$$

it suffices to show

$$\limsup_{N\to\infty} \frac{|S_N^{\pm}(\theta)|}{\sqrt{2\,\mathrm{Var}(S_N^{\pm}(\theta))\ln\ln \mathrm{Var}(S_N^{\pm}(\theta))}} \leqslant 1 \qquad \text{for any } \theta \tag{8}$$

$$\limsup \frac{\max_\theta |S_N^{\pm}|}{\sqrt{N^2 \ln(N^2)}} \leqslant 2\pi\sqrt{7|f|} \tag{9}$$

We only prove (8) and (9) for S_N^{+}. Proofs of (8) and (9) for S_N^{-} are similar. Let us rewrite $S_N^{+}(\theta)$ as

$$S_N^{+}(\theta) = \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} a_N(i,\, j,\, \theta)\epsilon(i,\, j)$$

where

$$a_N(i,\, j;\, \theta) = \sum_{m=1}^{N}\sum_{n=1}^{N} c(m-i,\, n-j)\{\cos(m\lambda + n\mu) + \sin(m\lambda + n\mu)\}$$

Then

$$\|a\| = \sup_{N,i,j,\theta} |a_N(i,\, j,\, \theta)| \leqslant 2\sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} |c(i,\, j)| < \infty \tag{10}$$

$$A_N \stackrel{\text{def}}{=} \sum_{i,j=-\infty}^{\infty} a_N^2(i,\, j;\, \theta) = \mathrm{Var}(S_N^{+}) < \infty \qquad \text{and} \qquad \frac{A_N}{(2\pi N)^2} \to f(\theta)$$

Note also, for any $N > M > 0$,

$$\sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty}(a_N(m, n; \theta) - a_M(m, n; \theta))^2 = \mathrm{E}(S_N^+(\theta) - S_M^+(\theta))^2$$

$$= \int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\left|\sum_{m,n\in D_N}(\cos(m\lambda + n\mu) + \sin(m\lambda + n\mu))\exp(\mathrm{i}(m\tilde{\lambda} + n\tilde{\mu}))\right|^2 f(\tilde{\lambda}, \tilde{\mu})\,\mathrm{d}\tilde{\lambda}\,\mathrm{d}\tilde{\mu}$$

$$\leqslant \|f\|\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\left|\sum_{m,n\in D_N}(\cos(m\lambda + n\mu) + \sin(m\lambda + n\mu))\exp(\mathrm{i}(m\tilde{\lambda} + n\tilde{\mu}))\right|^2 \mathrm{d}\tilde{\lambda}\,\mathrm{d}\tilde{\mu}$$

$$= \|f\|(2\pi)^2\sum_{m,n\in D_N}(\cos(m\lambda + n\mu) + \sin(m\lambda + n\lambda))^2$$

$$\leqslant 2(2\pi)^2\|f\|(N^2 - M^2) \leqslant \sum_{i=M+1}^{N}\alpha_i$$

where $D_N = \{(m, n) \in \mathbb{Z}^2,\ M < m \leqslant N,\ \text{or}\ M < n \leqslant N\}$, $\alpha_i = 4(2\pi)^2\|f\|i$, $i = M + 1, \ldots, N$. Applying Lemma 1 with $d = 1 > 2/r$, gives (8).

To prove (9), we need to truncate $\epsilon(m, n)$. Let

$$\tilde{\epsilon}(i, j) = \tilde{\epsilon}(i, j, \theta) = \epsilon(i, j)\{|a_N(i, j, \theta)\epsilon(i, j)| < N/\ln(N^2)\}$$

$$\tilde{S}_N(\theta) = \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty}a_N(i, j, \theta)\tilde{\epsilon}(i, j) \qquad \text{and} \qquad S_N^* = \tilde{S}_N(\theta) - \mathrm{E}(\tilde{S}_N(\theta))$$

where, for brevity, $\{|a_N(i, j, \theta)\epsilon(i, j)| < N/\ln(N^2)\}$ denotes the indicator function of the corresponding event.

We will first show, for any $\alpha > 7/2$,

$$\limsup_{N\to\infty}\frac{\max_\theta|S_N^*(\theta)|}{\sqrt{N^2\ln(N^2)}} \leqslant 2\pi\sqrt{2\alpha\|f\|} \tag{11}$$

For this end, let us show for any fixed θ, and for any $\beta > \alpha > 7/2$,

$$P(|S_N^*(\theta)| > 2\pi(2\beta\|f\|N^2\ln(N^2))^{0.5}) \leqslant 2N^{-2\alpha} \tag{12}$$

for all $N > N_0$, where N_0 does not depend on θ.

Since $\mathrm{E}\tilde{\epsilon}(i, j)^2 \leqslant \mathrm{E}\epsilon(i, j)^2 = 1$,

$$\mathrm{E}(S_N^*(\theta)^2) \leqslant \mathrm{E}(S_N(\theta)^2) = \int_{-\pi}^{\pi}\int_{\pi}^{\pi}\left|\sum_{m=1}^{N}\sum_{n=1}^{N}\mathrm{e}^{-\mathrm{i}(m\lambda+n\mu)}\right|^2 f(\lambda, \mu)\,\mathrm{d}\lambda\,\mathrm{d}\mu \leqslant (2\pi)^2\|f\|N^2$$

and $\sup_{i,j}|a_N(i, j, \theta)|\ |\tilde{\epsilon}(i, j) - \mathrm{E}\tilde{\epsilon}(i, j)| \leqslant 2N/\ln(N^2)$, we can apply Lemma 3 (i) of Lai and Wei (1982, p. 329) to $S_N^*(\theta)$ with

$$A = (2\pi)^2\|f\|N^2 \qquad c = \frac{1}{\pi\|f\|^{0.5}\ln(N^2)} \qquad \xi = (2\beta\ln(N^2))^{0.5}$$

to obtain that, for a sufficiently large N such that $c\xi < 1$,

$$
\begin{aligned}
&P(|S_N^*(\theta)| > 2\pi(2\beta\|f\|\, N^2 \ln(N^2))^{0.5}) \\
&\quad \leqslant 2\exp\left\{-\beta\ln(N^2)\left(1 - \frac{1}{2\pi\|f\|^{0.5}\ln(N^2)}(2\beta\ln(N^2))^{0.5}\right)\right\} \\
&\quad \leqslant 2\exp\{-\alpha \ln N^2\} = 2N^{-2\alpha}
\end{aligned}
$$

Now, choose a q such that $3 < q < (\alpha - 0.5)$ and divide $[-\pi, \pi] \times [-\pi, \pi]$ into $K_N = [N^q]^2$ equal-sized squares, each with a width $2\pi/[N^q]$, where $[N^q]$ is the integer part of N^q. Denote the squares by $\Delta_1, \Delta_2, \ldots, \Delta_{K_N}$ and their centers by $\theta_1, \theta_2, \ldots, \theta_{K_N}$. Note that $\mathrm{E}|\tilde{\epsilon}(i, j)| \leqslant \mathrm{E}|\epsilon(i, j)| \leqslant 1$ and

$$
\mathrm{E}\left(\sup_{\theta\in\Delta_k, k} |S_N^*(\theta) - S_N^*(\theta_k)|\right) \leqslant 2\sum_i \sum_j \sup_{\theta\in\Delta_k, k} |a_N(i, j, \theta) - a_N(i, j, \theta_k)|
$$

Since

$$
\begin{aligned}
&|\cos(m\lambda + n\mu) + \sin(m\lambda + n\mu) - \cos(m\lambda_k + n\mu_k) - \sin(m\lambda_k + n\mu_k)| \\
&\quad \leqslant 2|m(\lambda - \lambda_k) + n(\mu - \mu_k)| \\
&\quad \leqslant 2\sqrt{2}\, 2\pi[N^q]^{-1}\sqrt{m^2 + n^2} \qquad \forall k \text{ and } (\lambda, \mu) \in \Delta_k
\end{aligned}
$$

where $\theta_k = (\lambda_k, \mu_k)$, then

$$
\begin{aligned}
&\sup_{\theta\in\Delta_k, k} |a_N(i, j, \theta) - a_N(i, j, \theta_k)| \\
&\quad \leqslant \sqrt{2}\, 8\pi[N^q]^{-1} \sum_{m=1}^{N}\sum_{n=1}^{N} \sqrt{m^2 + n^2}\, |c(m - i, n - j)|
\end{aligned}
$$

From these inequalities and $\sum_{i=1}^N \sum_{j=1}^N \sqrt{m^2 + n^2} = O\,(N^3)$, and the fact that the array $c(i, j)$ is absolutely summable, we obtain

$$
\mathrm{E}\left(\sup_{\theta\in\Delta_k, k} |S_N^*(\theta) - S_N^*(\theta_k)|\right) = O\,(N^{3-q})
$$

Since $q > 3$, Markov Inequality and the Borel–Cantelli Lemma imply

$$
\sup_{\theta\in\Delta_k, k} |S_N^*(\theta) - S_N^*(\theta_k)| = o\,(N) \tag{13}
$$

From (12),

$$
\begin{aligned}
&P\left(\max_k |S_N^*(\theta_k)| > 2\pi\sqrt{2\beta\|f\| N^2 \ln(N^2)}\right) \\
&\quad \leqslant \sum_{k=1}^{K_N} P(|S_N^*(\theta_k)| > 2\pi\sqrt{2\beta\|f\| N^2 \ln(N^2)}) \\
&\quad \leqslant K_N \cdot 2N^{-2\alpha} = O\,(N^{2q-2\alpha}) \qquad \text{for any } \beta > \alpha
\end{aligned}
$$

Since $2\alpha - 2q > 1$, the Borel–Cantelli Lemma implies

$$
\limsup_{N\to\infty} \frac{\max_k |S_N^*(\theta_k)|}{\sqrt{N^2 \ln(N^2)}} \leqslant 2\pi/\sqrt{2\beta\|f\|} \tag{14}
$$

Since (14) is true for any $\beta > \alpha$, it is true for α. The inequality (11) now follows from (13) and (14).

Next, we show that $\mathrm{E}\tilde{S}_N(\theta)$ is negligible. Since $\mathrm{E}\epsilon(i, j) = 0$,

$$\mathrm{E}\tilde{S}_N(\theta) = \mathrm{E}\sum_i \sum_j a_N(i, j, \theta)\epsilon(i, j)\{|a_N(i, j, \theta)\epsilon(i, j)| \geqslant N/\ln(N^2)\}$$

Observing

$$\sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \sup_\theta |a_N(i, j, \theta)| \leqslant 2N^2 \sum_i \sum_j |c(i, j)|$$

and applying (10), gives

$$\begin{aligned}
&\sup_\theta |\mathrm{E}\tilde{S}_N(\theta)| \\
&\leqslant \sum_i \sum_j \mathrm{E}(\sup_\theta |a_N(i, j, \theta)\epsilon(i, j)|\{\sup_\theta |a_N(i, j, \theta)\epsilon(i, j)| > N/\ln(N^2)\}) \qquad (15)\\
&\leqslant \sum_i \sum_j \left(\frac{N}{\ln N^2}\right)^{1-r} \mathrm{E}(\sup_\theta |a_N(i, j, \theta)|\ |\epsilon(i, j)|)^r \\
&\leqslant \left(\frac{N}{\ln N^2}\right)^{1-r} \sup_{i,j} \mathrm{E}|\epsilon(i, j)|^r \|a\|^{r-1} \sum_i \sum_j \sup_\theta |a_\theta(i, j, \theta)| \\
&= O\,(N^{3-r}(\ln N)^{r-1})
\end{aligned}$$

To complete the proof, it suffices to show that, with probability 1,

$$\sup_\theta |S_N(\theta) - \tilde{S}_N(\theta)| = o\,(N) \qquad (16)$$

Since

$$\begin{aligned}
&\sup_\theta |S_N(\theta) - \tilde{S}_N(\theta)| \\
&\leqslant \sup_\theta |\sum_i \sum_j a_N(i, j, \theta)\epsilon(i, j)\{|a_N(i, j, \theta)\epsilon(i, j)| > N/\ln(N^2)\}
\end{aligned}$$

then

$$\begin{aligned}
&\mathrm{E}(\sup_\theta |S_N(\theta) - \tilde{S}_N(\theta)|) \\
&\leqslant \sum_i \sum_j \mathrm{E}(\sup_\theta |a_N(i, j, \theta)\epsilon(i, j)|\{\sup_\theta |a_N(i, j, \theta)\epsilon(i, j)| > N/\ln(N^2)\})
\end{aligned}$$

We see from the inequalities below (15) that

$$\mathrm{E}(\sup_\theta |S_N(\theta) - \tilde{S}_N(\theta)|) = O\,(N^{3-r}(\ln N)^{r-1})$$

Since $r > 3$, the Borel–Cantelli Lemma implies (16). The proof is completed. ∎

ACKNOWLEDGEMENT

We thank the referee for critical reading and constructive suggestions that have improved the presentation of this work.

REFERENCES

AN, H., CHEN, Z. and HANNAN, E. J. (1983) The maximum of periodogram. *Journal of Multivariate Analysis* 13, 383–400.

BANSAL, N., HAMEDANI, G. G. and ZHANG, H. (1999) Non-linear regression with multidimensional indices. *Statistics and Probability Letters* 45, 175–86.

BRILLINGER, D. (1981) *Time Series: Data Analysis and Theory* (Expanded edn). San Francisco: Holder-Day, Inc.

—— (1987) Fitting cosines: some procedures and some physical examples. In *Applied Probability, Stochastic Processes and Sampling Theory* (eds. I. MacNeill and G. J. Umphrey). Dordrecht: Reidel, 75–100.

CHEN, Z. (1988) Consistent estimates for hidden frequencies in a linear process. *Advances in Applied Probability* 20, 295–314.

FRANCOS, J. M., MEIRI, A. Z. and PORAT, B. (1993) A united texture model based on a 2-D Wold like decomposition. *IEEE Trans. Signal Processing* 41, 2665–78.

FISHER, R. A. (1929) Tests of significance in harmonic analysis. *Proceedings of the Royal Society, Series A*. 125, 54–9.

HARTLEY, H. O. (1949) Tests of significance in harmonic analysis. *Biometrika* 36, 194–201.

HE, S. (1995) Uniform convergence for weighted periodogram of stationary linear random fields. *Chin. Ann. of Math.* 16B, (3), 331–40.

HELSON, H. and LOWDENSLAGER, D. (1958) Prediction theory and Fourier series in several variables. *Acta Mathematica* 99, 165–202.

—— and —— (1962) Prediction theory and Fourier Series in several variables. *Acta Mathematica* 106, 175–213.

KUNDU, D. and MITRA, A. (1996) Asymptotic properties of least squares estimates of 2-D exponential signals. *Multidimensional Systems and Signal Processing* 7, 135–50.

LAI, T. L. and WEI, C. Z. (1982) A law of the iterated logarithm for double array of independent random variables with applications to regression and time series models. *Ann. Probability* 10, 320–35.

PRIESTLEY, M. B. (1962a) The analysis of stationary processes with mixed spectra-I. *Journal of the Royal Statistical Society, Series B* 24, 215–33.

—— (1962b) The analysis of stationary processes with mixed spectra-II. *Journal of the Royal Statistical Society, Series B* 24, 511–29.

—— (1964) The analysis of two-dimensional stationary processes with discontinuous spectra. *Biometrika* 51, 195–217.

—— (1997) Detection of periodicities. In *Applications of Time Series Analysis in Astronomy and Meteorology* (eds. T. Subba Rao, M. B. Priestley and O. Lessi). London: Chapman and Hall, 65–88.

QUINN, B. G. (1989) Estimating the number of terms in a sinusoidal regression. *Journal of Time Series Analysis* 10, 71–6.

RAO, C. R., ZHAO, L. and ZHOU, B. (1994) Maximum likelihood estimation of 2-D superimposed exponential signals. *IEEE Trans. Acoust., Speech, Signal Processing* 42 (7), 1795–802.

RIPLEY, B. (1981) *Spatial Statistics*. New York: John Wiley & Sons.

SCHUSTER, A. (1898) On the investigation of hidden periodicities with application to a supposed 26-day period of meteorological phenomena. *Terr. Mag. Atmos. Elect.* 3, 13–41.

WANG, X. B. (1993) An AIC type estimator for the number of cosinusoids. *Journal of Time Series Analysis* 14, 433–40.

WHITTLE, P. (1952) The simultaneous estimation of time series harmonic components and covariance structure. *Trabafos Estadistica* 3, 43–57.

—— (1954) The statistical analysis of a Seiche record. *J. Marine Research* 13, 76–100.

YUAN, J. and SUBBA RAO, T. (1992) Classification of textures using second order spectra. *Journal*

of Time Series Analysis 13, (6), 547–62.
—— and —— (1993) Spectrum estimation for random fields with applications to Markov modeling and texture classification. In *Markov Random Fields, Theory and Applications* (eds R. Chellappa and A. K. Jain). Boston: Academic Press.