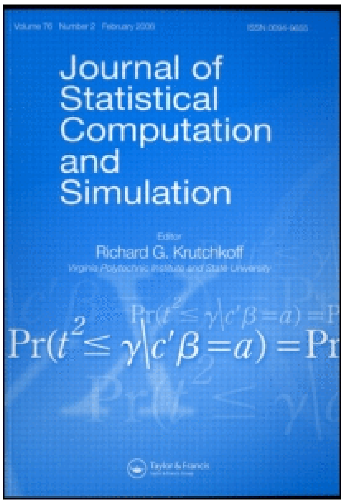

Journal of Statistical Computation and Simulation
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713650378

Consistent method for estimating sinusoidal frequencies: a non-iterative approach

Amit Mitra [a]; Debasis Kundu [a]
[a] Department of Mathematics, Indian Institute of Technology, Kanpur

To cite this Article Mitra, Amit and Kundu, Debasis'Consistent method for estimating sinusoidal frequencies: a non-iterative approach', Journal of Statistical Computation and Simulation, 58: 2, 171 — 194
To link to this Article: DOI: 10.1080/00949659708811829
URL: http://dx.doi.org/10.1080/00949659708811829

J. Statist. Comput. Simul., 1997, Vol. 58, pp. 171–194
Reprints available directly from the publisher
Photocopying permitted by license only

CONSISTENT METHOD FOR ESTIMATING SINUSOIDAL FREQUENCIES: A NON-ITERATIVE APPROACH

AMIT MITRA and DEBASIS KUNDU*

Department of Mathematics, Indian Institute of Technology, Kanpur-208016

(Received 1 May 1995; In final form 29 January 1997)

In this paper, we consider the problem of estimating the sinusoidal frequencies by a non-iterative technique. We establish the strong consistency of the proposed estimate. We further propose a modification of the non-iterative technique. It is observed in the simulation study that the proposed method works very well for reasonably small sample sizes. The mean squared errors of the proposed method reaches the Cramer-Rao lower bound in many situations. We also propose three different confidence intervals and compare their performances by simulation.

Keywords: Sinusoidal frequency; bootstrap confidence intervals; TLS-ESPRIT

AMS Subject Classification: 62J02, 62E25

1. INTRODUCTION

We consider the following time series model;

$$y_t = \sum_{k=1}^{M} (A_k \cos(\omega_K t) + B_K \sin(\omega_K t)) + \varepsilon_t \tag{1.1}$$

Here y_t's are observed at equidistant time points, for $t = 1, \ldots, N$. $(\omega_1, \ldots, \omega_M)$, $(A_1, \ldots, A_M)$ and $(B_1, \ldots, B_M)$ are the unknown parameters, ω_K's are distinct real numbers lying in $(0, \pi)$. M, the number of signals is assumed to be known apriori. $\{\varepsilon_t\}$ is a sequence of real

*Corresponding author.

valued i.i.d. random variables with

$$E(\varepsilon_t) = 0 \quad \text{and} \quad V(\varepsilon_t) = \sigma^2 \tag{1.2}$$

The problem is to estimate the unknown parameters ω_k, A_k and B_k for $k = 1, \ldots, M$ and σ^2. The estimation of the parameters of the model (1.1) is a fundamental problem in signal processing (Kay and Marple; 1981) and time series analysis. The asymptotic theory of the least squares estimates (LSE) for this model has a long history. Whittle (1953) obtained some of the earliest results. More recent results are by Hasan (1982), Hannan (1973) and Walker (1971). They formalized and extended Whittle's results. Walker (1971) introduced the concept of an approximate LSE for the model (1.1). He first estimated the frequencies by finding the maximum of the periodogram and then computing the estimates of the amplitudes. The approximate LSE were shown to be strongly consistent and the asymptotic normality of the estimates were also obtained. It may be noted that although asymptotically the approximate LSE estimates are equivalent to the exact LSE, for finite sample sizes the performance of the exact LSE are better than the approximate ones in terms of lower mean squared errors (Kundu and Mitra; 1996). Kundu (1993a) was the first one to give a direct proof of consistency of the exact LSE for the model (1.1) under the assumption of normality of the error random variables, the consistency and asymptotic normality for general error random variables can be found in (Kundu and Mitra, 1996).

It may be noted that although the least squares estimates are the most desired estimates, the problem of finding the estimates is well known to be numerically difficult. Rice and Rosenblatt (1988) discussed the computational complexities involved to obtain the LSE. The model (1.1) being a nonlinear one, to obtain the LSE some sort of iterative search procedure must be employed. Typically, search methods start from an initial guess value and then proceed by a sequence of Gauss-Newton steps. For this nonlinear least squares problem it turns out that there are many local minima with a separation in frequency of about N^{-1} which makes the stationary point to which the iterative scheme converges extremely sensitive to the starting values. This problem gets worse as the sample size increases. It is also observed (Rice and Rosenblatt; 1988) that unless the frequency is resolved at the first step with order $o(N^{-1})$, the failure to converge to the global minimum may give a very poor estimate of the amplitude. The problem becomes especially severe if one is estimating the parameters of several harmonic components simultaneously, since in that situation the iteration is taking place in a higher dimensional space with many local minima. The method of Walker (1971) for estimating the initial values by finding the maximum of the periodogram turns out to have drawbacks. A bias can arise for moderate sample sizes that is appreciable compared to the standard deviation suggested by asymptotic theory (Rice and Rosenblatt; 1988). The initial values provided to the search algorithms are thus critical. A direct search of the periodogram at a fine grid of points substantially finer than that given by the frequencies $2\pi i/N$ used by fast Fourier transform is appealing, but unfortunately has its drawbacks as well. Thus the problem to estimate efficiently the initial values to be provided to the search methods remains.

Recently the total least squares approach becomes quite popular to estimate the sinusoidal frequencies, references may be made to the works of Hua and Sarkar (1990) and if the errors are correlated then the generalized least squares approach can be used, see for example Mackisack and Poskitt (1989) and Dragosevic and Stankovic (1989). Another interesting problem is to estimate the number of sinusoidal components, i.e., 'M', of (1.1). It is a very important but difficult problem, see for example Ensor and Newton (1988), Fuchs (1988), Reddy and Biradar (1993), Hannan (1993), Kavalieris and Hannan (1994) and Kundu (1996).

In this article, we propose a new non-iterative method for estimating the frequencies of the model (1.1) which can be used as an efficient initial values. First we transform the model (1.1) to an undamped superimposed exponential signals model and then use extended order modeling and singular value decomposition technique to estimate the frequencies. We call the new estimate as **Noise Space Decomposition (NSD)** estimates. The linear parameters can then be obtained using separable regression technique of Richards (1961). The proposed method is shown to give strongly consistent estimates. Since the proposed method is strongly consistent, a further one step modified estimate is also proposed which already have the same asymptotic properties as the exact LSE. Some confidence intervals of the frequencies are also proposed.

The organization of the paper is as follows; in Section 2 we describe the estimation procedure and the consistency results are stated. Modified estimators are proposed in Section 3, different confidence intervals are discussed in Section 4. Some Monte Carlo simulations study is presented in Section 5 and finally we draw conclusions in Section 6. We provide the proof of consistency in the Appendix.

2. ESTIMATION PROCEDURE

Observe that the model (1.1) can be written as a linear combination of $2M$ complex exponential terms in the following way;

$$y_t = \sum_{k=1}^{M} C_r \exp(i\omega_r t) + \sum_{k=1}^{M} D_r \exp(-i\omega_r t) + \varepsilon_t, \quad t = 1, \ldots, N \quad (2.1)$$

where $C_r = (1/2)(A_r - iB_r)$ and $D_r = (1/2)(A_r + iB_r)$ and $i = \sqrt{-1}$.

It is well known (Prony; 1795, see Kundu; 1993b also) that in the noiseless case there exists an unique vector $\boldsymbol{C} = (c_1, \ldots, c_{2M+1})$ such that;

$$\sum_{k=1}^{2M+1} c_k y_{t+k} = 0 \quad (2.2)$$
$$\text{for all} \quad t = 0, \ldots, N - 2M - 1, \ \|\boldsymbol{C}\| = 1 \quad \text{and} \quad c_1 > 0.$$

The unknown constants $(c_1, \ldots, c_{2M+1})$ are such that the roots of the polynomial equation

$$c_1 Z^{2M} + c_2 Z^{2M-1} + \cdots + c_{2M+1} = 0 \quad (2.3)$$

are of the form $\lambda_k = \exp(\pm i\omega_K)$. Thus if we can estimate $\boldsymbol{C}$, we can estimate the unknown frequencies ω_k's using (2.3). Now observe that (Kahn *et. al.*, 1993), the condition that the roots be purely imaginary means (2.3) must factorize in the form

$$\eta_o \prod_{k=1}^{M} (\lambda^2 - (2 - \eta_K^2)\lambda + 1) \quad (2.4)$$

This implies that $c_k = c_{2M+2-K}$; $k = 1, \ldots, 2M + 1$ and the roots are $\lambda_K = \exp(\pm i\omega_K)$, where $2 - \eta_K^2 = 2\cos(\omega_K)$.

Consider the following $N-L \times L+1$ data matrix

$$\boldsymbol{A} = \begin{bmatrix} y_1 & \cdots & y_{L+1} \\ \vdots & \cdots & \vdots \\ y_{N-L} & \cdots & y_N \end{bmatrix} \tag{2.5}$$

for any positive integer L such that $2M \leq L \leq N-2M$.

Let's denote by $\boldsymbol{T}$ the $L+1 \times L+1$ matrix given by

$$\boldsymbol{T} = \frac{1}{N} A^* A \tag{2.6}$$

where '$*$' denotes the conjugate transpose of a matrix or a vector. Observe that in the noiseless case the matrix $\boldsymbol{T}$ has rank $2M$. Let the singular value decomposition of $\boldsymbol{T}$ (see Rao; pp. 42, 1973) be as follows;

$$\boldsymbol{T} = \sum_{i=1}^{L+1} \hat{\sigma}_i^2 \hat{\boldsymbol{U}}_i \hat{\boldsymbol{U}}_i^* \tag{2.7}$$

where $\hat{\sigma}_1^2 > \hat{\sigma}_z^2 > \ldots > \hat{\sigma}_{L+1}^2$ are the ordered eigenvalues of $\boldsymbol{T}$ and $\hat{\boldsymbol{U}}_i$ is the normalized eigenvector corresponding to $\hat{\sigma}_i^2$. The subspace generated by $\{\hat{\boldsymbol{U}}_1, \ldots, \hat{\boldsymbol{U}}_{2M}\}$ is denoted by $\mathbb{S}$ and that of $\{\hat{\boldsymbol{U}}_{2M+1}, \ldots, \hat{\boldsymbol{U}}_{L+1}\}$ is denoted by $\mathbb{N}$. We call $\mathbb{S}$ the signal subspace and $\mathbb{N}$ the noise subspace. Let $\boldsymbol{B}_1$ be any basis of the noise subspace $\mathbb{N}$. We write

$$\boldsymbol{B}_1 = \begin{bmatrix} b_{1,1} & \cdots & b_{1,\, L+1-2M} \\ b_{L+1,1} & \cdots & b_{L+1,L+1-2M} \end{bmatrix} \tag{2.8}$$

Observe that because of (2.2), in the noiseless situation there exists an unique basis of $\mathbb{N}$ which has the following form

$$\boldsymbol{B}_2 = \begin{bmatrix} c_1 & 0 & \cdots & 0 \\ \cdot & c_1 & \cdots & 0 \\ c_M & \cdot & & 0 \\ c_{M+1} & c_M & & 0 \\ c_M & c_{M+1} & & c_1 \\ \cdot & c_M & & \\ c_1 & \cdot & & \cdot \\ & & & c_M \\ 0 & c_1 & & c_{M+1} \\ \cdot & 0 & & c_M \\ 0 & 0 & \cdots & c_1 \end{bmatrix} \tag{2.9}$$

Now observe that $\boldsymbol{B}_1 = [\hat{\boldsymbol{U}}_{2M+1}, \ldots, \hat{\boldsymbol{U}}_{L+1}]$ forms a basis of the estimated noise space. Our main aim is to obtain a basis of $\mathbb{N}$ which has the form similar to (2.9) and to estimate $\boldsymbol{C}$ from these.

Let's partition the matrix $\boldsymbol{B}_1$ as follows;

$$\underset{L+1-2M\,\times\,L+1}{\boldsymbol{B}_1^T} = \left[\underset{L+1-2M\times K}{\boldsymbol{B}_{11}^T} \;:\; \underset{L+1-2M\times 2M+1}{\boldsymbol{B}_{12}^T} \;:\; \underset{L+1-2M\times L-K-2M}{\boldsymbol{B}_{13}^T}\right] \tag{2.10}$$

for $k = 0, 1, \ldots, L-2M$. Now consider the matrix

$$\left[\underset{L+1-2M\times K}{\boldsymbol{B}_{11}^T} \;:\; \underset{L+1-2M\times L-K-2M}{\boldsymbol{B}_{13}^T}\right].$$

Since the above is a random matrix, it is of rank $L-2M$. Therefore we can conclude that there exists an unique $L+1-2M$ vector $\boldsymbol{X_{k+1}} \neq 0$, such that

$$\begin{bmatrix} \boldsymbol{B}_{11}^T \\ K \times L+1-2M \\ \boldsymbol{B}_{13}^T \\ L-K-2M \times L+1-2M \end{bmatrix} \boldsymbol{X}_{K+1} = \boldsymbol{0} \tag{2.11}$$

Consider the $2M+1$ vector $\widehat{\boldsymbol{C}}^{K+1}$, where

$$\widehat{\boldsymbol{C}}^{K+1} = (\hat{c}_{K+1,1}, \ldots, \hat{c}_{K+1,2M+1}) = \boldsymbol{B}_{12}\boldsymbol{X}_{K+1} \tag{2.12}$$

By properly normalizing we can make $\hat{c}_{K+1,1} > 0$ and $\|\hat{c}^{\,K+1}\| = 1$ for $k = 0, \ldots, L-2M$. Therefore we can conclude that there exist vectors $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_{L+1-2M}$ such that

$$\boldsymbol{B}_1[\boldsymbol{X}_1 : \ldots : \boldsymbol{X}_{L-2M+1}] = \begin{bmatrix} \hat{c}_{1,1} & 0 & \cdot & 0 \\ \vdots & \hat{c}_{2,1} & \cdot & 0 \\ \cdot & & \cdot & 0 \\ \hat{c}_{1,2M+1} & \vdots & \cdot & \hat{c}_{L-2M+1,1} \\ 0 & \hat{c}_{2,2M+1} & & \cdot \\ 0 & 0 & \vdots & \cdot \\ \cdot & \cdot & & \\ 0 & 0 & & \hat{c}_{L-2M+1,2M+1} \end{bmatrix} \tag{2.13}$$

where $\hat{c}_{k,1} > 0$ and $\|\widehat{C}^k\| = 1$ for $k = 1, \ldots, L+1-2M$. Observe that in the noiseless situation

$$\widehat{C}^1 = \widehat{C}^2 = \ldots = \widehat{C}^{L-2M+1} = C. \tag{2.14}$$

Let J be the $L+1 \times L+1$ exchange matrix given by

$$J = \begin{bmatrix} 0 & 0 & \ldots & 1 \\ 0 & 0 & \cdot\cdot 1 & 0 \\ \cdot & \cdot & \ldots & \cdot \\ 1 & 0 & \ldots & 0 \end{bmatrix} \tag{2.15}$$

Consider the matrix $\tilde{T}$ given by

$$\tilde{T} = J\,T\,J \tag{2.16}$$

Observe that the eigenvalues of T and $\tilde{T}$ are same and if x is an eigenvector of T corresponding to the eigenvalue λ, i.e.,

$$Tx = \lambda x \Rightarrow J\,T\,J\,J\,x = \lambda\,Jx \Rightarrow \tilde{T}(Jx) = \lambda(Jx)$$

then Jx is an eigenvector for $\tilde{T}$ corresponding to λ. Let's denote by $\tilde{\mathbb{N}}$ the subspace generated by $\{J\hat{U}_{2M+1}, \ldots, J\hat{U}_{L+1}\}$, we call $\tilde{\mathbb{N}}$ the noise space of $\tilde{T}$.

It can be easily seen that in the noiseless situation there exists an unique basis of the noise space of $\tilde{T}$ of the form

$$\tilde{B}_2 = \begin{bmatrix} 0 & 0 & \ldots & c_1 \\ \cdot & \cdot & \ldots & \cdot \\ & 0 & & c_M \\ 0 & c_1 & & c_{M+1} \\ c_1 & \cdot & & c_M \\ \cdot & c_M & & c_1 \\ c_M & c_{M+1} & & 0 \\ c_{M+1} & c_M & & \cdot \\ c_M & \cdot & & \\ \cdot & c_1 & & \\ c_1 & 0 & \ldots & 0 \end{bmatrix} \tag{2.17}$$

In this case also our aim is to obtain the basis of the estimated noise space $\tilde{\mathbb{N}}$, i.e.,

$$\tilde{\boldsymbol{B}}_1 = [\boldsymbol{J}\widehat{\boldsymbol{U}}_{2M+1} : \cdots : \boldsymbol{J}\widehat{\boldsymbol{U}}_{L+1}] \tag{2.18}$$

to the form similar to (2.17). Proceeding exactly as in the case of $\mathbb{N}$, we reduce the basis to the following form

$$\mathbb{G} = \begin{bmatrix} 0 & 0 & \cdot & \hat{c}^*_{L-2M+1,1} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & 0 & & \cdot \\ 0 & \hat{c}^*_{2,1} & \cdot & \hat{c}^*_{L-2M+1,2M+1} \\ \hat{c}^*_{1,1} & \vdots & \cdot & 0 \\ \vdots & \hat{c}^*_{2,2M+1} & & \vdots \\ \hat{c}^*_{1,2M+1} & 0 & \cdot \quad \cdot & 0 \end{bmatrix} \tag{2.19}$$

such that for each $\widehat{\boldsymbol{C}}^*_K = (\hat{c}^*_{k,1}, \ldots, \hat{c}^*_{k,2M+1})$, $k = 1, \ldots, L-2M+1$; $\hat{c}^*_{k,1} > 0$ and $\|\widehat{\boldsymbol{C}}^*_K\| = 1$. As in the case of $\mathbb{N}$, in the noiseless situation

$$\widehat{\boldsymbol{C}}^*_1 = \widehat{\boldsymbol{C}}^*_2 = \ldots = \widehat{\boldsymbol{C}}^*_{L+1-2M} = \boldsymbol{C} \tag{2.20}$$

It is further observed that $\mathbb{G} = \boldsymbol{J}\boldsymbol{B}_1[\boldsymbol{X}_1 : \ldots : \boldsymbol{X}_{L-2M+1}]$, i.e.,

$$\begin{bmatrix} 0 & 0 & \cdot & \hat{c}^*_{L-2M+1,1} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & 0 & & \cdot \\ 0 & \hat{c}^*_{2,1} & \cdot & \hat{c}^*_{L-2M+1,2M+1} \\ \hat{c}^*_{1,1} & \vdots & \cdot & 0 \\ \vdots & \hat{c}^*_{2,2M+1} & \cdot & \cdot \\ \hat{c}^*_{1,2M+1} & 0 & \cdot \quad \cdot & 0 \end{bmatrix}$$

$$= \boldsymbol{J}\begin{bmatrix} \hat{c}_{1,1} & 0 & \cdot & 0 \\ \vdots & \hat{c}_{2,1} & \cdot & 0 \\ \cdot & & & 0 \\ \hat{c}_{1,2M+1} & \vdots & \cdot & \hat{c}_{L-2M+1,1} \\ 0 & \hat{c}_{2,2M+1} & & \cdot \\ 0 & 0 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \\ 0 & 0 & & \hat{c}_{L-2M+1,2M+1} \end{bmatrix} \tag{2.21}$$

Now observe that since (2.14) and (2.20) are true, it is quite natural that any one of the $\widehat{C}^K$ for $k = 1, \ldots, L - 2M + 1$ or $\widehat{C}_K^*$ for $k = 1, \ldots, L - 2M + 1$ can be used to estimate the frequencies. In fact the use of $\widehat{C}_K^{**} = \frac{1}{2}(\widehat{C}^K + \widehat{C}_K^*)$; $k = 1, \ldots, L - 2M + 1$ always ensure that the estimated coefficients of the polynomial prediction equation (2.3) satisfy the symmetry constraint and roots of

$$\hat{c}_{k,1}^{**} Z^{2M} + \hat{c}_{k,2}^{**} Z^{2M-1} + \cdots + \hat{c}_{k,2M+1}^{**} = 0 \qquad (2.22)$$

are of the form $\exp(\pm i\hat{\omega}_K)$, for $k = 1, \ldots, L - 2M + 1$. We use all $\widehat{C}_K^{**}$; $k = 1, \ldots, L - 2M + 1$ to estimate ω. We take the average of all $\widehat{C}_K^{**}$'s and use (2.3) to get the final estimate $\widehat{\omega}$ of ω. We call the resulting estimate $\widehat{\omega}$, the **Noise Space Decomposition** (**NSD**) estimates.

The following consistency result can be established.

THEOREM 1 *Under the assumptions of the model* (1.1), *the estimate* $\widehat{\omega}$ *of* ω *obtained by the method described above is strongly consistent, i.e.,*

$$\widehat{\omega}_{NSD} \xrightarrow{a.s.} \omega \qquad (2.23)$$

For the proof of Theorem 1, *see Appendix.*

3. MODIFIED NOISE SPACE DECOMPOSITION METHOD

It is well known (Harvey; 1981, Ch. 4.5) that when a regular likelihood (differentiable up to third order) is maximized through the Newton-Raphson, scoring or a related algorithm, the estimates obtained after one single round of iteration already have the same asymptotic properties as the exact least squares estimates. This holds, if the starting values have been chosen $\sqrt{N}$ – consistently. Now, since the NSD method is strongly consistent we combine the NSD method with one single round of scoring algorithm. This way the asymptotic error variances should (in theory, at least) coincide with the asymptotic variance irrespective of the distributional form of the error term. We call the resulting estimates obtained after one round of iteration with NSD as starting values, the Modified Noise Space Decomposition (**MNSD**) estimates.

One way of implementing this idea would be the following:

Let us write the model (1.1) in the vector form

$$Y = A(\omega)\alpha + \varepsilon \tag{3.1}$$

where $A(\omega) = [A_1(\omega)A_2(\omega)\ldots A_N(\omega)]^T$ with $A_k(\omega) = [\cos(\omega_1 k)\ \sin(\omega_1, k)\ldots\cos(\omega_M k)\ \sin(\omega_M k))^T$, $\alpha = (A_1B_1\ldots A_MB_M)$, $Y = (Y_1\ldots\ Y_N)^T$ and $\varepsilon = (\varepsilon_1\ldots\varepsilon_N)^T$.

Now consider the concentrated residual sum of squares

$$Y^*[I - P_A(\omega)]Y = Y^*[I - A(\omega)[A^*(\omega)A(\omega)]^{-1}A^*(\omega)]Y \tag{3.2}$$

To obtain the least squares estimates first (3.2) can be minimized with respect to ω and then the estimate of α can be obtained using linear regression technique. For details see Kundu (1993b). We obtain the MNSD estimates after one step minimization of (3.2) using NSD estimates as starting values.

4. CONFIDENCE INTERVALS

In this section we propose different confidence intervals for the frequencies. We propose an asymptotic confidence interval and two bootstrap confidence intervals.

4.1. Asymptotic Confidence Intervals

In this subsection we discuss the confidence intervals for the frequencies based on their asymptotic distribution. It may be observed that (Kundu and Mitra; 1996) that the asymptotic distribution of the exact LSE of the frequencies is of the following form

$$N^{-3/2}[\hat{\omega}_k - \omega_k] \sim N\left(0, \frac{24\sigma^2}{(A_k^2 + B_k^2)}\right) \tag{4.1.1}$$

which eventually coincides with the distribution of the approximate LSE proposed by Walker (1971).

Since the MNSD estimates proposed in Section 3 has the same asymptotic properties as the exact LSE, we take the MNSD estimates as $\hat{\omega}_k$ in (4.12).

4.2. Percentile Bootstrap Confidence Intervals

In this subsection we construct the percentile bootstrap confidence intervals for ω_k's using the method suggested by Efron (1982).

Suppose we have a sample of size N; $y_1, \ldots, y_N$ coming from (1.1). We propose the following algorithm to obtain the confidence intervals

(1) Estimate $(\omega_1, \ldots, \omega_M)$ from $y_1, \ldots, y_N$ using MNSD method.
(2) Estimate $\hat{\varepsilon}_i = y_i - \hat{y}_i$, $i = 1, \ldots, N$, where

$$\hat{y}_i = \sum_{k=1}^{M} (\hat{A}_k \cos(\hat{\omega}_k t) + \hat{B}_k \sin(\hat{\omega}_k t)).$$

(Here $\hat{A}_k$ and $\hat{B}_k$ are obtained using linear regression, see Kundu (1994))
(3) Draw a random sample of size N from $\{\hat{\varepsilon}_1, \ldots, \hat{\varepsilon}_N\}$ with replacement, let it be $\{\hat{\varepsilon}_{B_1}, \ldots, \hat{\varepsilon}_{B_N}\}$.
(4) Obtain bootstrap sample $y_1^*, \ldots, y_N^*$; where

$$y_i^* = \hat{y}_i + \hat{\varepsilon}_{B_i}, \quad i = 1, \ldots, N.$$

(5) Estimate $(\omega_1, \ldots, \omega_M)$ from $y_1^*, \ldots, y_N^*$ using MNSD method. Denote it by ω_k^*, $K = 1, \ldots, M$.
(6) Repeat the steps (3) to (5) NBOOT times.
(7) Order these NBOOT estimates corresponding to each ω_k.
(8) Estimate $\hat{L}_{PB}(\alpha/2)$ by NBOOT $\alpha/2$th order statistics and $\hat{U}_{PB}(\alpha/2)$ by NBOOT (1-$\alpha/2$)th order statistics for each set of ω_k^* and claim that $(\hat{L}_{PB}(\alpha/2), \hat{U}_{PB}(\alpha/2))$ to be the $100(1-\alpha)\%$ percentile bootstrap-t confidence intervals for ω_k.

4.3. Bootstrap-*t* Confidence Intervals

In this subsection we construct the bootstrap-t confidence intervals based on the method suggested by Hall (1988).

We propose the following algorithm for computing the bootstrap-t confidence intervals,
Step (1) to (4) same as Percentile Boot Strap method.

(5) Estimate $(\omega_1, \ldots, \omega_M)$ from $y_1^*, \ldots, y_N^*$ using MNSD method, denote it by $\hat{\omega}_k^*$ and also the estimate of σ^2 as $\hat{\sigma}_B^2$.

6) Obtain for each ω_k, $k = 1, \ldots, M$

$$T_{k,i} = \frac{\sqrt{N}(\hat{\omega}_k^* - \hat{\omega}_k)}{\hat{\sigma}_B}$$

(7) Repeat the steps (3) to (6) NBOOT times.

(8) For each ω_k, order the NBOOT number of $T_{k,i}$'s. Estimate $\hat{L}_{TB}(\alpha/2)$ by $\hat{\omega}_k + \sqrt{N}\hat{\sigma}$ (NBOOT $\alpha/2$ th order statistics from $T_{k,i}$'s) and $\hat{U}_{TB}(\alpha/2)$ by $\hat{\omega}_k + \sqrt{N}\hat{\sigma}$ (NBOOT $(1-\alpha/2)$ th order statistics from $T_{k,i}$'s). Now claim that $(\hat{L}_{TB}(\alpha/2), \hat{U}_{TB}(\alpha/2))$ to be the $100(1-\alpha)\%$ bootstrap-t confidence interval for ω_k.

5. MONTE CARLO SIMULATIONS

We have performed Monte Carlo simulation study to ascertain the behavior of NSD and MNSD estimates for moderate sample sizes and different ranges of the error variances σ^2. All these simulations have been done on the HP-9000 computer at the Indian Institute of Technology Kanpur, using the IMSL random deviate generator.

We consider the following models;

$$y_t = 1.5\cos(\omega t) + 1.5\sin(\omega t) + \varepsilon_t, \qquad t = 1, \ldots, 25 \tag{5.1}$$

$$y_t = 1.5\cos(\omega_1 t) + 1.5\sin(\omega_1 t) + 2.5\cos(\omega_2 t) + 2.5\sin(\omega_2 t) + \varepsilon_t, \quad t = 1, \ldots, 25 \tag{5.2}$$

The error random variable $\{\varepsilon_t\}$ is white and Gaussian with variances σ^2. The frequency ω is taken to be 0.25π, 0.50π, 0.75π for (5.1) and $\omega_1 = 0.5$, $\omega_2 = 2.5$ for (5.2). In each case, 500 independent trials using different ε_t sequences are performed. The variance of the error random variables is varied from 0.01 to 1.5. In each case we computed $(\omega_1, \ldots, \omega_M)$ by NSD, MNSD, TLS-ESPRIT (Roy and Kailath; 1989) and Quinn's (Quinn; 1994) methods. For each ω, we computed the average estimates and the mean squared errors (MSE) over 500 replications and also the corresponding Cramer-Rao lower bound (CRLB).

It is observed that the performance of the NSD estimates changes with the different values of L. We observed that the MSE starts decreasing as L increases and for $N = 25$, the best performance (min

MSE) of the NSD occurs at $L = 15$ ($\cong (3/5)N$). The performance of the MNSD estimates does not seem to be much affected with the variation of L. We report the best performance of NSD and MNSD along with their CRLB in Table I for the model (5.1) and that of model (5.2) in Table III.

TABLE I

σ^2		$\omega = .25\pi$	$\omega = .50\pi$	$\omega = .75\pi$
	CRLB	0.785398	1.570796	2.356194
		3.41333E-06	3.41333E-06	3.41333E-06
	NSD	0.785453	1.570718	2.356080
		4.33998E-06	4.12378E-06	4.08629E-06
0.01	MNSD	0.785389	1.570732	2.356134
		3.73869E-06	3.16699E-06	2.67494E-06
	TLS-ESP	0.785421	1.570707	2.356098
		4.69288E-06	4.03914E-06	3.17762E-06
	QUINN	0.780732	1.570961	2.366450
		2.89467E-05	4.59882E-06	1.08402E-04
	CRLB	0.785398	1.570796	2.356194
		6.82667E-06	6.82667E-06	6.82667E-06
	NSD	0.785500	1.570686	2.356013
		8.69172E-06	8.25931E-06	8.20251E-06
0.02	MNSD	0.785375	1.570705	2.356124
		7.48658E-06	6.35441E-06	5.35563E-06
	TLS-ESP	0.785439	1.570672	2.356061
		9.39960E-06	8.09771E-06	6.36954E-06
	QUINN	0.780695	1.570911	2.366391
		3.65142E-05	9.16851E-06	1.10426E-04
	CRLB	0.785398	1.570796	2.356194
		1.02400E-05	1.02400E-05	1.02400E-05
	NSD	0.785546	1.570662	2.355952
		1.30522E-05	1.24045E-05	1.23413E-05
0.03	MNSD	0.785359	1.570683	2.356123
		1.12405E-05	9.55837E-06	8.04419E-06
	TLS-ESP	0.785455	1.570645	2.356034
		1.41192E-05	1.21739E-05	9.57469E-06
	QUINN	0.780658	1.570867	2.366341
		4.40984E-05	1.37549E-05	1.12635E-04
	CRLB	0.785398	1.570796	2.356194
		1.36533E-05	1.36533E-05	1.36533E-05
	NSD	0.785592	1.570642	2.355897
		1.74222E-05	1.65586E-05	1.65001E-05
0.04	MNSD	0.785342	1.570665	2.356126
		1.50005E-05	1.27775E-05	1.07415E-05
	TLS-ESP	0.785471	1.570623	2.356011
		1.88515E-05	1.62672E-05	1.27937E-05
	QUINN	0.780623	1.570826	2.366294
		5.17059E-05	1.83559E-05	1.14946E-04

TABLE I (Continued)

σ^2		$\omega = .25\pi$	$\omega = .50\pi$	$\omega = .75\pi$
	CRLB	0.785398	1.570796	2.356194
		1.70667E-05	1.70667E-05	1.70667E-05
	NSD	0.785638	1.570623	2.355843
		2.18019E-05	2.07213E-05	2.06775E-05
0.05	MNSD	0.785326	1.570648	2.356132
		1.87662E-05	1.60110E-05	1.34479E-05
	TLS-ESP	0.785487	1.570604	2.355992
		2.35964E-05	2.03773E-05	1.60256E-05
	QUINN	0.780588	1.570786	2.366251
		5.73365E-05	2.29709E-05	1.17325E-04
	CRLB	0.785398	1.570796	2.356194
		1.70667E-04	1.70667E-04	1.70667E-04
	NSD	0.787822	1.570307	2.353656
		2.31700E-04	2.15971E-04	2.26471E-04
0.05	MNSD	0.784518	1.570232	2.356872
		1.93809E-04	1.74121E-04	1.47636E-04
	TLS-ESP	0.786194	1.570276	2.355635
		2.53689E-04	2.24849E-04	1.77509E-04
	QUINN	0.782626	1.567484	2.360877
		5.05633E-04	3.63172E-04	6.84565E-04
	CRLB	0.785398	1.570796	2.356194
		3.41333E-04	3.41333E-04	3.41333E-04
	NSD	0.791625	1.570268	2.348618
		5.85014E-04	4.65771E-04	2.00918E-03
1.0	MNSD	0.783479	1.569923	2.356048
		4.05486E-04	3.78684E-04	1.85945E-03
	TLS-ESP	0.785290	1.573467	2.357093
		1.88270E-03	5.50305E-04	1.66025E-03
	QUINN	0.784584	1.562216	2.355021
		1.25684E-03	9.45444E-04	2.38975E-03
	CRLB	0.785398	1.570796	2.356194
		5.12000E-04	5.12000E-04	5.12000E-04
	NSD	0.801783	1.571409	2.337188
		9.85854E-03	4.55059E-03	1.88180E-02
1.5	MNSD	0.786734	1.571403	2.351391
		9.60798E-03	4.89990E-04	1.68852E-02
	TLS-ESP	0.799084	1.581543	2.372139
		5.34451E-03	3.26071E-02	1.42323E-02
	QUINN	0.787966	1.560561	2.358988
		4.46255E-03	3.07516E-03	2.13763E-02

We also performed a simulation study to investigate the performance of the different confidence intervals discussed in Section 4 with respect to their average length and coverage probabilities. We consider the simulation model (5.1) with ε_t white and Gaussian having error variance σ^2. Results are obtained for $\omega = 0.25\pi$, 0.50π and 0.75π. For each ω, simulations were performed for $\sigma^2 = 0.01$, 0.05 and 0.1.

TABLE II

	$\omega = .25\pi$	$\omega = .50\pi$	$\omega = .75\pi$
		$\sigma^2=0.01$	
Asymptotic	0.005740	0.005716	0.005714
(Cov. Prob.)	(.820)	(.849)	(.818)
Per – Boot	0.006130	0.005764	0.005496
(Cov. Prob.)	(.881)	(.840)	(.811)
Bootstrap-*t*	0.006780	0.006375	0.006060
(Cov. Prob.)	(.889)	(.840)	(.820)
		$\sigma^2=0.05$	
Asymptotic	0.012812	0.012800	0.012794
(Cov. Prob.)	(.810)	(.850)	(.819)
Per – Boot	0.013699	0.012915	0.012346
(Cov. Prob.)	(.881)	(.840)	(.810)
Bootstrap-*t*	0.015173	0.014288	0.013620
(Cov. Prob.)	(.890)	(.849)	(.821)
		$\sigma^2=0.10$	
Asymptotic	0.040690	0.040699	0.040577
(Cov. Prob.)	(.799)	(.830)	(.850)
Per – Boot	0.044166	0.042175	0.040585
(Cov. Prob.)	(.850)	(.800)	(.860)
Bootstrap-*t*	0.048556	0.047148	0.044970
(Cov. Prob.)	(.883)	(.860)	(.860)

TABLE III

σ^2		$\omega_1 = 0.5$	$\omega_2 = 2.5$
	CRLB	0.500000	2.500000
		3.41333E-06	1.92000E-06
	NSD	0.500032	2.499929
		3.31704E-06	5.46915E-06
0.01	MNSD	0.499997	2.499955
		2.09698E-06	3.27784E-06
	TLS-ESP	0.500009	2.499931
		2.50940E-06	4.37303E-06
	QUINN	0.497793	2.497986
		1.04618E-05	1.30069E-05
	CRLB	0.500000	2.500000
		6.82667E-06	3.84000E-06
	NSD	0.500056	2.499883
		6.63565E-06	1.09256E-05
0.02	MNSD	0.499992	2.499951
		4.19578E-06	6.55044E-06
	TLS-ESP	0.500019	2.499903
		5.02088E-06	8.73761E-06
	QUINN	0.497850	2.497994
		1.68289E-05	2.19167E-05

TABLE III (Continued)

σ^2		$\omega_1 = 0.5$	$\omega_2 = 2.5$
	CRLB	0.500000	2.500000
		1.02400E-05	5.76000E-06
	NSD	0.500080	2.499841
		9.95597E-06	1.63778E-05
0.03	MNSD	0.499986	2.499954
		6.29682E-06	9.82452E-06
	TLS-ESP	0.500029	2.499880
		7.53475E-06	1.31020E-05
	QUINN	0.497932	2.498008
		2.36584E-05	3.08083E-05
	CRLB	0.500000	2.500000
		1.36533E-05	7.68000E-06
	NSD	0.500103	2.499802
		1.32780E-05	2.18287E-05
0.04	MNSD	0.499980	2.499961
		8.40023E-06	1.31025E-05
	TLS-ESP	0.500038	2.499861
		1.00511E-05	1.74692E-05
	QUINN	0.497916	2.498022
		2.93299E-05	3.95915E-05
	CRLB	0.500000	2.500000
		1.70667E-05	9.60000E-06
	NSD	0.500125	2.499763
		1.66019E-05	2.72801E-05
0.05	MNSD	0.499974	2.499970
		1.05061E-05	1.63857E-05
	TLS-ESP	0.500048	2.499845
		1.25700E-05	2.18412E-05
	QUINN	0.497958	2.498039
		3.55309E-05	4.84084E-05
	CRLB	0.500000	2.500000
		1.70667E-04	9.60000E-05
	NSD	0.501137	2.498171
		1.68378E-04	2.80284E-04
0.5	MNSD	0.499671	2.500758
		1.08109E-04	1.74437E-04
	TLS-ESP	0.500501	2.499336
		1.28815E-04	2.32245E-04
	QUINN	0.498717	2.499234
		2.96343E-04	4.58838E-04
	CRLB	0.500000	2.500000
		3.41333E-04	1.92000E-04
	NSD	0.502438	2.495981
		3.45122E-04	6.00205E-04
1.0	MNSD	0.499327	2.501724
		2.23072E-04	3.81547E-04
	TLS-ESP	0.501026	2.499141
		2.65772E-04	1.54417E-03
	QUINN	0.499399	2.500623
		5.77437E-04	9.58823E-04

TABLE III (Continued)

	CRLB	0.500000	2.500000
		5.12000E-04	2.88000E-04
	NSD	0.504321	2.493360
		5.72005E-04	1.45885E-03
1.5	MNSD	0.499018	2.503483
		3.43910E-04	1.05221E-03
	TLS-ESP	0.496528	2.495730
		2.47949E-03	3.26896E-02
	QUINN	0.499606	2.500833
		8.53128E-04	2.51304E-03

Average length of the confidence intervals (with nominal level 0.90) and the corresponding coverage probabilities over 500 simulations are reported for all the methods in Table II. The bootstrapping number NBOOT is taken as 100 for both the bootstrap methods.

6. CONCLUSIONS

In this article, we propose a new non-iterative method for estimating the frequencies of the model (1.1) when the number of frequencies is known apriori. If the number of harmonic components is unknown, then we can first estimate the number of frequencies by the method of Kundu (1992) and then use the proposed method to estimate the frequencies. First we transform the model (1.1) to an undamped superimposed exponential signals model, then use extended order modeling and decompose the noise space by singular value decomposition technique. It is further proved that the proposed non-iterative technique yields estimates that are strongly consistent.

Simulation results show very satisfactory performance of NSD estimates even at high values of error variance and moderate sample sizes for both single sinusoid as well as multiple sinusoids. The proposed one step modified estimate MNSD, performs even much better than the NSD in the simulations study. The performance of MNSD almost attains the CRLB in the cases considered.

The choice of L obviously affects the performance of the NSD estimates. Clearly L should be at least $M+1$, but the natural question is why it should be larger than that? Although no theoretical justifications have been given in the literature, but it is observed that extended order modeling always helps to improve the performance of the estimators. Some heuristic justifications can be found in Tufts and Kumaresan (1982). It seems more theoretical work is needed in this direction. Here we have observed that as L increases the MSE starts decreasing for NSD. It reaches a minimum at $L = 15\ (\cong 3/5N)$, when the sample size is 25. The performance of the MNSD estimates does not seem to be affected much with variation in L.

Among the three confidence intervals for the frequencies discussed in Section 4, the bootstrap-t confidence intervals gives the highest coverage probabilities although the average length of these intervals is marginally larger than the other two. It is also observed that the asymptotic confidence interval performs better than the percentile bootstrap intervals in terms of shorter average length and higher coverage probabilities when $\omega = 0.50\pi$. But the percentile bootstrap gives higher coverage probabilities and almost same length intervals when $\omega = 0.25\pi$ or 0.75π as compared with the asymptotic confidence intervals. It is further observed that all the three confidence intervals are symmetric about the true parameter in the cases considered. Based on the results of the simulations, we recommend to use bootstrap-t method for finding confidence intervals.

Acknowledgement

Work is partly supported by a grant (No. SR/OY/M-06/93) of the Department of Science and Technology, Government of India.

References

Bai, Z. D., Krishnaiah, P. R. and Zhao, L. C. (1986) "On the 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.

Bai, Z. D., Miao, B. Q. and Rao, C. R. (1990) "Estimation of direction of arrival of signals: Asymptotic results," *Advances in Spectrum Analysis and Array Processing*, vol II, 327–347, Ed. S. Haykins, Prentice-Hall, NJ.

Dragosevic, M. V. and Stankovic, S. S. (1989) "A generalized least squares method for frequency estimation," *IEEE Trans. on Acoust. Speech and Signal Processing*, **37**, 805–819.

Efron, B. (1982) "The Jacknife, the Bootstrap and other resampling plans," *Society for Industrial and Applied Mathematics*, Philadelphia, Pennsylvania, USA.

Ensor, K. B. and Newton, H. J. (1988) "The effect of order estimation on estimating the peak frequency of an autoregressive spectral density," *Biometrika*, **75**, 587–589.

Fuchs, J. J. (1988) "Estimating the number of sinusoids in additive white noise", *IEEE Trans. Acoust. Speech and Signal Processing*, **34**, 1201–1209.

Hall, P. (1988) "Theoretical comparison of bootstrap confidence intervals," *Ann. Stat.*, **16**, 3, pp. 927–953.

Hannan, E. J. (1973) "The estimation of frequency," *Jour. Appl. Prob.*, **10**, pp. 510–519.

Hannan, E. J. (1993) "Determining the number of jumps in a spectrum," *Developments in Time Series Analysis*, Ed. T. Subba Rao, Chapman and Hall, London, 127–138.

Harvey, A. C. (1981) "The Econometric Analysis of Time Series," Philip Allan, New York.

Hasan, T. (1982) "Nonlinear time series regression for a class of amplitude modulated cosinusoids," *Jour. Time Series Analysis*, **3**, 109–122.

Hua, Y. and Sarkar, T. K. (1990) "Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise", *IEEE Trans. Signal Processing*, **38**, 814–824.

Khan, M. S., Mackisack, M. C., Osborne, M. R. and Smyth, G. K. (1993) "On the consistency of Prony's method and related algorithms," *Jour. Graphical and Computational Stat.*, **1**, 4, 329–349.

Kavalieris, L. and Hannan, E. J. (1994) "Determining the number of terms in trigonometric regression," *Journal of Time Series Analysis*, **15**, 613–625.

Kay, S. M. and Marple, S. L. (1981) "Spectral analysis – A modern prospective," *Proc. IEEE*, **69**, 1380–1419.

Kundu, D. (1992) "Estimating the number of signals using information theoretical criterion," *Journal of Statistical Computation and simulation*, **44**, 117–131.

Kundu, D. (1993b) "Estimating the parameters of undamped exponential signals," *Technometrics*, **35**(2), 215–218.

Kundu, D. (1994) "A modified Prony algorithm to estimate damped or undamped exponential signals," *Sankhya A*, **56**(3), 524–544.

Kundu, D. (1993a) "Asymptotic theory of least squares estimates of a particular nonlinear regression model," *Statistics and Prob. Letters*, **18**(1), 13–17.

Kundu, D. (1996) "Estimating the number of sinusoids and its performance analysis," Technical Report, Indian Institute of Technology, Kanpur.

Kundu, D. and Mitra, A. (1996) "Asymptotic behavior of the least squares estimates of a particular nonlinear time series model," *Communications in Statistics, Theory and Methods*, **25**(1), 133–142.

Mackisack, M. S. and Poskitt, D. S. (1989) "Autoregressive frequency estimation," *Biometrika*, **76**, 565–575.

Petrov, V. V. (1975) "Sum of Independent Random Variables," Springer–Varlag, New York.

Prony, R. (1795) "Essai experimentale et analytique," *Jour. Ecole Polytechnique* (Paris), **1**, 24–76.

Quinn, B. G. (1989) "Estimating the number of terms in a sinusoidal regression", *Jour. Time Series Anal.*, **10**(1), 71–75.

Rao, C. R. (1973) "Linear Statistical inference and it's applications," 2nd Ed., Wiley Eastern.

Reddy, V. U. and Biradar, L. S. (1993). "SVD based information theoretic criterion for detection of the number of damped/undamped sinusoids and their performance analysis", *IEEE Trans. Signal Processing*, **41**, 2872–2881.

Rice, J. A. and Rosenblatt, M. (1988) "On frequency estimation," *Biometrika*, **75**(3), 477–484.

Richards, F. S. G. (1961) "A method of maximum likelihood estimation," *Jour. Royal Stat. Soc. B*, **23**, 469–475.

Roy, R. and Kailath, T. (1988) "ESPRIT – Estimation of signal paramters via rotational invarinace techniques", Ed. Deprettre, E. F., SVD and Signal Processing Algorithms, *Applications and Architectures, Elsevier Amsterdam*, 233–265.

Tufts, D. W. and Kumaresan, R. (1982) "Estimation of frequencies of multiple sinusoids: Making linear prediction perform like maximum likelihood," *Proc IEEE*, **70**, 975–989.

Von-Neumann, J. (1937) "Some matrix inequalities and metrization of metric space," *Tomsk Univ. Rev.*. **1**. 386–400.

Walker, A. M. (1971) "On the estimation of a Harmonic component in a time series with stationary independent residuals," *Biometrika*, **58**, 21–36.
Whittle, P. (1953) "The simultaneous estimation of a time series Harmonic component and covariance structure," *Trab. Estad.*, **3**, pp. 43–57.

APPENDIX

For the proof of Theorem 1, we need the following lemmas.

LEMMA 1 *Let* $\boldsymbol{P}=((p_{ij}))$ *and* $\boldsymbol{Q}=((q_{ij}))$ *are two Hermitian* $m \times m$ *matrices with spectral decompositions*

$$\begin{aligned} \boldsymbol{P} &= \sum_{i=1}^{m} \delta_i \boldsymbol{p}_i \boldsymbol{p}_i^* \quad \delta_1 \geq \delta_2 \geq \cdots\cdots \geq \delta_m \\ \boldsymbol{Q} &= \sum_{i=1}^{m} \lambda_i \boldsymbol{q}_i \boldsymbol{q}_i^* \quad \lambda_1 \geq \lambda_2 \geq \cdots\cdots \geq \lambda_m \end{aligned} \tag{A.1}$$

where δ_i*'s and* λ_i*'s are eigen values of* $\boldsymbol{P}$ *and* $\boldsymbol{Q}$ *respectively,* $\boldsymbol{p}_i$ *and* $\boldsymbol{q}_i$ *are orthogonal normalized eigenvalues associated with* δ_i *and* λ_i *respectively for* $i = 1, \ldots, m$. *Further assume that*

$$\lambda_{n_{h-1}} + 1 = \ldots\ldots = \lambda_{n_h} = \tilde{\lambda}_h, \quad n_o = 0 < n_1 < \ldots < n_s = p;$$
$$h = 1, \ldots, s \ \tilde{\lambda}_1 > \tilde{\lambda}_2 > \ldots > \tilde{\lambda}_s$$

and that $|p_{ik} - q_{ik}| < \alpha, \quad i, k = 1, \ldots, m$ *then there is a constant* M *independent of* α *such that*

(1) $|\delta_i - \lambda_i| < M\alpha \quad i = 1, \ldots, m$

$$\text{(2)} \quad \sum_{i=n_{h-1}}^{n_h} \boldsymbol{p}_{i+1} \boldsymbol{p}_i^* = \sum_{i=n_{h-1}}^{n_h} \boldsymbol{q}_{i+1} \boldsymbol{q}_i^* + \boldsymbol{C}^{(h)} \tag{A.2}$$

$$\textit{with} \quad C^{(h)} = ((C_{lk}^{(h)})), \quad |C_{lk}^{(h)}| \leq M\alpha$$

Proof It follows from Von-Neumann's (1937) inequality, see also Bai, Miao and Rao (1991). ■

LEMMA 2

$$\frac{1}{N}\boldsymbol{A}^*\boldsymbol{A} = \sigma^2\boldsymbol{I}_{L+1} + \Omega^{(L)}\boldsymbol{D}\,\Omega^{(L)*} + O\left(\frac{\sqrt{\log\log N}}{N}\right) \quad \text{a.s.} \tag{A.3}$$

where $\boldsymbol{D}_{2M\times 2M} = diag\,\{|C_1^2|,\dots,|C_M^2|,\ |D_1^2|,\dots,|D_M^2|\}$ *and*

$$\Omega^{(L)}_{L+1\times M} = \begin{bmatrix} e^{-i\omega_1} & & e^{-i\omega_M} & e^{i\omega_1} & & e^{i\omega_M} \\ \vdots & \vdots\ \vdots & \vdots & \vdots & \vdots\ \vdots & \vdots \\ e^{-i(L+1)\omega_1} & \vdots & e^{-i(L+1)\omega_M} & e^{-i(L+1)\omega_1} & \vdots & e^{-i(L+1)\omega_M} \end{bmatrix}$$

Proof We have

$$\frac{1}{N}\boldsymbol{A}^*\boldsymbol{A} = \boldsymbol{T} = \frac{1}{N}((t_{iK})) \tag{A.4}$$

with the following renaming of the parameters of (2.2) as

$$\alpha_i = \begin{cases} C_i & \text{for} \quad i = 1,\dots,M \\ D_{i-M} & \text{for} \quad i = M+1,\dots,2M \end{cases} \tag{A.5}$$

$$\beta_i = \begin{cases} \omega_i & \text{for} \quad i = 1,\dots,M \\ -\omega_{i-M} & \text{for} \quad i = M+1,\dots,2M \end{cases} \tag{A.6}$$

we have the following

$$\begin{aligned} \frac{1}{N}t_{ik} &= \frac{1}{N}\sum_{l=0}^{N-L-1}\bar{y}_{l+i}\,y_{l+k} \\ &= \frac{1}{N}\sum_{l=0}^{N-L-1}\left(\sum_{u=1}^{2M}\bar{\alpha}_u\exp(-i\beta_u(1+i)) + \bar{\varepsilon}_{l+i}\right) \\ &\quad \left(\sum_{u=1}^{2M}\alpha_u\exp(i\beta_u(1+k)) + \varepsilon_{l+k}\right) \\ &= \frac{1}{N}\sum_{l=0}^{N-L-1}\left[\sum_{u=1}^{2M}|\alpha_u|^2\exp(i\beta_u(k-i)\right) \\ &\quad + \sum_{u\neq v}\bar{\alpha}_u\,\alpha_v\exp(i(\beta_v(1+k) - \beta_u(1+i))) \end{aligned}$$

$$
\begin{aligned}
&+ \varepsilon_{l+k}\left(\sum_{u=1}^{2M} \bar{\alpha}_u \exp(-i\beta_u(1+i))\right) \\
&+ \bar{\varepsilon}_{l+i}\left(\sum_{u=1}^{2M} \alpha_u \exp(i\beta_u(1+k)) + \varepsilon_{l+k}\right) \\
&+ \bar{\varepsilon}_{l+i}\,\varepsilon_{l+k}\Bigg] \\
&= R_1 + R_2 + R_3 + R_4 + R_5 \quad \text{(say)} \qquad \text{(A.7)}
\end{aligned}
$$

Observe that $R_2 = O(1/N)$ and $R_3 = O(\sqrt{\log\log N}/N)$ a.s. and $R_4 = O(\sqrt{\log\log N}/N)$ a.s. (see Petrove (1975, page 375) and by the law of iterative logarithm of M dependent sequence

$$
R_5 = \begin{cases} O\left(\sqrt{\frac{\log\log N}{N}}\right) & \text{if } \; i \neq k \\ \sigma^2 + O\left(\sqrt{\frac{\log\log N}{N}}\right) & \text{if } \; i = k \end{cases}
$$

this proves lemma 2. ■

LEMMA 3 *Let $g_n(x)$ be a sequence of polynomials of degree k, with roots $x_1^{(n)}, \ldots, x_k^{(n)}$ for each n. Let $g(x)$ be a polynomial of degree Q, with roots $x_1, \ldots, x_q$, $Q \leq k$. If $g_n(x) \to g(x)$ and $n \to \infty$ then with proper rearrangement the roots of $g_n(x)$, $x_j^{(n)}$ converge to the roots of $g(x)$, i.e., to x_j.*

Proof See Bai *et al.* (1986)

Proof of Theorem 1 From Lemma 2 it follows that;

$$
\boldsymbol{T} \xrightarrow{\text{a.s.}} \sigma^2 \boldsymbol{I}_{L+1} + \Omega^{(L)} \boldsymbol{D} \Omega^{(L)^*} = \boldsymbol{S}
$$

and

$$
\tilde{\boldsymbol{T}} \xrightarrow{\text{a.s.}} \sigma^2 \boldsymbol{I}_{L+1} + \boldsymbol{J}\Omega^{(L)} \boldsymbol{D} \Omega^{(L)^*} \boldsymbol{J} = \boldsymbol{S}_J^*
$$

Observe that the eigen values of $\boldsymbol{S}$ are of the form

$$
\lambda_1 \geq \lambda_2 \geq \ldots\ldots\ldots \geq \lambda_{2M} > \lambda_{2M+1} = \ldots = \lambda_{L+1} = \sigma^2 \qquad \text{(A.8)}
$$

since $\Omega^{(L)} \boldsymbol{D} \Omega^{(L)^*}$ is of rank $2M$. Let the singular value decomposition of $\boldsymbol{S}$ be

$$\boldsymbol{S} = \sum_{i=1}^{L+1} \lambda_i \boldsymbol{s}_i \boldsymbol{s}_i^* \tag{A.9}$$

where $\boldsymbol{s}_i$ is the normalized eigenvector corresponding to the eigenvalue λ_i and $\boldsymbol{s}_i$'s are orthogonal to each other. Therefore using Lemma 1

$$\sum_{i=2M+1}^{L+1} \hat{\boldsymbol{U}}_i \hat{\boldsymbol{U}}_i^* \xrightarrow{\text{a.s.}} \sum_{i=2M+1}^{L+1} \boldsymbol{s}_i \boldsymbol{s}_i^* \tag{A.10}$$

(A.9) implies that the vector space generated by $\{\hat{\boldsymbol{U}}_{2M+1}, \ldots, \hat{\boldsymbol{U}}_{L+1}\}$ converges to the vector space generated by $\{\hat{\boldsymbol{s}}_{2M+1}, \ldots, \boldsymbol{s}_{L+1}\}$.

Now the former one has a unique basis of the form

$$\begin{bmatrix} \hat{c}_{1,1} & 0 & \cdot & 0 \\ \vdots & \hat{c}_{2,1} & \cdot & 0 \\ \hat{c}_{1,2M+1} & \vdots & \cdot & \dot{0} \\ 0 & \hat{c}_{2,2M+1} & & \hat{c}_{L-2M+1,1} \\ 0 & 0 & \cdot & \cdot \\ \cdot & \cdot & \vdots & \cdot \\ 0 & 0 & & \hat{c}_{L-2M+1,2M+1} \end{bmatrix} \tag{A.11}$$

with $\hat{c}_{k,1} > 0$ and $\|\hat{c}^k\| = 1$ where $\hat{\boldsymbol{c}}^k = (\hat{c}_{k,1}, \ldots, \hat{c}_{k,2M+1})$ for $k = 1, \ldots, L - 2M + 1$, whereas the later one has a unique basis of the form

$$\begin{bmatrix} c_1 & 0 & \cdot\ \cdot\ \cdot & 0 \\ c_2 & c_1 & \cdot\ \cdot\ \cdot & 0 \\ \vdots & c_2 & \cdot\ \cdot\ \cdot & \cdot \\ c_{2M+1} & \vdots & \cdot\ \cdot\ \cdot & 0 \\ 0 & c_{2M+1} & \cdot\ \cdot & c_1 \\ 0 & 0 & & \cdot \\ \cdot & \cdot & \vdots\ \vdots\ \vdots & \vdots \\ 0 & 0 & & c_{2M+1} \end{bmatrix} \tag{A.12}$$

where $c_1 > 0$ and $\|\boldsymbol{C}\| = 1$ with $\boldsymbol{C} = (c_1, \ldots, c_{2M+1})$.

This implies that

$$\widehat{\boldsymbol{C}}^k \xrightarrow{\text{a.s.}} \boldsymbol{C} \quad \text{for} \quad k = 1, \ldots, L - 2M + 1 \tag{A.13}$$

Similar analysis for $\widetilde{\mathbf{T}}$ shows that

$$\widehat{\boldsymbol{C}}_k^* \xrightarrow{\text{a.s.}} \boldsymbol{C} \quad \text{for} \quad k = 1, \ldots, L - 2M + 1 \tag{A.14}$$

Thus we have

$$\widehat{\boldsymbol{C}}_k^{**} \xrightarrow{\text{a.s.}} \boldsymbol{C} \quad \text{for} \quad \boldsymbol{k = 1, \ldots, L - 2M + 1} \tag{A.15}$$

Therefore from Lemma 3 we can say that the roots obtained using $\widehat{\boldsymbol{C}}_k^{**}$ are consistent estimators of $\underset{\sim}{\omega}$'s for all $k = 1, \ldots, L - 2M + 1$.