Proc. Natl. Acad. Sci. USA
Vol. 96, pp. 11106–11110, September 1999
Statistics

The simultaneous estimation of the number of signals and frequencies of multiple sinusoids when some observations are missing: I. Asymptotics

(strong consistency/limiting distribution/rate of convergence/signal processing/undamped exponential signal model)

ZHIDONG BAI*, CALYAMPUDI R. RAO†, YUEHUA WU‡, MEI-MEI ZEN§, AND LINCHENG ZHAO¶

*Department of Statistics and Applied Probability, The National University of Singapore, Singapore; †Department of Statistics, Pennsylvania State University, University Park, PA 16802; ‡Department of Mathematics and Statistics, York University, Toronto, Ontario, Canada M3J 1P3; §Department of Statistics, National Cheng-Kung University, Tainan, Taiwan; and ¶Department of Statistics and Finance, University of Science and Technology, Hefei, Anhui, China

Contributed by Calyampudi R. Rao, July 26, 1999

ABSTRACT The problem of simultaneous estimation of the number of signals and frequencies of multiple sinusoids is considered in the case when some observations are missing. The number of signals is estimated with an information theoretic criterion, and the frequencies are estimated with eigenvariation linear prediction. The strong consistency of the estimates of the number of signals and the frequencies is established and the rate of convergence of these estimates is provided. Besides, the limiting distributions of various estimates are given.

1. Introduction

In resource prospecting and earthquake detection, telecommunications, biomedical engineering, radio location of objects, etc., it is often necessary to detect the number of signals. A commonly used model for signal processing is the undamped exponential model

$$y(n) = \sum_{j=1}^{p_0} \alpha_j e^{i\omega_j n} + w(n), \quad n = 1, \ldots, N, \qquad [1.1]$$

where $i = \sqrt{-1}$, $\{\alpha_j\}$ is a set of unknown complex amplitudes, $\{\omega_j\}$ is a set of unknown angular frequencies, and $\{w(n)\}$ is a sequence of independently and identically distributed complex random noise variables, usually assumed to have mean zero and finite variance σ^2. Associated with this model, two interesting problems are to determine the number of signals and estimate the unknown parameters. Even when the number of signals p_0 is known, it is not easy to find the least squares estimates of α_j and ω_j values because it would involve solving a system of nonlinear equations with exponential functions. To avoid this difficulty, various methods have been developed in the literature. Among others, references may be made to Bressler and MaCovski (1), Kay (2), Kumaresan *et al.* (3), Kundu (4), Rao (5), Stoica (6), Tufts and Kumaresan (7), and Ulrych and Clayton (8).

Missing or incomplete data of model **1.1** are usually failure of sensors or recording. Because of this failure, the estimation of both the number of signals and the amplitudes and frequencies of the signals is of both theoretical and practical interest. Kundu and Kundu (9) proposed a consistent estimation of the frequency parameters. In this paper, we propose a method for simultaneously estimating the number of signals and the frequencies of the signals, and the asymptotic properties of the proposed method will be discussed.

In Section 2, a method for simultaneous estimation of the number of signals and the amplitudes and frequencies of the signals is proposed when some observations in model **1.1** are missing. In Section 3, the strong consistency of the estimators is proved. In Section 4, we establish the limiting distributions of the newly proposed estimators. In Section 5, the rates of convergence of the estimators of the number of signals and the frequencies of the signals are given.

In the following, $\bar{A}$ denotes the complex conjugate of the matrix A, A^* denotes the complex conjugate transpose of A and $\|A\|^2 = \text{trace}(A^*A)$.

2. Determination of the Number of Signals and Estimation of the Frequency Parameters When Some Observations Are Missing

Suppose that the data sequence $\{y(n)\}$ is given as follows:

$$y(n) = \sum_{j=1}^{p_0} \alpha_j e^{i\omega_j n} + w(n), \quad n = 1, \ldots, N, \qquad [2.1]$$

where $i = \sqrt{-1}$, $\{\alpha_j\}$ is a set of unknown complex amplitudes, $\{\omega_j\}$ is a set of unknown angular frequencies, and $\{w(n)\}$ is a sequence of independently and identically distributed complex random noise variables such that

$$E(w(1)) = 0, \quad Ew(1)\overline{w(1)} = \sigma^2, \quad E|w(1)|^4 < \infty, \qquad [2.2]$$

with σ^2 unknown. We assume that $\omega_j \neq \omega_k$ if $j \neq k$, $\omega_j \in [0, 2\pi)$ for any j, and $p_0 \leq P < \infty$.

In this paper, we are primarily interested in determining p_0 and estimating the frequency parameters ω_j, $j = 1, \ldots, p_0$. Once ω_j values are estimated, the α_j values can be found by the linear least squares fit to the data.

Denote

$$Y = \begin{bmatrix} y(p+1) & y(p) & \cdots & y(1) \\ y(p+2) & y(p+1) & \cdots & y(2) \\ \cdots & \cdots & \cdots & \cdots \\ y(N) & y(N-1) & \cdots & y(N-p) \end{bmatrix} = [\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_{N-p}]'. \qquad [2.3]$$

Suppose that the observations $\{y(k), k \in \kappa_N\}$ are missing, where κ_N is a subset of the set $\{1, \ldots, N\}$. Let $A_{N,p}$ be the matrix obtained from the matrix Y in **2.3** with the rows having missing observations removed. Denote $\theta_{N,p} = \{n: (y(n+p), \ldots, y(n))$ is a row of $A_{N,p}\}$, and let $r_{N,p}$ be the number of rows of $A_{N,p}$. For brevity of notation, we omit N in the above notations.

Define $\hat{\Gamma}^{(p)} = r_p^{-1} A_p^* A_p = (\hat{\gamma}_{m\ell}^{(p)})$. Let

$$S_p = \min\{r_p^{-1}\|A_p \boldsymbol{b}^{(p)}\|^2: \|\boldsymbol{b}^{(p)}\| = 1\},$$

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

PNAS is available online at www.pnas.org.

for $p = 0, 1, \ldots, P$, where $\boldsymbol{b}^{(p)} = (b_0^{(p)}, b_1^{(p)}, \ldots, b_p^{(p)})'$. It is clear that S_p is the smallest eigenvalue of $\hat{\Gamma}^{(p)}$.

Let

$$R_p = S_p + pC_N, \quad p = 0, 1, \ldots, P, \qquad \textbf{[2.4]}$$

where C_N is assumed to satisfy

$$\lim_{N\to\infty} C_N = 0, \quad \lim_{N\to\infty} \frac{\sqrt{r_p}C_N}{\sqrt{\log\log r_p}} = \infty. \qquad \textbf{[2.5]}$$

Then, we find a nonnegative integer $\hat{p} \le P$ such that $R_{\hat{p}} = \min_{0\le p\le P} R_p$. We use $\hat{p}$ as an estimate of p_0.

Further, we find a unit $(\hat{p}+1) \times 1$ complex vector $\hat{\boldsymbol{b}}^{(\hat{p})} = (b_0^{(\hat{p})}, b_1^{(\hat{p})}, \ldots, b_p^{(\hat{p})})'$ such that

$$S_{\hat{p}} = r_{\hat{p}}^{-1} \| A_{\hat{p}} \hat{\boldsymbol{b}}^{(\hat{p})} \|^2. \qquad \textbf{[2.6]}$$

Let $\hat{\rho}_j e^{-i\hat{\omega}_j}$, $j = 1, \ldots, \hat{p}$, be the solutions to the equation $\hat{B}(z) \hat{=} \sum_{j=0}^{\hat{p}} \hat{b}_j z^j = 0$, where $\hat{\rho}_j > 0$, $\hat{\omega}_j \in [0, 2\pi)$, $j = 1, \ldots, \hat{p}$. Then we use $\hat{\omega}_j$ values as estimates of ω_j values.

We need the following assumption:

ASSUMPTION (A). *Assume that* $r_p \to \infty$ *and* $\sum_{j\in\theta_p} e^{ija} = O(1)$ *for any real number* $a \neq 0$.

Remark 2.1: It is easy to show that r_p tends to ∞ and $\sum_{j\in\theta_p} e^{ija} = O(1)$ for any real number $a \neq 0$ when κ_N is bounded.

3. Strong Consistency of the Detection and Estimation Procedures

The following theorem contains the main results of this section.

THEOREM 3.1. *Suppose that* $\{w(n)\}$ *is an iid sequence of complex random variables satisfying* **2.2** *and that Assumption (A) holds. Then*

(i) $\hat{p} = p_0$ for large N;

(ii) there exists a unique $(p_0 + 1) \times 1$ unit vector $\hat{\boldsymbol{b}}$ (up to a complex factor with modulus one) which satisfies **2.6**, and

(iii) for appropriate ordering

$$\hat{\omega}_j \to \omega_j, \quad j = 1, \ldots, p_0, \quad S_{\hat{p}} \to \sigma^2, \quad \text{as} \quad N \to \infty.$$

The following two lemmas are needed in the proof of *Theorem 3.1*.

LEMMA 3.1. *Let* $\{x_n, \ n \ge 1\}$ *be a sequence of independent real random variables with zero means. Write* $s_n^2 = \sum_{j=1}^n E(x_j^2)$ *and* $S_n = \sum_{j=1}^n x_j$. *If* $\liminf_{n\to\infty} s_n^2/n > 0$ *and* $E(|x_j|^{2+\mu}) \le K < \infty$, $j \ge 1$ *for some constants* K *and* $\mu > 0$, *then* $\limsup_{n\to\infty} S_n/\sqrt{2s_n^2 \log\log s_n^2} = 1$, *a.s.*

Proof: For a proof, see Petrov (10). ■

Let

$$\Gamma^{(p)} = \sigma^2 I_{p+1} + \Omega D D^* \Omega^* = (\gamma_{m\ell}^{(p)}), \qquad \textbf{[3.1]}$$

where $D = \mathrm{diag}(\alpha_1, \ldots, \alpha_{p_0})$ and

$$\Omega = \begin{bmatrix} 1 & \cdots & 1 \\ e^{i\omega_1} & \cdots & e^{i\omega_{p_0}} \\ \cdots & \cdots & \cdots \\ e^{ip\omega_1} & \cdots & e^{ip\omega_{p_0}} \end{bmatrix}. \qquad \textbf{[3.2]}$$

LEMMA 3.2. *Suppose that* $\{w(n)\}$ *is an iid sequence satisfying* **2.2** *and that Assumption (A) holds. Then, as* $N \to \infty$,

$$\hat{\Gamma}^{(p)} = \Gamma^{(p)} + O\left(r_p^{-1/2}\sqrt{\log\log r_p}\right), \ a.s. \qquad \textbf{[3.3]}$$

Proof: For $0 \le p \le P$, and $m, \ell = 0, 1, \ldots, p$, $\hat{\gamma}_{m\ell}^{(p)} = r_p^{-1} \sum_{n\in\theta_p} \overline{y(n-m)} y(n-\ell) = \sum_{j,\,h=1,\,j\neq h}^{p_0} \alpha_j \bar{\alpha}_h e^{i(m\omega_h - \ell\omega_j)} r_p^{-1} \sum_{n\in\theta_p} e^{in(\omega_j-\omega_h)} + \sum_{j=1}^{p_0} \alpha_j e^{i(m-\ell)\omega_j} r_p^{-1} \times \sum_{n\in\theta_p} e^{i(n-m)\omega_j} \overline{w(n-m)} + \sum_{j=1}^{p_0} \bar{\alpha}_j e^{i(m-\ell)\omega_j} r_p^{-1} \sum_{n\in\theta_p} e^{-i(n-\ell)\omega_j} w(n-\ell) + r_p^{-1} \sum_{n\in\theta_p} w(n-\ell)\overline{w(n-m)} + \sum_{j=1}^{p_0} |\alpha_j|^2 e^{i(m-\ell)\omega_j} \hat{=} J_{1N} + J_{2N} + J_{3N} + J_{4N} + \nu_{m\ell}$. Because $r_p \to \infty$ as $N \to \infty$ and $\sum_{j\in\theta_p} e^{ija} = O(1)$ for any real number $a \neq 0$,

$$J_{1N} = O\left(r_p^{-1}\right). \qquad \textbf{[3.4]}$$

Using the condition **2.2** and *Lemma 3.1*, we obtain

$$J_{2N} = O\left(r_p^{-1/2}\sqrt{\log\log r_p}\right), \ a.s.$$

$$J_{3N} = O\left(r_p^{-1/2}\sqrt{\log\log r_p}\right), \ a.s. \qquad \textbf{[3.5]}$$

In terms of the law of iterated logarithm of M-dependent sequence, it follows that

$$J_{4N} = \begin{cases} O\left(\sqrt{\frac{\log\log r_p}{r_p}}\right), \ a.s. & \text{for } m \neq \ell, \\ \sigma^2 + O\left(\sqrt{\frac{\log\log r_p}{r_p}}\right), \ a.s. & \text{for } m = \ell. \end{cases} \qquad \textbf{[3.6]}$$

By **3.4–3.6**, **3.3** follows. ■

We begin to prove the *Theorem 3.1*.

Proof of Theorem 3.1: Because both $\Gamma^{(p)}$ and $\hat{\Gamma}^{(p)}$ are positive-definite Hermitian matrices, their eigenvalues and $\mathrm{trace}(\Gamma^{(p)}\hat{\Gamma}^{(p)})$ are nonnegative. Let $\hat{\lambda}_1^{(p)} \ge \cdots \ge \hat{\lambda}_{p+1}^{(p)}$ be the eigenvalues of $\hat{\Gamma}^{(p)}$, and let $\lambda_1^{(p)} \ge \cdots \ge \lambda_{p+1}^{(p)}$ be the eigenvalues of $\Gamma^{(p)}$. Then, by von Neumann (11), $\sum_{j=1}^{p+1} \lambda_j^{(p)} \hat{\lambda}_j^{(p)} \ge \mathrm{trace}(\Gamma^{(p)}\hat{\Gamma}^{(p)})$. Hence,

$$\sum_{j=1}^{p+1} (\hat{\lambda}_j^{(p)} - \lambda_j^{(p)})^2 \le \mathrm{trace}(\hat{\Gamma}^{(p)} - \Gamma^{(p)})^2. \qquad \textbf{[3.7]}$$

By **3.1**, it is easy to see that

$$\lambda_{p+1}^{(p)} > \sigma^2 \text{for } p < p_0, \text{ and } \lambda_{p+1}^{(p)} = \sigma^2 \text{for } p \ge p_0. \qquad \textbf{[3.8]}$$

Hence, by *Lemma 3.2*, **3.7**, and $S_p = \hat{\lambda}_{p+1}^{(p)}$, we have

$$\lim_{N\to\infty} S_p = \lambda_{p+1}^{(p)} > \sigma^2, \ a.s. \quad \text{for } p < p_0, \qquad \textbf{[3.9]}$$

and

$$|S_p - \sigma^2| = O\left(r_p^{-1/2}\sqrt{\log\log r_p}\right), \ a.s. \quad \text{for } p \ge p_0. \qquad \textbf{[3.10]}$$

Assume that $p < p_0$. Then by **3.9**, **3.10**, **2.4**, and **2.5**,

$$\lim_{N\to\infty} (R_{p_0} - R_p) = \sigma^2 - \lambda_{p+1}^{(p)} < 0, \ a.s. \qquad \textbf{[3.11]}$$

Hence, with probability one, for large N,

$$R_{p_0} < R_p \quad \text{for } p < p_0.$$

Now we assume that $p > p_0$. Then by **3.10**, **2.4**, and **2.5**, with probability one, for large N,

$$R_{p_0} - R_p = S_{p_0} - S_p - (p - p_0)C_N$$
$$= O\left(\sqrt{\frac{\log\log r_p}{r_p}}\right) - (p - p_0)C_N < 0. \qquad \textbf{[3.12]}$$

Therefore, with probability one, for large N, $\hat{p} = p_0$, which establishes (*i*).

To prove (*ii*) and (*iii*), we can use the similar methods found in the literature, e.g., Bai and Rao (12). The details are omitted.

4. Limiting Distributions of $S_{\hat{p}}$ and $\{\hat{\omega}_j, j \le \hat{p}\}$.

Because $\hat{p} \to p_0$, a.s., as $N \to \infty$, we use p_0 instead of $\hat{p}$ when we consider the limiting properties of various estimates involving $\hat{p}$. For further simplicity, we shall use p for p_0, but we keep all other notations defined in previous sections. In the following, Ω in **3.2** is a $(p+1) \times p$ matrix.

Throughout this section, we assume that $\{w(n)\}$ is a sequence of iid complex r.v.'s such that

$$E(w(1)) = 0, \quad Ew(1)\overline{w(1)} = \sigma^2,$$
$$\mathrm{Var}(|w(1)|^2) = c\sigma^4 \text{ with } c > 0. \qquad \textbf{[4.1]}$$

Let $v_j \sim N(0, \sigma^2)$, $j = 1, \ldots, p$, $u_0 \sim N(0, c\sigma^4)$, and $u_j \sim N(0, \sigma^4)$, $j = 1, \ldots, p$. Assume that v_j and u_j values are independent of each other. Denote

$$U = \begin{bmatrix} u_0 & u_1 & \cdots & u_p \\ \bar{u}_1 & u_0 & \cdots & u_{p-1} \\ \cdots & \cdots & \cdots & \cdots \\ \bar{u}_p & \bar{u}_{p-1} & \cdots & u_0 \end{bmatrix},$$

$V = \mathrm{diag}(v_1, \ldots, v_p)$, $\zeta_N = \sqrt{r_p}(S_p - \sigma^2)$, $\tau_{Nj} = \sqrt{r_p}(\hat{\rho}_j - 1)$, $\Delta_{Nj} = \sqrt{r_p}(\hat{\omega}_j - \omega_j)$, $\boldsymbol{\tau}_N = (\tau_{N1}, \ldots, \tau_{Np})'$, and $\boldsymbol{\Delta}_N = (\Delta_{N1}, \ldots, \Delta_{Np})'$. Define $B(z) = \sum_{j=0}^{p} b_j z^j$ where $\boldsymbol{b} = (b_0, b_1, \ldots, b_p)'$ is the eigenvector of $\Gamma^{(p)}$ corresponding to the eigenvalue σ^2 such that $\|\boldsymbol{b}\| = 1$. Write $G = \mathrm{diag}(i\frac{d}{d\,\omega_1}B(e^{-i\omega_1}), \ldots, i\frac{d}{d\,\omega_p}B(e^{-i\omega_p}))$. Then the limiting distributions of $\sqrt{r_p}(S_p - \sigma^2)$ and $\sqrt{r_p}((\hat{\rho}_1 - 1) - i(\hat{\omega}_1 - \omega_1), \ldots, (\hat{\rho}_p - 1) - i(\hat{\omega}_p - \omega_p))$ are given in the following theorem.

THEOREM 4.1. *Suppose that condition* **4.1** *is satisfied and that Assumption (A) holds. Then,*

$$\zeta_N \xrightarrow{\mathcal{L}} \zeta \hat{=} \boldsymbol{b}^* U \boldsymbol{b}, \qquad \textbf{[4.2]}$$

and

$$\boldsymbol{\tau}_N - i\boldsymbol{\Delta}_N \xrightarrow{\mathcal{L}} G^{-1}(DD^*)^{-1}(\Omega^*\Omega)^{-1}\Omega^* U \boldsymbol{b}. \qquad \textbf{[4.3]}$$

The following lemma is needed to prove *Theorem 4.1*.

LEMMA 4.1. *Suppose that condition* **4.1** *is satisfied and that Assumption (A) holds. Then* $r_p^{-1/2}\sum_{n\in\theta_p} e^{-i(n-\ell)\omega_j}\, w(n-\ell) \xrightarrow{\mathcal{L}} v_j$, *for* $j = 1, \ldots, p$, $\ell = 0, 1, \ldots, p$, $r_p^{-1/2}\sum_{n\in\theta_p}(|w(n-\ell)|^2 - \sigma^2) \xrightarrow{\mathcal{L}} u_0$, *for* $\ell = 0, 1, \ldots, p$, $r_p^{-1/2}\sum_{n\in\theta_p} w(n-\ell)\overline{w(n-m)} \xrightarrow{\mathcal{L}} u_{\ell-m}$, *for* $0 \le m < \ell \le p$.

Proof: Since $\sum_{j\in\theta_p} e^{ija} = O(1)$ for any real number $a \ne 0$, the normality of v_j, u_0, and u_j follows from the central limit theorem. ■

Proof of Theorem 4.1: Let $\hat{\Psi} = r_p^{-1/2}(\hat{\Gamma}^{(p)} - \Gamma^{(p)}) = (\hat{\psi}_{m\ell})$. By *Lemma 4.1*, we have $\hat{\Psi} \xrightarrow{\mathcal{L}} \Psi \hat{=} \Omega(D\bar{V} + \bar{D}V)\Omega^* + U$. Because S_p is the smallest eigenvalue of $\hat{\Gamma}^{(p)}$, we have

$$0 = \det[\hat{\Gamma}^{(p)} - S_p I_{p+1}]$$
$$= \det[\Omega DD^*\Omega^* + r_p^{-1/2}\hat{\Psi} - (S_p - \sigma^2)I_{p+1}]. \qquad \textbf{[4.4]}$$

Let Q be a unitary matrix such that $Q^*\Omega DD^*\Omega^* Q = \mathrm{diag}(\xi_1, \ldots, \xi_p, 0)$, $\xi_1 \ge \ldots \ge \xi_p > 0$. Note that the last column of Q is the eigenvector of $\Gamma^{(p)}$ corresponding to the eigenvalue σ^2. We choose this column as $\boldsymbol{b}$. Write $\Xi_N = \mathrm{diag}(\xi_1 - (S_p - \sigma^2), \ldots, \xi_p - (S_p - \sigma^2), -(S_p - \sigma^2)) + r_p^{-1/2}Q^*\hat{\Psi}Q$. By **4.4**, $\det[\Xi_N] = 0$. Multiply by $\sqrt{r_p}$ the last row of the matrix Ξ_N. Because $\hat{\Psi} \xrightarrow{\mathcal{L}} \Psi$, by Skorohod's representation theorem [see Skorohod (13)], $\hat{\Psi} \to \Psi$, a.s. as $N \to \infty$. Therefore,

$$\zeta_N \to \zeta = \boldsymbol{b}^*\Psi\boldsymbol{b} = \boldsymbol{b}^* U \boldsymbol{b}, \quad a.s. \qquad \textbf{[4.5]}$$

Note that **4.5** reveals only that there exist some versions of ζ_N, ζ, Ψ, and U which have the same distributions as ζ_N, ζ, Ψ, and U, respectively, such that **4.5** holds. Hence we only get **4.2**. The principle of this statement also applies to the following proof of **4.3** and so on.

Because $(\hat{\Gamma}^{(p)} - S_p I_{p+1})\hat{\boldsymbol{b}} = \boldsymbol{0}$ and $(\Gamma^{(p)} - \sigma^2 I_{p+1})\boldsymbol{b} = \boldsymbol{0}$ with the choice of $\hat{\boldsymbol{b}}$ such that $\hat{\boldsymbol{b}} \to \boldsymbol{b}$, *a.s.* and $\hat{\Psi} \to \Psi$, *a.s.*, by **4.5**, we have

$$\boldsymbol{0} = (\Gamma^{(p)} - \sigma^2 I_{p+1})\boldsymbol{\eta}_N + (I_{p+1} - \boldsymbol{b}\boldsymbol{b}^*)U\boldsymbol{b}$$
$$+ o(1), \quad a.s. \qquad \textbf{[4.6]}$$

where $\boldsymbol{\eta}_N = \sqrt{r_p}(\hat{\boldsymbol{b}} - \boldsymbol{b}) \hat{=} (\eta_{N0}, \ldots, \eta_{Np})'$.

Write $\boldsymbol{\eta}_N = \boldsymbol{\eta}_{N1} + \boldsymbol{\eta}_{N2}$ such that $\boldsymbol{\eta}_{N1} = \Omega\boldsymbol{\beta}_N$ for some complex random vector $\boldsymbol{\beta}_N$ and $\Omega^*\boldsymbol{\eta}_{N2} = \boldsymbol{0}$. Because $(\Gamma^{(p)} - \sigma^2 I_{p+1})\boldsymbol{\eta}_N = \Omega DD^*\Omega^*\boldsymbol{\eta}_N = \Omega DD^*\Omega^*\boldsymbol{\eta}_{N1} = \Omega DD^*\Omega^*\Omega\boldsymbol{\beta}_N$, by **4.6**, we obtain

$$\boldsymbol{\beta}_N \to -(\Omega^*\Omega)^{-1}(DD^*)^{-1}(\Omega^*\Omega)^{-1}\Omega^* U\boldsymbol{b}, \ a.s.$$

which implies that $\boldsymbol{\eta}_{N1} \to -\Omega(\Omega^*\Omega)^{-1}(DD^*)^{-1}(\Omega^*\Omega)^{-1} \cdot \Omega^* U\boldsymbol{b}$, a.s. Hence

$$\Omega^*\boldsymbol{\eta}_N = \Omega^*\boldsymbol{\eta}_{N1} \to -(DD^*)^{-1}(\Omega^*\Omega)^{-1}\Omega^* U\boldsymbol{b}, \quad a.s. \qquad \textbf{[4.7]}$$

Because $\hat{\boldsymbol{b}} \to \boldsymbol{b}$, *a.s.*, $\hat{\rho}_j \to 1$, *a.s.*, and $\hat{\omega}_j \to \omega_j$, *a.s.*, for appropriate ordering and for $j = 1, \ldots, p$, with probability one, we have

$$\Omega^*\boldsymbol{\eta}_N + G(\boldsymbol{\tau}_N - i\boldsymbol{\Delta}_N) + o(\|\boldsymbol{\tau}_N\| + \|\boldsymbol{\Delta}_N\|) = \boldsymbol{0} \qquad \textbf{[4.8]}$$

for large N. Because $e^{-i\omega_j}$ is a simple root of the multinomial $B(z)$, we have $\frac{d}{d\omega_j}B(e^{-i\omega_j}) \ne 0$ for $j = 1, \ldots, p$, and thus G is nonsingular, which, together with **4.7** and **4.8**, implies that

$$\boldsymbol{\tau}_N - i\boldsymbol{\Delta}_N \to tG^{-1}(DD^*)^{-1}(\Omega^*\Omega)^{-1}\Omega^* U\boldsymbol{b}, \quad a.s. \qquad \textbf{[4.9]}$$

Hence, **4.3** follows. ■

5. Rate of Convergence of the Estimates of the Number of Signals and the Frequencies

The main results of this section are included in the following four theorems. The first theorem gives the rates of convergence of $\hat{p}$.

THEOREM 5.1. *Suppose that* $\{w(n)\}$ *is an iid sequence of complex random variables with zero mean and finite variance and that Assumption (A) holds. If* C_N *in* **2.4** *satisfies the following conditions*

$$\lim_{N\to\infty} C_N = 0, \quad \lim_{N\to\infty} \sqrt{r_p}C_N = \infty, \qquad \textbf{[5.1]}$$

then $E(|w(1)|^{2\mu}) < \infty$, $\mu > 1$, *implies that*

$$P(\hat{p} \ne p_0) = O(r_p(r_p C_n)^{-\mu}) + O((r_p C_N^2)^{-s}), \qquad \textbf{[5.2]}$$

as $N \to \infty$, *for any* $s > \mu$, *and* $E(e^{\tau|w(1)|^2}) < \infty$, $\tau > 0$ *implies that*

$$P(\hat{p} \ne p_0) = O(e^{-\eta r_p C_N^2}), \qquad \textbf{[5.3]}$$

as $N \to \infty$, *for some* $\eta > 0$.

Let

$$\hat{\boldsymbol{b}}^{(p_0)} = a_N[I - (\hat{\Gamma}^{(p_0)} - S_{p_0}I)^+(\hat{\Gamma}^{(p_0)} - S_{p_0}I)]\boldsymbol{b}^{(p_0)}, \qquad \textbf{[5.4]}$$

where $a_N > 0$ such that $\|\hat{\boldsymbol{b}}^{(p_0)}\| = 1$, and G^+ denotes Moore–Penrose inverse of the matrix G. By **3.3** and **3.10**, $\hat{\boldsymbol{b}}^{(p_0)}$ is well defined and $a_N = \|[I - (\hat{\Gamma}^{(p_0)} - S_{p_0}I)^+(\hat{\Gamma}^{(p_0)} - S_{p_0}I)]\boldsymbol{b}^{(p_0)}|^{-1}$. The following theorem gives the rates of convergence of $\hat{\boldsymbol{b}}^{(p_0)}$.

THEOREM 5.2. *Suppose that $\{w(n)\}$ is an iid sequence of complex random variables with zero mean and finite varianc e and that Assumption (A) holds. If C_N satisfies* **2.5**, *then $E(|w(1)|^{2\mu}) < \infty$, $\mu > 1$, implies that $P(|\hat{\boldsymbol{b}}^{(p_0)} - \boldsymbol{b}^{(p_0)}| \geq \varepsilon) = O\left(r_p(r_pC_n)^{-\mu}\right) + O\left((r_pC_N^2)^{-s}\right)$, as $N \to \infty$, for some $s > \mu$, and $E(e^{\tau|w(1)|^2}) < \infty$, $\tau > 0$ implies that $P(|\hat{\boldsymbol{b}}^{(p_0)} - \boldsymbol{b}^{(p_0)}| \geq \varepsilon) = O\left(e^{-\eta r_p C_N^2}\right)$, as $N \to \infty$, for some $\eta > 0$.*

Because the roots of a polynomial are continuous functions of coefficients of the polynomial, we have the following theorem, which describes the rates of convergence of $(\hat{\rho}e^{-i\hat{\omega}_1}, \ldots, \hat{\rho}e^{-i\hat{\omega}_{\hat{p}}})$.

THEOREM 5.3. *Suppose that $\{w(n)\}$ is an iid sequence of complex random variables with zero mean and finite varianc e and that Assumption (A) holds. Assume that $\hat{\omega}_j$ and ω_j $j = 1, \ldots, \hat{p}$, are both arranged in increasing order. Let $\hat{z} = (\hat{\rho}e^{-i\hat{\omega}_1}, \ldots, \hat{\rho}e^{-i\hat{\omega}_{\hat{p}}})'$ and $z = (e^{-i\omega_1}, \ldots, e^{-i\omega_{p_0}})'$. If C_N satisfies* **2.5**, *then $E(|w(1)|^{2\mu}) < \infty$, $\mu > 1$, implies that $P(|\hat{z} - z| \geq \varepsilon) = O\left(r_p(r_pC_n)^{-\mu}\right) + O\left((r_pC_N^2)^{-s}\right)$, as $N \to \infty$, for some $s > \mu$, and $E(e^{\tau|w(1)|^2}) < \infty$, $\tau > 0$ implies that $P(|\hat{z} - z| \geq \varepsilon) = O\left(e^{-\eta r_p C_N^2}\right)$, as $N \to \infty$, for some $\eta > 0$.*

Hence it is obvious that the following results regarding the rates of convergence of $\{\hat{\rho}_j\}$ and $\{\hat{\omega}_j\}$ hold.

THEOREM 5.4. *Suppose that $\{w(n)\}$ is an iid sequence of complex random variables with zero mean and finite variance. Assume that r_p, the number of rows of A_p, tends to ∞ and $\sum_{j\in\theta_p} e^{ija} = O(1)$ for any real number $a \neq 0$. Assume that $\hat{\omega}_j$ and ω_j $j = 1, \ldots, \hat{p}$, are both arranged in increasing order. Let $\hat{z} = (\hat{\rho}e^{-i\hat{\omega}_1}, \ldots, \hat{\rho}e^{-i\hat{\omega}_{\hat{p}}})'$ and $z = (e^{-i\omega_1}, \ldots, e^{-i\omega_{p_0}})'$. If C_N satisfies* **2.5**, *then $E(|w(1)|^{2\mu}) < \infty$, $\mu > 1$, implies that for some $s > \mu$, $P(\max_{1\leq j\leq \hat{p}} |\hat{\rho}_j - 1| \geq \varepsilon) = O\left(r_p(r_pC_n)^{-\mu}\right) + O\left((r_pC_N^2)^{-s}\right)$, and $P(\max_{1\leq j\leq \hat{p}\vee p_0} |\hat{\omega}_j - \omega_j| \geq \varepsilon) = O\left(r_p(r_pC_n)^{-\mu}\right) + O\left((r_pC_N^2)^{-s}\right)$, as $N \to \infty$, and $E(e^{\tau|w(1)|^2}) < \infty$, $\tau > 0$ implies that for some $\eta > 0$, $P(\max_{1\leq j\leq \hat{p}} |\hat{\rho}_j - 1| \geq \varepsilon) = O\left(e^{-\eta r_p C_N^2}\right)$, and $P(\max_{1\leq j\leq \hat{p}\vee p_0} |\hat{\omega}_j - \omega_j| \geq \varepsilon) = O\left(e^{-\eta r_p C_N^2}\right)$, as $N \to \infty$.*

The following two lemmas are needed in the proofs of the theorems.

LEMMA 5.1. *Let $z_1, z_2, \ldots$ be independent and identically distributed complex random variables with mean 0 and $E|z_1|^\mu < \infty$ for some $\mu > 1$ and $a_1, a_2, \ldots$ be a sequence of real numbers. Assume that $n(n\alpha_n)^{-\mu} \to 0$ and $n\alpha_n^2 \to \infty$. Then for any $s > \mu$ and, we have*

$$P\left(\left|\sum_{j=1}^{n} e^{ia_j} z_j\right| \geq n\alpha_n\right) = O\left(n(n\alpha_n)^{-\mu}\right) + O\left((n\alpha_n^2)^{-s}\right).$$

Proof: The lemma can be proved following the same lines of the proof of lemma 2.2 of Bai, Krishnaiah, and Zhao (14). ■

LEMMA 5.2. *Let $z_1, z_2, \ldots$ be independent and identically distributed complex random variables with mean 0 and $Ee^{\tau|z_1|} < \infty$ for some $\tau > 0$ and $a_1, a_2, \ldots$ be a sequence of real numbers. Assume that $\alpha_n \to 0$ and $n\alpha_n^2 \to \infty$. Then*

$$P\left(\left|\sum_{j=1}^{n} e^{ia_j} z_j\right| \geq n\alpha_n\right) \leq ce^{-\eta n\alpha_n^2}$$

for some $\eta > 0$ and $c > 0$.

Proof: The lemma can be proved in the same way as in lemma 2.3 of Bai, Krishnaiah, and Zhao (14).

Proof of Theorem 5.1: In view of the structure of $\Gamma^{(p)}$, it is clear that for $p \geq 1$,

$$\lambda_1^{(p)} \geq \lambda_1^{(p-1)} \geq \lambda_2^{(p)} \geq \lambda_2^{(p-1)} \geq \ldots$$
$$\geq \lambda_p^{(p)} \geq \lambda_p^{(p-1)} \geq \lambda_{p+1}^{(p)} \geq \sigma^2$$

and for $p \geq p_0$,

$$\lambda_{p_0+1}^{(p)} = \lambda_{p_0+2}^{(p)} = \ldots = \lambda_{p+1}^{(p)} = \sigma^2, \quad [5.5]$$
$$\lambda_{p_0}^{(p)} \geq \lambda_{p_0}^{(p_0-1)} > \sigma^2.$$

Now, let $\nu = \lambda_{p_0}^{(p_0-1)} - \sigma^2$. Because $\hat{\lambda}_{p+1}^{(p)} = S_p$, by **3.7** we have

$$|S_p - \lambda_{p+1}^{(p)}| \leq \sqrt{\mathrm{trace}(\hat{\Gamma}^{(p)} - \Gamma^{(p)})^2}, \quad p = 0, 1, \ldots, P. \quad [5.6]$$

If $p < p_0$, then by **5.5**,

$$\mathrm{trace}(\hat{\Gamma}^{(p)} - \Gamma^{(p)})^2 < C_N^2/4,$$

and

$$\mathrm{trace}(\hat{\Gamma}^{(p_0)} - \Gamma^{(p_0)})^2 < C_N^2/4, \quad [5.7]$$

which imply that $R_p - R_{p_0} = S_p - S_{p_0} + (p - p_0))C_N \geq C_N - \left(\sqrt{\mathrm{trace}(\hat{\Gamma}^{(p)} - \Gamma^{(p)})^2} + \sqrt{\mathrm{trace}(\hat{\Gamma}^{(p_0)} - \Gamma^{(p_0)})^2}\right) > 0$. Therefore,

$$\hat{p} \neq p, \quad \text{for } p > p_0. \quad [5.8]$$

If $p < p_0$, then by **5.1**, $\mathrm{trace}(\hat{\Gamma}^{(p)} - \Gamma^{(p)})^2 < C_N^2/4$, which, together with **5.7**, implies that $R_p - R_{p_0} = S_p - S_{p_0} - (p_0 - p))C_N \geq \lambda_{p+1}^{(p)} - \sigma^2 - \left(\sqrt{\mathrm{trace}(\hat{\Gamma}^{(p)} - \Gamma^{(p)})^2} + \sqrt{\mathrm{trace}(\hat{\Gamma}^{(p_0)} - \Gamma^{(p_0)})^2} + (p_0 - p)C_N\right) \geq \nu - (p_0 + 1)C_N > 0$, for large N. Therefore,

$$\hat{p} \neq p, \quad \text{for } p < p_0. \quad [5.9]$$

In view of **5.8** and **5.9**, we have

$$P(\hat{p} \neq p_0) \leq P\left(\bigcup_{p=0}^{P}[\mathrm{trace}\left(\hat{\Gamma}^{(p)} - \Gamma^{(p)}\right)^2 \geq C_N^2/4]\right).$$

By the definition of $\hat{\gamma}_{m\ell}^{(p)}$, for $0 \leq p \leq P$, and $m, \ell = 0, 1, \ldots, p$,

$$\begin{aligned}\hat{\gamma}_{m\ell}^{(p)} &= r_p^{-1} \sum_{n\in\theta_p} \overline{y(n-m)}y(n-\ell) \\ &= \gamma_{m\ell}^{(p)} \quad [5.10] \\ &\quad + \sum_{j,h=1,j\neq h}^{p_0} \alpha_j\bar{\alpha}_h e^{i(m\omega_h - \ell\omega_j)} r_p^{-1} \sum_{n\in\theta_p} e^{in(\omega_j-\omega_h)} \\ &\quad + \sum_{j=1}^{p_0} \alpha_j e^{i(m-\ell)\omega_j} r_p^{-1} \sum_{n\in\theta_p} e^{i(n-m)\omega_j} \overline{w(n-m)} \\ &\quad + \sum_{j=1}^{p_0} \bar{\alpha}_j e^{i(m-\ell)\omega_j} r_p^{-1} \sum_{n\in\theta_p} e^{-i(n-\ell)\omega_j} w(n-\ell) \\ &\quad + r_p^{-1} \sum_{n\in\theta_p} [w(n-\ell)\overline{w(n-m)} - \sigma^2\delta_{m\ell}], \quad [5.11]\end{aligned}$$

where $\delta_{m\ell}$ denotes the Kronecker delta. If $E(|w(1)|^{2\mu}) < \infty$, $\mu > 1$, by **5.11** and *Lemma 5.1*, **5.2** follows. If $E(e^{\tau|w(1)|^2}) < \infty$, $\tau > 0$, by **5.11** and *Lemma 5.2*, **5.3** holds. The proof of *Theorem 5.1* is complete. ■

Proof of Theorem 5.2: The proof of *Theorem 5.2* can be given based on the same idea of the proof of *Theorem 5.1*. The details are omitted.

Y.W. was partially supported by the Natural Sciences and Engineering Research Council of Canada, and L.Z. was partially supported by National Natural Science Foundation of China, Ph.D. Program Foundation of National Education Committee of China, and Special Foundation of Academia Sinica. C.R.R. was supported by Army Research Office Grant DAAH04-96-1-0082.

1. Bressler, Y. & MaCovski, A. (1986) *IEEE Trans. Acoust. Speech Signal Process.* **34**, 1081–1089.
2. Kay, S. M. (1984) *IEEE Trans. Acoust. Speech Signal Process.* **32**, 540–547.
3. Kumaresan, R., Sharf, L. L. & Shaw, A. K. (1986) *IEEE Trans. Acoust. Speech Signal Process.* **34**, 833–840.
4. Kundu, D. (1993) *Technometrics* **35**, 215–218.
5. Rao, C. R. (1988) *Statistical Decision Theory and Related Topics, IV*, eds. Gupta, S. S. & Berger, J. O. (Springer, New York), Vol. 2, pp. 319–332.
6. Stoica, P. (1993) *Signal Process.* **37**, 720–741.
7. Tufts, D. W. & Kumaresan, R. (1982) *Proc. IEEE Special Issue Spectral Estimation* **70**, 975–989.
8. Ulrych, T. J. & Clayton, R. W. (1976) *Physics Earth Planet. Inter.* **12**, 188–200.
9. Kundu, D. & Kundu, R. (1995) *J. Statist. Planning Inference* **44**, 205–218.
10. Petrov, V. V. (1975) *Sums of Independent Random Variables* (Springer, New York).
11. von Neumann, J. (1937) *Tomsk University Rev.* **1**, 286–300.
12. Bai, Z. D. & Rao, C. R. (1989) *Spectral Analysis in One or Two Dimensions*, eds. Prasad, S. & Kashyap, R. L. (Oxford & IBH Publications Co., New Delhi), pp. 493–507.
13. Skorokhod, A. V. (1956) *Theo. Probab. App.* **1**, 261–290.
14. Bai, Z. D., Krishnaiah, P. R. & Zhao, L. C. (1989) *IEEE Trans. Inf. Theory* **35**, 380–388.