Statistica Sinica 8(1998), 559-570

ON STRONG CONSISTENCY OF A 2-DIMENSIONAL FREQUENCY ESTIMATION ALGORITHM

B. Q. Miao, Y. Wu* and L. C. Zhao

*University of Sciences and Technology of China and * York University*

Abstract: In this paper, we prove that an algorithm proposed in Rao, Zhao and Zhou (1993) gives a strongly consistent estimation of 2-dimensional frequencies.

Key words and phrases: Estimation, 2-dimensional frequencies, signal model, strong consistency.

1. Introduction

Consider the following signal model

$$y(m,n)=\sum_{k=1}^{p}x_k e^{j(m\mu_k+n\nu_k)}+w(m,n),\quad m=0,1,\ldots,M-1,\ \ n=0,1,\ldots,N-1,$$

where $y(m,n)$ consists of p distinct noise contaminated 2-D sinusoids in the (μ,ν)-plane and $w(m,n)$'s are independent and identically distributed random variables with zero mean and variance σ^2. This signal model can be found in synthetic aperture radar imaging, frequency and wave number estimation in array signal processing and nuclear magnetic resonance imaging and so on. To recover the p 2-D sinusoids, the techniques often used are periodogram, autoregressive spectral estimation and maximum entropy spectral estimation etc. (See, say Dudgeon and Merseresau (1984).) But all these methods require an exhaustive search in a 2-D space to obtain the estimates of the 2-D angular frequencies and need plenty of data. In the paper by Kung, Arun and Rao (1993) and Shaw and Kumaresan (1986), two sets of 1-D frequencies $\{\mu_k\}$ and $\{\nu_k\}$ are estimated separately instead of one set of 2-D frequencies $\{(\mu_k,\nu_k)\}$. Hua (1992) proposed a new algorithm by a matrix enhancement and matrix pencil (MEMP) approach. However, in the set $\{\mu_k\}$ (or $\{\nu_k\}$), two or more components may be equal, but they are treated as different parameters in the estimation of the set $\{\mu_k\}$ (or $\{\nu_k\}$). This concludes that the number of parameters estimated by MEMP may be larger than is necessary. In pairing the components in the sets $\{\mu_k\}$ and $\{\nu_k\}$, the computation cost by MEMP in an optimum pairing process is proportional to $p!$. In order to reduce the cost, Hua (1992) suggested a pairing process with the computation cost proportional to p^2 but its performance is not clear. As pointed out by Hua (1992), it is desirable to have some a priori information about $\{\mu_k\}$ (or $\{\nu_k\}$) in order to get the best mate. Rao, Zhao and Zhou (1993) developed an algorithm for 2-D frequency estimation. This approach does not need a priori information on $\{\mu_k\}$ (or $\{\nu_k\}$) for the pairing. The pairing process is consistent and its computational cost is only proportional to $q_1 q_2$, where q_1 and q_2 are the numbers of distinct components in the set $\{\mu_k\}$ and $\{\nu_k\}$, respectively. Besides, this approach has the inherent capability of determining the number of signals and can be easily extended to estimate higher dimensional frequencies. The simulation showed that the algorithm works well (see Rao, Zhao and Zhou (1993)).

In this paper, the theoretical foundation for this algorithm will be established. In Section 2, the algorithm is introduced briefly. In Section 3, it is proved that the algorithm gives strongly consistent estimates of the 2-D frequencies.

2. Procedure of Frequency Estimation

The algorithm by Rao, Zhao and Zhou (1993) is divided into two stages. First $\{\mu_k\}$ and $\{\nu_k\}$ are estimated separately, then they are paired together to get the p 2-D frequencies. In the following $\sqrt{-1}$ is denoted by j.

For fixed frequencies $\{\nu_k\}$, $y(m,i)$, $1 \le m \le M-1$, are observations of frequencies $\mu_1,\ldots,\mu_p$. Its covariance matrix, say Γ^{iq}, is

$$\Gamma^{iq} = [\gamma^{iq}_{\ell h}],$$
$$\gamma^{iq}_{\ell h} = (1/(M-q)) \sum_{m=q}^{M-1} y(m-\ell,i)y^*(m-h,i), \quad \ell,h = 0,1,\ldots,q,$$

where $i = 0,1,\ldots,N-1$, "*" denotes complex conjugate operator (or complex conjugate transpose operator if y is a vector or matrix), and $q \le T_1 < M-1$. Assume that there are q_1 distinct angular frequencies of $\{\mu_k\}$. Let $\prod_{t=1}^{q_1}(1 - e^{j\mu_t}z) = 1 + g_1 z + \cdots + g_{q_1} z^{q_1}$ and denote $\sum_{k=1}^{p} x_k e^{j(m\mu_k + n\nu_k)}$ by $\xi_{m,n}$. It can be observed that for each fixed i,

$$\xi_{q_1+\ell,i} + g_1\xi_{q_1+\ell-1,i} + \cdots + g_{q_1}\xi_{\ell,i} = 0, \quad \ell = 0,\ldots,M-1-q_1. \tag{2.1}$$

Once $\boldsymbol{g} = (g_{q_1},\ldots,g_1,1)'$ is known, $\{\mu_k\}$ can be obtained by finding the roots of the polynomial equation $\sum_{t=0}^{q_1} g_t z^t = 0$. Let $\boldsymbol{y}_i = (y(0,i),\ldots,y(M-1,i))'$ for each i and G be the $(M-q_1)\times M$ matrix with each row containing the row vector of $\boldsymbol{g}'$ and a number of zeros, where the position of $\boldsymbol{g}'$ is shifted by one element when we go from one row to the next row. By (2.1), $E(G\boldsymbol{y}_i) = \mathbf{0}$. When $\boldsymbol{g}$ is unknown, it may be estimated by minimizing $\boldsymbol{y}_i^* G^* G \boldsymbol{y}_i/(M-q_1)$. It can be shown that $\min \boldsymbol{y}_i^* G^* G\boldsymbol{y}_i/(M-q_1)$ is the smallest eigenvalue of the matrix Γ^{iq_1}. Therefore, when q_1 is unknown, it may be estimated by minimizing the log function of the smallest eigenvalue of Γ^{iq} plus the penalty term (see Bai, Krishnaiah and Zhao (1986)). Since any single Γ^{iq} will not lead to satisfactory estimation, it is natural to define their mean covariance:

$$\Gamma^q = (1/N) \sum_{i=0}^{N-1} \Gamma^{iq}.$$

Write the smallest eigenvalue of Γ^q as S^q and the corresponding unit eigenvector as $\boldsymbol{b}^q = (b_0, b_1, \ldots, b_q)'$, where the prime is the matrix transpose operator. There exists an integer $\hat{q}_1 \le T_1$ such that

$$R^q = \log S^q + qC_N, \quad q = 0, 1, \ldots, T_1, \tag{2.2}$$

is minimized at $q = \hat{q}_1$, where we assume that M and N are of the same magnitude and C_N is some sequence satisfying conditions:

$$\lim_{n\to\infty} C_N = 0, \quad \lim_{n\to\infty} NC_N/\sqrt{\log\log N} = \infty. \tag{2.3}$$

For estimating the number of distinct angular frequencies of $\{\nu_k\}$, similar to above, we can get an integer $\hat{q}_2 \le T_2 < N-1$ by replacing S^q with the smallest eigenvalue $S^{(q)}$ of $\Gamma^{(q)}$ in (2.2), where

$$\Gamma^{(q)} = (1/M) \sum_{k=0}^{M-1} \Gamma^{(k,q)}, \quad \Gamma^{(k,q)} = [\gamma_{\ell h}^{(k,q)}],$$
$$\gamma_{\ell h}^{(k,q)} = (1/(N-q)) \sum_{n=q}^{N-1} y(k, n-\ell) y^*(k, n-h), \quad \ell, h = 0, 1, \ldots, q,$$

and $q \le T_2 < N-1$.

Now $(\hat{q}_1, \hat{q}_2)$ is used as the estimate of (q_1, q_2), where q_1 and q_2 are the numbers of distinct angular frequencies of $\{\mu_k\}$ and $\{\nu_k\}$, respectively.

Theorem 2.1. *Suppose that there exist two positive constants d_1 and d_2 such that $d_1 N \le M \le d_2 N$ and $w(m,n)$, $m, n = 1, 2, \ldots$ are independent and identically distributed random variables with mean zero and finite variance σ^2 such that $E(|w(m,n)|^4 \log^+ |w(m,n)|) < \infty$. Take $\{C_N\}$ satisfying Equation* (2.3)*; then with probability one, we have $(\hat{q}_1, \hat{q}_2) = (q_1, q_2)$ for large N.*

In the following, we assume that q_1 and q_2 are estimated. In order to estimate distinct angular frequencies of μ_k's, say $\mu_1', \ldots, \mu_{q_1}'$, solve the polynomial equation

$$H_{\boldsymbol{b}^{\hat{q}_1}}(z) = \sum_{k=0}^{\hat{q}_1} b_k z^k = 0.$$

Let $\hat{\rho}_k e^{j\hat{\mu}'_k}$, $k = 1, 2, \ldots, \hat{q}_1$, be the solutions, where $\hat{\rho}_k > 0$, $\hat{\mu}'_k \in [0, 2\pi)$, and take $\hat{\mu}'_k$ as the estimates of μ'_k, $k = 1, 2, \ldots, \hat{q}_1$. Note that on average the number of distinct angular frequencies, q_1, is less than p, the total number of 2-D angular frequencies.

In a similar way the estimates of the distinct $\nu'_1, \ldots, \nu'_{\hat{q}_2}$ can be obtained.

After $\hat{q}_1$, $\hat{q}_2$, $\{\mu'_k\}$ and $\{\nu'_k\}$ are obtained, we choose two integers $K(> \hat{q}_1)$ and $L(> \hat{q}_2)$, and define

$$B^{k\ell} = \begin{pmatrix} y(k,\ell) & \cdots & y(k,\ell+L-1) \\ y(k+1,\ell) & \cdots & y(k+1,\ell+L-1) \\ & \cdots & \\ y(k+K-1,\ell) & \cdots & y(k+K-1,\ell+L-1) \end{pmatrix},$$

$$W^{(k,\ell)} = \begin{pmatrix} w(k,\ell) & \cdots & w(k,\ell+L-1) \\ w(k+1,\ell) & \cdots & w(k+1,\ell+L-1) \\ & \cdots & \\ w(k+K-1,\ell) & \cdots & w(k+K-1,\ell+L-1) \end{pmatrix}.$$

Set $\boldsymbol{u}_I(\tau) = (1, e^{j\tau}, \ldots, e^{j(I-1)\tau})'$, $X = \operatorname{diag}(x_1, \ldots, x_p)$, and $F_1^k(\boldsymbol{\eta}) = \operatorname{diag}(e^{jk\eta_1}, \ldots, e^{jk\eta_p})$, where $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_p)'$. It is easy to see that $B^{k\ell}$ can be rewritten as

$$B^{k\ell} = (\boldsymbol{u}_K(\mu_1), \ldots, \boldsymbol{u}_K(\mu_p))F_1^k(\boldsymbol{\mu})XF_1^\ell(\boldsymbol{\nu})(\boldsymbol{u}_L(\nu_1), \ldots, \boldsymbol{u}_L(\nu_p))' + W^{(k,\ell)},$$

where $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_p)'$ and $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_p)'$. Define $\boldsymbol{y}^{k\ell} = \operatorname{vec}((B^{k\ell})')$, $\boldsymbol{w}^{k\ell} = \operatorname{vec}((W^{(k,\ell)})')$, where $\operatorname{vec}(A)$ denotes the vector form of a matrix A. Let "$\otimes$" denote the Kronecker product operator. For $B = \operatorname{diag}(b_1, \ldots, b_p)$, $A_{m\times p} = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_p)$, and $C_{n\times p} = (\boldsymbol{c}_1, \ldots, \boldsymbol{c}_p)$, $\operatorname{vec}(ABC') = (\boldsymbol{c}_1 \otimes \boldsymbol{a}_1, \ldots, \boldsymbol{c}_p \otimes \boldsymbol{a}_p)(b_1, \ldots, b_p)'$. Therefore,

$$\begin{aligned} \boldsymbol{y}^{k\ell} &= ((\boldsymbol{u}_K(\mu_1)\otimes\boldsymbol{u}_L(\nu_1), \ldots, \boldsymbol{u}_K(\mu_p)\otimes\boldsymbol{u}_L(\nu_p))(x_1 e^{j(k\mu_1+\ell\nu_1)}, \ldots, x_p e^{j(k\mu_p+\ell\nu_p)})' \\ &\quad + \boldsymbol{w}^{k\ell} \\ &\hat{=} A(\boldsymbol{\mu},\boldsymbol{\nu})X\boldsymbol{f}^{k\ell}(\boldsymbol{\mu},\boldsymbol{\nu}) + \boldsymbol{w}^{k\ell}, \end{aligned} \tag{2.4}$$

where

$$\begin{aligned} A(\boldsymbol{\mu},\boldsymbol{\nu}) &= (\boldsymbol{a}(\mu_1,\nu_1), \ldots, \boldsymbol{a}(\mu_p,\nu_p)), \quad \boldsymbol{a}(\mu_i,\nu_i) = \boldsymbol{u}_K(\mu_i)\otimes\boldsymbol{u}_L(\nu_i), \\ \boldsymbol{f}^{k\ell}(\boldsymbol{\mu},\boldsymbol{\nu}) &= (e^{j(k\mu_1+\ell\nu_1)}, \ldots, e^{j(k\mu_p+\ell\nu_p)})'. \end{aligned}$$

Set

$$\hat{R} = [(M-K+1)(N-L+1)]^{-1}\sum_{k=0}^{M-K}\sum_{\ell=1}^{N-L}\boldsymbol{y}^{k\ell}(\boldsymbol{y}^{k\ell})^*. \tag{2.5}$$

Write the eigenvalues and corresponding orthonormal (or unit) eigenvectors of $\hat{R}$ as $\lambda_1 \geq \cdots \geq \lambda_{KL}$ and $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_p, \boldsymbol{e}_{p+1}, \ldots, \boldsymbol{e}_{KL}$, respectively. Set $E_n =$

$[\boldsymbol{e}_{p+1}, \ldots, \boldsymbol{e}_{KL}]$ which is called the noise subspace. The p distinct 2-D angular frequencies (μ_i, ν_i) may be obtained by minimizing

$$\boldsymbol{a}^*(\mu, \nu) E_n E_n^* \boldsymbol{a}(\mu, \nu) \tag{2.6}$$

over $\{(\hat{\mu}_k, \hat{\nu}_\ell) : k = 1, \ldots, \hat{q}_1,\ \ell = 1, \ldots, \hat{q}_2\}$. We have the following theorem.

Theorem 2.2. *Suppose that the assumptions made in Theorem* 2.1 *hold. If* $(\tilde{\mu}_1, \tilde{\nu}_1)$, $\ldots$, $(\tilde{\mu}_p, \tilde{\nu}_p)$ *are* p *pairs of* $\{(\hat{\mu}_k, \hat{\nu}_\ell) : k = 1, \ldots, \hat{q}_1,\ \ell = 1, \ldots, \hat{q}_2\}$, *which reach the first* p *minimum values of* (2.6), *then* $(\tilde{\mu}_k, \tilde{\nu}_k) \to (\mu_k, \nu_k)$ *a.s.* $k = 1, 2, \ldots, p$.

Remark 2.1. When $w(m, n)$, $m, n = 1, 2, \ldots$ are independent and identically normally distributed with zero mean, the assumptions on $w(m, n)$ made in Theorem 2.1 are satisfied, which, jointly with other assumptions made in Theorem 2.1, will guarantee the results of Theorem 2.1 and Theorem 2.2.

3. Proofs

Let $\{x_n : n \in \mathcal{Z}_+^d\}$ be a sequence of independent random variables, where $\mathcal{Z}_+^d$, $d \geq 2$, is the positive integer d-dimensional set of lattice points with coordinatewise partial ordering "<". Set $S_n = \sum_{k \leq n} x_k$. For $n = (n_1, \ldots, n_d)$, write $|n| = \prod_{i=1}^d n_i$. In the following, $n \to \infty$ means that $n_i \to \infty$ for all i, $\limsup_n \alpha_n$ is interpreted as $\lim_n \sup_{n<m} \alpha_m$ and similarly for $\liminf_n \alpha_n$.

We need the following two lemmas.

Lemma 3.1. *Let* $\{x_n, n \in \mathcal{Z}_+^d\}$ *be a sequence of independent and identically distributed random variables with zero mean and finite variance* σ^2. *Assume that for any* $n \in \mathcal{Z}_+^d$, $E(x_n^2 (\log^+ |x_n|)^{d-1} / \log^+ \log^+ |x_n|) < \infty$, *where* $\log^+ |b| = \log(\max\{|b|, e\})$. *If* $\{c_n, n \in \mathcal{Z}_+^d\}$ *is a sequence of constants with* $|c_n| = 1$, *then*

$$\limsup_n \sum_{k \leq n} c_k x_k / (2\sigma^2 |n| \log\log |n|)^{1/2} = \sqrt{d} \quad \text{a.s.}$$
$$\liminf_n \sum_{k \leq n} c_k x_k / (2\sigma^2 |n| \log\log |n|)^{1/2} = -\sqrt{d} \quad \text{a.s.}$$

Proof. By Theorem 5 and Lemma 5.1 of Wichura (1973), the lemma follows.

Lemma 3.2. *Let* $\{x_n : n \in \mathcal{Z}_+^d\}$ *be independent random variables with zero mean. Suppose that there exists a constant* $\delta > 0$ *such that for any* $n \in \mathcal{Z}_+^d$, $E(|x_n| (\log^+ |x_n|)^{d+\delta}) \leq c < \infty$ *for a constant* c. *Then* $\lim_{|n| \uparrow \infty} S_n / |n| = 0$ *a.s. where* $|n| \uparrow \infty$ *means* $\prod_{i=1}^d n_i \uparrow \infty$.

Proof. By the theorem of Section 3 of Smythe (1973), the result follows.

It is easy to see that when x_n, $n \in \mathcal{Z}_+^d$ are independent and identically distributed normal variables with zero mean, the assumptions made in both lemmas are satisfied.

The proofs of Theorem 2.1 and Theorem 2.2 are as follows:

Proof of Theorem 2.1. Assume that there are q_1 different elements in the set $\{\mu_1, \ldots, \mu_p\}$, say $\mu'_1, \ldots, \mu'_{q_1}$. Write $L_s = \{k : 1 \le k \le p, \mu_k = \mu'_s\}$, $\quad s = 1, \ldots, q_1$. For fixed ℓ and h,

$$\begin{aligned}
N^{-1}\sum_{i=0}^{N-1}\gamma_{\ell h}^{iq} &= N^{-1}\sum_{i=0}^{N-1}(M-q)^{-1}\sum_{t=q}^{M-1} y(t-\ell,i)y^*(t-h,i)\\
&= N^{-1}\sum_{i=0}^{N-1}(M-q)^{-1}\sum_{k,m=1}^{p} x_k\bar{x}_m e^{ji(\nu_k-\nu_m)}\sum_{t=q}^{M-1} e^{j(t-\ell)\mu_k - j(t-h)\mu_m}\\
&\quad +N^{-1}\sum_{i=0}^{N-1}\sum_{k=1}^{p} x_k e^{ji\nu_k}(M-q)^{-1}\sum_{t=q}^{M-1} e^{j(t-\ell)\mu_k}\bar{w}(t-h,i)\\
&\quad +N^{-1}\sum_{i=0}^{N-1}\sum_{m=1}^{p} \bar{x}_m e^{-ji\nu_m}(M-q)^{-1}\sum_{t=q}^{M-1} e^{-j(t-h)\mu_m}w(t-\ell,i)\\
&\quad +N^{-1}\sum_{i=0}^{N-1}(M-q)^{-1}\sum_{t=q}^{M-1} w(t-\ell,i)\bar{w}(t-h,i)\\
&\hat{=} J_1+J_2+J_3+J_4, \qquad (3.1)
\end{aligned}$$

and

$$\begin{aligned}
J_1 &= \sum_{s=1}^{q_1}\sum_{k,m\in L_s} x_k\bar{x}_m N^{-1}\sum_{i=0}^{N-1} e^{ji(\nu_k-\nu_m)}e^{-j(\ell-h)\mu'_s}\\
&\quad +\sum_{s\ne s_1}\sum_{k\in L_s}\sum_{m\in L_{s_1}} x_k\bar{x}_m N^{-1}\sum_{i=0}^{N-1} e^{ji(\nu_k-\nu_m)}\Big((M-q)^{-1}\sum_{t=q}^{M-1} e^{j(t-\ell)\mu_k - j(t-h)\mu_m}\Big)\\
&\hat{=} J_{11}+J_{12}.
\end{aligned}$$

Since for $k \ne m$, $\mu_k = \mu_m$ implies $\nu_k \ne \nu_m$, $|\sum_{i=0}^{N-1} e^{ji(\nu_k-\nu_m)-j(\ell-h)\mu'_s}|$ is a bounded quantity. Thus,

$$\begin{aligned}
J_{11} &= \sum_{s=1}^{q_1}\sum_{k\in L_s}|x_k|^2 e^{-j(\ell-h)\mu'_s} + \sum_{s=1}^{q_1}\sum_{k,m\in L_s, k\ne m} x_k\bar{x}_m N^{-1}\sum_{i=0}^{N-1} e^{ji(\nu_k-\nu_m)-j(\ell-h)\mu'_s}\\
&= \sum_{s=1}^{q_1}\sum_{k\in L_s}|x_k|^2 e^{-j(\ell-h)\mu'_s} + O(N^{-1}), \quad \text{as } N\to\infty.
\end{aligned}$$

It is easy to see that $J_{12} = O(M^{-1}), \quad$ as $N \to \infty$. Noting that $d_1 N \le M \le d_2 N$, we have

$$J_1 = \sum_{s=1}^{q_1} \sum_{k \in L_s} |x_k|^2 e^{-j(\ell-h)\mu_s'} + O(N^{-1}), \quad \text{as } N \to \infty. \tag{3.2}$$

By Lemma 3.1 and the fact that $|e^{-j(t-h)\mu_k - ji\nu_k}| = 1$, we have

$$\begin{aligned} |J_3| &= \Big| \sum_{m=1}^{p} \bar{x}_m (N(M-q))^{-1} \sum_{i=0}^{N-1} \sum_{t=q}^{M-1} e^{-j(t-h)\mu_m - ji\nu_m} w(t-\ell, i) \Big| \\ &= O\Big((MN)^{-1/2} (\log\log(MN))^{1/2} \Big) \\ &= O\Big(N^{-1} (\log\log N)^{1/2} \Big) \quad \text{a.s.} \end{aligned} \tag{3.3}$$

In the same way, we have

$$|J_2| = O\left(N^{-1} (\log\log N)^{1/2} \right) \quad \text{a.s.} \tag{3.4}$$

To estimate J_4 for $\ell = h$, by Lemma 3.1 we have

$$\begin{aligned} J_4 &= (N(M-q))^{-1} \sum_{i=0}^{N-1} \sum_{t=q}^{M-1} |w(t-\ell, i)|^2 \\ &= \sigma^2 + O\left((MN)^{-1/2} (\log\log(MN))^{1/2} \right) \\ &= \sigma^2 + O\left(N^{-1} (\log\log N)^{1/2} \right) \quad \text{a.s.} \end{aligned} \tag{3.5}$$

In order to estimate J_4 for $\ell \ne h$, we divide the sum into two parts such that each part is the sum of independent and identically distributed random variables. For example, consider the case where $\ell = 2$ and $h = 0$. In this case, we have

$$\begin{aligned} J_4 &= (N(M-q))^{-1} \sum_{i=0}^{N-1} \sum_{t=q}^{M-1} w(t-2, i) \bar{w}(t, i) \\ &= \Big[(N(M-q))^{-1} \sum_{k=0}^{k_1} \sum_{i=1}^{N-1} w(q+4k-2, i) \bar{w}(q+4k, i) \\ &\qquad + (N(M-q))^{-1} \sum_{k=0}^{k_2} \sum_{i=1}^{N-1} w(q+4k-1, i) \bar{w}(q+4k+1, i) \Big] \\ &\quad + \Big[(N(M-q))^{-1} \sum_{k=0}^{k_3} \sum_{i=1}^{N-1} w(q+4k, i) \bar{w}(q+4k+2, i) \\ &\qquad + N(M-q))^{-1} \sum_{k=0}^{k_4} \sum_{i=1}^{N-1} w(q+4k+1, i) \bar{w}(q+4k+3, i) \Big] \\ &\hat{=} J_{41} + J_{42}, \end{aligned}$$

where $k_1 = \lfloor (M-q)/4 \rfloor$, the integer part of $(M-q)/4$, $k_2 = k_3 = k_4 = k_1 - 1$ if $(M-q)/4$ is an integer; $k_2 = k_3 = k_4 = k_1 - 1$ if $M - q = 4k_1 + 1$; $k_2 = k_1$ and $k_3 = k_4 = k_1 - 1$ if $M - q = 4k_1 + 2$; $k_3 = k_2 = k_1$ and $k_4 = k_1 - 1$ if $M - q = 4k_1 + 3$. Both J_{41} and J_{42} are sums of independent and identically distributed random variables. By Lemma 3.1, it follows that

$$\begin{aligned} |J_4| = |J_{41} + J_{42}| &= O\left((MN)^{-1/2}(\log\log(MN))^{1/2}\right) \\ &= O\left(N^{-1}(\log\log N)^{1/2}\right) \quad \text{a.s.} \end{aligned} \tag{3.6}$$

Write

$$\boldsymbol{\omega}(i) = (1, e^{-j\mu_i'}, \ldots, e^{-jq\mu_i'})', \quad i = 1, \ldots, q_1,$$

$$\Omega(q+1, q_1) = [\boldsymbol{\omega}(1), \ldots, \boldsymbol{\omega}(q_1)], \tag{3.7}$$

$$\Psi = \operatorname{diag}\Big(\sum_{k\in L_1} |x_k|^2, \ldots, \sum_{k\in L_{q_1}} |x_k|^2\Big). \tag{3.8}$$

Combining (3.1) to (3.6), it follows that

$$\Gamma^q = \sigma^2 I_{q+1} + \Omega(q+1, q_1)\Psi\Omega^*(q+1, q_1) + O\left(N^{-1}(\log\log N)^{1/2}\right) \quad \text{a.s.} \tag{3.9}$$

which implies that

$$\lim_{N\to\infty} S^q > \sigma^2 \quad \text{a.s.} \tag{3.10}$$

for $q < q_1$ and

$$S^q = \sigma^2 + O\Big(N^{-1}(\log\log N)^{1/2}\Big) \quad \text{a.s.} \tag{3.11}$$

for $q \geq q_1$.

Let us consider two cases.

Case 1. $q > q_1$. For this case, we have, with probability one,

$$\begin{aligned} (R^{q_1} - R^q) &= (\log S^{q_1} - \log S^q + (q_1 - q)C_N) \\ &= O\left(N^{-1}(\log\log N)^{1/2}\right) - (q - q_1)C_N < 0 \end{aligned} \tag{3.12}$$

for large N by (3.11) and (2.3).

Case 2. $q < q_1$. By (3.10) and (2.3), we have

$$\begin{aligned} \lim_{N\to\infty} (R^{q_1} - R^q) &= \lim_{N\to\infty} (\log S^{q_1} - \log S^q + (q_1 - q)C_N) \\ &= \log\sigma^2 - \lim_{N\to\infty} S^q < 0 \quad \text{a.s.} \end{aligned} \tag{3.13}$$

In view of (3.12) and (3.13), we get tha t $\hat{q}_1 = q_1$ with probability one for large N.

Similarly, it can be shown that $\hat{q}_2 = q_2$ with probability one for large N.

Proof of Theorem 2.2. Assume that q_1 and q_2 are known. By (3.9), we have

$$\begin{aligned}\Gamma^{q_1} &= \sigma^2 I_{q_1+1} + \Omega(q_1+1,q_1)\Psi\Omega^*(q_1+1,q_1) + O\left(N^{-1}(\log\log N)^{1/2}\right)\\ &\to \sigma^2 I_{q_1+1} + \Omega(q_1+1,q_1)\Psi\Omega^*(q_1+1,q_1) \hat{=} \Gamma_0, \end{aligned} \tag{3.14}$$

where Ψ and $\Omega(q_1+1,q_1)$ are defined in (3.7) and (3.8). Note that the rank of $\Omega(q_1+1,q_1)\Psi\Omega^*(q_1+1,q_1)$ is q_1. Therefore, the equation $(\Gamma_0 - \sigma^2 I_{q_1+1})\boldsymbol{\beta} = \mathbf{0}$, $\|\boldsymbol{\beta}\| = 1$ has a unique root $\boldsymbol{b}^0$ (up to a complex factor with module one). Now let $\boldsymbol{b}^{q_1}$ be a unit eigenvector corresponding to the smallest eigenvalue of Γ^{q_1}. By (3.14), with probability one, this smallest eigenvalue of Γ^{q_1} is strictly less than other eigenvalues of Γ^{q_1} for large N. Hence, with appropriate choice of a complex factor, $\lim_{N\to\infty}\boldsymbol{b}^{q_1} = \boldsymbol{b}^0$. Let $z_1,\ldots,z_{q_1}$ be the q_1 roots of the polynomial equation

$$H_{\boldsymbol{d}}(z) = \sum_{k=0}^{q_1} d_k z^k = 0, \tag{3.15}$$

where $\boldsymbol{d} = (d_0, d_1, \ldots, d_{q_1})'$. By the definition of $\Omega(q_1+1,q_1)$ and Γ_0, μ_1', $\ldots$, μ_{q_1}' are solutions of the equation (3.15), when $\boldsymbol{d} = (b_0^0,\ldots,b_{q_1}^0)' = \boldsymbol{b}^0$. For $\boldsymbol{d} = \boldsymbol{b}^{q_1}$, the unit eigenvector corresponding to the smalleast eigenvalue of Γ^{q_1}, write the solutions of (3.15) as $\hat\rho_k e^{j\hat\mu_k'}$, $k = 1,\ldots,q_1$. Note that $\boldsymbol{b}^{q_1} \to \boldsymbol{b}^0$. The relation between the roots and coefficients of a polynomial yields $\hat\rho_k e^{j\hat\mu_k'} \to e^{j\mu_k'}$ a.s. $k = 1,\ldots,q_1$ and $\hat\mu_k' \to \mu_k'$ a.s. $k = 1,\ldots,q_1$ with properly rearranged subscripts of $\hat\rho_k e^{j\hat\mu_k'}$.

In the same way, it can be shown that $\hat\nu_k' \to \nu_k'$ a.s. $k = 1,\ldots,q_2$. By (2.4) and (2.5),

$$\begin{aligned}\hat R &= \tfrac{1}{(M-K+1)(N-L+1)} \sum_{k=0}^{M-K}\sum_{\ell=0}^{N-L} (AX\boldsymbol{f}^{k\ell} + \boldsymbol{w}^{k\ell})((\boldsymbol{f}^{k\ell})^* X^* A^* + (\boldsymbol{w}^{k\ell})^*)\\ &\hat{=} Q_1 + Q_2 + Q_3 + \bar Q_3, \end{aligned} \tag{3.16}$$

where

$$Q_1 = A(\boldsymbol{\mu},\boldsymbol{\nu})X\Big\{\tfrac{1}{(M-K+1)(N-L+1)}\Big(\sum_{k=0}^{M-K}\sum_{\ell=0}^{N-L}\boldsymbol{f}^{k\ell}(\boldsymbol{\mu},\boldsymbol{\nu})(\boldsymbol{f}^{k\ell}(\boldsymbol{\mu},\boldsymbol{\nu}))^*\Big)\Big\}X^*A^*(\boldsymbol{\mu},\boldsymbol{\nu}),$$

$$Q_2 = \tfrac{1}{(M-K+1)(N-L+1)} \sum_{k=0}^{M-K}\sum_{\ell=0}^{N-L} \boldsymbol{w}^{k\ell}(\boldsymbol{w}^{k\ell})^*,$$

$$\begin{aligned}Q_3 &= A(\boldsymbol{\mu},\boldsymbol{\nu})X\Big\{\tfrac{1}{(M-K+1)(N-L+1)} \sum_{k=0}^{M-K}\sum_{\ell=0}^{N-L} \boldsymbol{f}^{k\ell}(\boldsymbol{\mu},\boldsymbol{\nu})(\boldsymbol{w}^{k\ell})^*\Big\}\\ &\hat{=} A(\boldsymbol{\mu},\boldsymbol{\nu})XG_{MN}.\end{aligned}$$

It is obvious that

$$Q_1 \to A(\boldsymbol{\mu},\boldsymbol{\nu})XX^*A^*(\boldsymbol{\mu},\boldsymbol{\nu}) \quad \text{as } N \to \infty. \tag{3.17}$$

By Lemma 3.2, we have

$$Q_2 \to \sigma^2 I_{KL} \quad \text{a.s.} \quad \text{as } N \to \infty. \tag{3.18}$$

In order to show that $Q_3 \to 0$ a.s. as $N \to \infty$, we take the (1,1) element of G_{MN} as an example. By Lemma 3.2, we have

$$G_{MN}(1,1) = \tfrac{1}{(M-K+1)(N-L+1)} \sum_{k=0}^{M-K} \sum_{\ell=0}^{N-L} e^{j(k\mu_1+\ell\nu_1)} \bar{w}(k,\ell) \to 0 \quad \text{a.s.}$$

as $N \to \infty$. Therefore,

$$Q_3 \to 0 \;\text{ a.s. and }\; \bar{Q}_3 \to 0 \text{ a.s.} \tag{3.19}$$

as $N \to \infty$. In view of (3.16) to (3.19), it follows that

$$\hat{R} \to A(\boldsymbol{\mu},\boldsymbol{\nu})XX^*A^*(\boldsymbol{\mu},\boldsymbol{\nu}) + \sigma^2 I_{KL} \hat{=} R_0 \quad \text{a.s.} \tag{3.20}$$

as $N \to \infty$. Let a spectral decomposition of R_0 be $R_0 = \sum_{i=1}^{KL} \lambda_i^0 \boldsymbol{e}_i^0 (\boldsymbol{e}_i^0)^*$, where $\lambda_1^0 \geq \cdots \geq \lambda_p^0 > \lambda_{p+1}^0 = \cdots = \lambda_{KL}^0 = \sigma^2$ are the eigenvalues of R^0 and $\boldsymbol{e}_i^0$ is the eigenvector corresponding to the eigenvalue λ_i^0 for $i = 1,\ldots,KL$. Denote the eigenvalues of $\hat{R}$ by $\lambda_1,\ldots,\lambda_{KL}$ and the eigenvector corresponding to the eigenvalue λ_i by $\boldsymbol{e}_i$ for $i = 1,\ldots,KL$. By (3.20), we have $\lambda_i \to \lambda_i^0$ a.s. for $i \leq p$ and $\lambda_i \to \sigma^2$ a.s. for $i = p+1,\ldots,KL$, which implies that

$$E_n E_n^* = \sum_{k=p+1}^{KL} \boldsymbol{e}_k \boldsymbol{e}_k^* \to \sum_{k=p+1}^{KL} \boldsymbol{e}_k^0 (\boldsymbol{e}_k^0)^* \hat{=} E_n^0 (E_n^0)^* \quad \text{a.s.} \tag{3.21}$$

By the definition of $\boldsymbol{a}(\mu_k,\nu_k)$ and E_n^0, it follows that

$$\boldsymbol{a}^*(\mu_k,\nu_k) E_n^0 (E_n^0)^* \boldsymbol{a}(\mu_k,\nu_k) = 0, \quad k = 1,\ldots,p,$$

and

$$\boldsymbol{a}^*(\mu_k,\nu_k) E_n E_n^* \boldsymbol{a}(\mu_k,\nu_k) \to 0 \quad \text{a.s.} \quad k = 1,\ldots,p. \tag{3.22}$$

Therefore,

$$\begin{aligned} &\{\mu_1,\ldots,\mu_p\} = \{\mu_1',\ldots,\mu_{q_1}'\}, \quad \{\nu_1,\ldots,\nu_p\} = \{\nu_1',\ldots,\nu_{q_2}'\}, \\ &\hat{\mu}_k \to \mu_k' \quad \text{a.s.} \quad k = 1,\ldots,q_1, \quad \hat{\nu}_k \to \nu_k' \quad \text{a.s.} \quad k = 1,\ldots,q_2. \end{aligned}$$

For the nontrue match of (μ_i', ν_k'), say $(\mu_t', \nu_\tau') \notin \{(\mu_1, \nu_1), \ldots, (\mu_p, \nu_p)\} \hat{=} \Theta$, (3.21) gives

$$\lim_{N\to\infty} \boldsymbol{a}^*(\mu_t', \nu_\tau') E_n E_n^* \boldsymbol{a}(\mu_t', \nu_\tau') = \boldsymbol{a}^*(\mu_t', \nu_\tau') E_n^0 (E_n^0)^* \boldsymbol{a}(\mu_t', \nu_\tau') > 0 \quad \text{a.s.} \tag{3.23}$$

Write $D = \{(i,k) : 1 \le i \le q_1,\ 1 \le k \le q_2$; then there exists a pair, say $(\mu_t, \nu_t) \in \Theta$, such that $(\mu_i', \nu_k') = (\mu_t, \nu_t)\}$. By (3.22) and (3.23), with probability one,

$$\inf_{(i,k)\notin D} \boldsymbol{a}^*(\mu_i', \nu_k') E_n E_n^* \boldsymbol{a}(\mu_i', \nu_k') > \sup_{(i,k)\in D} \boldsymbol{a}^*(\mu_i', \nu_k') E_n E_n^* \boldsymbol{a}(\mu_i', \nu_k'),$$

for large N, which implies that with probability one, $\boldsymbol{a}^*(\hat{\mu}_i, \hat{\nu}_k) E_n E_n^* \boldsymbol{a}(\hat{\mu}_i, \hat{\nu}_k)$, $(i,k) \in D$, reach the first p minimum values among $\boldsymbol{a}^*(\hat{\mu}_i, \hat{\nu}_k) E_n E_n^* \boldsymbol{a}(\hat{\mu}_i, \hat{\nu}_k)$, $1 \le i \le q_1$, $1 \le k \le q_2$ for large N. Write $(\hat{\mu}_i, \hat{\nu}_k)$, $(i,k) \in D$, as $(\tilde{\mu}_1, \tilde{\nu}_1), \ldots, (\tilde{\mu}_p, \tilde{\nu}_p)$. Then we have proved that $(\tilde{\mu}_k, \tilde{\nu}_k) \to (\mu_k, \nu_k)$ a.s. $k = 1, \ldots, p$.

Acknowledgement

The first author's research was partially supported by the Natural Sciences and Engineering Research Council of Canada and the National Natural Science Foundation of China and Doctorial Foundation of Department of Education of China. The second author's research was supported by the Natural Sciences and Engineering Research Council of Canada. The third author's research was partially supported by the National Natural Science Foundation of China, Doctorial Foundation of the Education Commission of China and Special Foundation of Academia Sinica. The authors are very grateful to Professor C. F. Jeff Wu and an anonymous referee for their valuable comments and suggestions.

References

Bai, Z. D., Krishnaiah, P. R. and Zhao, L. C. (1986). On simultaneous estimation of the number of signals and frequencies under a model with multiple sinusoids. Technical Report 86-37, Center for Multivariate Analysis, University of Pittsburgh.

Dudgeon, D. E. and Merseresau, R. M. (1984). *Multidimensional Digital Signal Processing.* Englewood Cliffs, NJ: Prentice-Hall, Inc.

Hua, Y. (1992). Estimating two-dimensional frequencies by matrix enhancement and matrix pencil. *IEEE Trans. Signal Processing* **40**, 2267-2280.

Kung, S. Y., Arun, K. S. and Rao, D. V. B. (1993). State-space and singular-value decomposition-based approximation methods for the harmonic retrieval problem. *J. Opt. Soc. Amer.* **73**, 1799-1811.

Rao, C. R., Zhao, L. C. and Zhou, B. (1993). A novel algorithm for 2-dimensional frequency estimation. *Conference Record of The Twenty-seventhAsilomar Conference on Signals, Systems and Computers*, Vol. 1 of 2, November 1-3, 1993, IEEE Computer Society Press.

Shaw, A. K. and Kumaresan, R. (1986). Frequency-wave number estimation by structured matrix approximation. *Proceeding of IEEE ASSP 3rd Workship on Spectral Estimation Modelling* (Boston, MA), November, 1986, pp. 81-84.

Smythe, S. T. (1973). Strong laws of large numbers for r-dimensional arrays of random variables. *Ann. Probab.* **1**, 164-170.

Wichura, M. J. (1973). Some strassen-type laws of the iterated logarithm for multiparameter stochastic processes with independent increments. *Ann. Probab.* **1**, 272-296.

Department of Statistics and Finance, University of Sciences and Technology, Hefei, Anhui, China.

E-mail: baiqi@math.ustc.edu.cn

E-mail: lczhao@math.ustc.edu.cn

Department of Mathematics and Statistics, York University, 4700 Keele Street, Toronto, Ontario, M3J 1P3 Canada.

E-mail: wuyh@mathstat.yorku.ca

(Received July 1996; accepted March 1997)