ELSEVIER

Signal Processing 72 (1999) 129–139

SIGNAL PROCESSING

# On asymptotic behavior of least squares estimators and the confidence intervals of the superimposed exponential signals

Debasis Kundu*[,1], Amit Mitra[2]

*Department of Mathematics, I.I.T. Kanpur, Pin 208016, India*

Received 7 October 1994; received in revised form 7 September 1998

**Abstract**

The problem of estimation of the parameters of complex sinusoids in complex white noise arises in many fields. Although many papers appeared in the last twenty years dealing with the estimation of the complex sinusoids, not much attention has been paid to obtain the confidence intervals of the unknown parameters, without which the estimation may not make much sense. Recently Rao and Zhao (1993) obtained the asymptotic distribution of the least squares estimators of the frequencies of the complex sinusoidal model under the assumption of the Gaussian white noise, which can be used to obtain the confidence interval of the unknown parameters for finite sample. However, they did not perform any numerical study to see the validity of the asymptotic results for finite sample sizes. Moreover, in many situations it is observed that the error distributions need not be Gaussian. In this paper we consider the superimposed exponential model when the error distributions may not be Gaussian. We prove the strong consistency of the least squares estimators and derive the asymptotic distributions of the least squares estimators, which can be used to obtain the confidence interval of the unknown parameters. Finally some numerical experiments are performed to see how the asymptotic results behave for finite sample sizes and for different error distributions. © 1999 Elsevier Science B.V. All rights reserved.

**Zusammenfassung**

Das Problem der Schätzung von Parametern von komplexen Sinusoidalen in komplexem weißen Rauschen ergibt sich in vielen Anwendungsbereichen. Obwohl in den letzten zwanzig Jahren eine Vielzahl von Arbeiten erschienen sind, die sich mit der Schätzung komplexer sinusförmiger Signale beschäftigen, wurde der Bestimmung von Konfidenzintervallen für die unbekannten Parameter, ohne die die Parameterschätzung wenig Sinn macht, nicht viel Aufmerksamkeit gewidmet. Nicht zuletzt ermittelten Rao und Zhao (1993) die Konfidenzintervalle des Kleinste-Quadrate-Schätzers der Frequenzen einer komplexen, sinusoidalen Schwingung unter der Annahme, daß Gauss'sches, weißes Rauschen vorliegt.

* Corresponding author. Address for correspondence: Department of Statistics, Pennsylvania State University, State College, PA 16802-2111, USA, e-mail: kundu@stat.psu.edu.

[1] Part of the work of Debasis Kundu is supported by a grant No. ($SR/OY/M$ – 06/93) of the Department of Science and Technology, Government of India.

[2] Part of the work of Amit Mitra is supported by the National Board of Higher Mathematics, Department of Atomic Energy, Government of India. Presently at the Reserve Bank of India.



PII: S0165-1684(98)00171-6

Dies kann zur Bestimmung der Konfindenzintervalle für die unbekannten Parameter bei einer begrenzte Anzahl Abtastwerte herangezogen werden. Leider wurden an dieser Stelle keine numerischen Untersuchungen der Gültigkeit des asymptotischen Ergebnisses für eine begrenzte Anzahl von Abtastwerten durchgeführt. Mehr noch, in vielen Situationen läßt sich beobachten, daß die Störungen nicht Gauß-verteilt sein müssen. In dieser Arbeit betrachten wir das überlagerte Exponentialmodell, wenn die Fehler nicht Gauß-verteilt sind. Wir beweisen die strenge Konsistenz des Kleinste-Quadrate-Schätzers und leiten die asymptotische Verteilung für diesen Schätzer ab. Die asymptotische Verteilung kann dann zur Bestimmung der Konfindenzintervalle für die unbekannten Parameter herangezogen werden. Abschließend werden numerische Experimente durchgeführt, um zu sehen, wie sich die asymptotischen Resultate bei begrenzter Anzahl von Abstastwerten und für unterschiedliehe Fehlerwahrscheinlichkeiten auswirkt. 

**Résumé**

Le problème de l'estimation des paramètres de sinusoïdes complexes dans du bruit blanc complexe apparaît dans de nombreux domaines. Malgré la parution durant les vingt dernières années de nombreux articles traitant de l'estimation de sinusoïdes complexes, peu d'attention a été portée à l'obtention d'intervalles de confiance pour les paramètres estimés, sans lesquels l'estimation peut ne pas avoir beaucoup de sens. Récemment, Rao et Zhao (1993) ont obtenu une distribution asymptotique des estimateurs des moindres carrés des fréquences d'un modèle de sinusoïdes complexes sous l'hypothèse d'un bruit blanc gaussien, qui peut être utilisée pour obtenir l'intervalle de confiance des paramètres inconnus pour un échantillon fini. Cependant, ils n'ont fait aucune étude numérique pour voir la validité des résultats asymptotiques pour des tailles d'échantillons finis. De plus, dans beaucoup de situations, on peut observer que la distribution de l'erreur n'est pas gaussienne. Dans cet article, nous considerérons le modèle exponentiel superposé lorsque la distribution de l'erreur peut ne pas être gaussienne. Nous prouvons la forte consistance de l'estimateur des moindres carrés et dérivons la distribution asymptotique des estimateurs des moindres carrés, qui peut être utilisée pour obtenir l'intervalle de confiance des paramètres inconnus. Finalement, des essais numériques sont réalisés pour voir comment les résultats numériques se comportent pour des tailles d'échantillons finis et pour différentes distributions d'erreur. 

*Keywords:* Asymptotic covariance matrix; Cramer–Rao bound; Complex normal distributions; Exponential signals; Least squares estimators; Maximum likelihood estimators

## 1. Introduction

We consider the following model of multiple superimposed exponential signals in noise:

$$Y_t = \sum_{j=1}^{M} \alpha_j^0 \exp(it\beta_j^0) + \varepsilon_t, \tag{1}$$

where $i=\sqrt{-1}$, $\beta_j^0 \in [0,2\pi]$ is the radian frequency, $\alpha_j^0$ is its complex amplitude and $\{\varepsilon_t\}$ is a sequence of i.i.d. complex random variables with mean zero and finite variance. It is assumed that both the real and imaginary parts of $\varepsilon_t$ have mean zero and finite variance $\sigma^2/2 < \infty$ and they are independent. In general $\beta^0 = (\beta_1^0, \ldots, \beta_M^0)$, $\alpha^0 = (\alpha_1^0, \ldots, \alpha_M^0)$ and $\sigma^2$ are unknown. $\beta_1^0, \ldots, \beta_M^0$ are different and it is desired to estimate these unknown parameters. Here we assume that '$M$' is known.

This is an important problem in signal processing. For example in electromagnetic pulse (EMP) situations [12,13], the EMP pickup can be characterized by a sum of complex exponentials whose parameters are to be determined. The parameters are a means of coding the various pulse wave forms, and the signal approximation thus obtained can be readily employed to analyze responses in various subsystems under EMP environment. In system identification problems, the characterization of the impulse responses of a linear system by a sum of complex exponentials and then identifying or approximating the complex amplitudes and natural frequencies with high degree of accuracy has its special importance in a wide variety of applications.

Quite a number of papers appeared in the last twenty years (see for example [1,4,7,10,14]) dealing with this model, but nowhere, at least not known to the authors, the theoretical properties of the least squares estimators (LSEs) of the model (11) has been obtained under this general setup.

Observe that under the normality assumption on $\{\varepsilon_t\}$, the maximum likelihood estimators (MLEs) of $\alpha^0$ and $\beta^0$ are the same as the non-linear LSEs. Theoretically it is interesting to give a close study of the statistical properties of the LSEs. It is important to note that the model does not satisfy [8] the standard sufficient conditions of [3], [16] or [5], therefore although the least squares method usually gives the statistical procedure with satisfactory performance, the complexity of the model makes it unclear how good the LSEs are under the present situation. Bai et al. [2] first proved the strong consistency of the LSEs under the normality assumption of the error random variables. Rao and Zhao [11] obtained the asymptotic distribution of the LSE's under the same set of assumptions, which can be used to compute the confidence interval of the unknown parameters for finite samples. However they did not perform any numerical study to see how the asymptotic results behave for finite sample size. In the present paper the result of [11] have been generalized for more general class of distributions. It was observed [11] that the asymptotic covariance matrix of the LSEs of the regression parameters attains the Cramer–Rao lower bound under the assumption of normality on the error random variables. We observe that the same result is true even for a more general class of distributions, which can be used to obtain the confidence intervals of the unknown parameters when the errors are not necessarily Gaussian. Our approach is different from their approach. Interestingly we observe that arranging the parameters properly results a very convenient representaion of the asymptotic covariance matrix and some of the properties of the LSEs become more transparent in this representation. We perform some numerical experiments to see how the asymptotic results behave for different sample sizes and for different error distributions. It is observed that even for moderate sample sizes the asymptotic results work quite satisfactorily.

## 2. Consistency and asymptotic normality of the LSE

### 2.1. Consistency of the LSE

In this subsection we prove the consistency of the LSEs of the unknown parameters. We denote the real and imaginary part of $\alpha_j$ as $\alpha_{jr}$ and $\alpha_{jc}$, respectively, $\theta_j = (\alpha_{jr},\alpha_{jc},\beta_j)$ for $j = 1,\dots,M$, $\theta = (\theta_1,\dots,\theta_M)$. We denote the real and the imaginary part of $\varepsilon_t$ as $\varepsilon_{tr}$ and $\varepsilon_{tc}$, respectively. Similarly $\theta_j^0$, $\alpha_{jr}^0$, $\alpha_{jc}^0$, $\beta_j^0$ for $j = 1,\dots,M$ and $\theta^0$ are also defined. Now the LSEs can be obtained by minimizing

$$Q(\theta) = \sum_{t=1}^{n} \left| Y_t - \sum_{j=1}^{M} \alpha_j \exp(\mathrm{i}t\beta_j) \right|^2 \tag{2}$$

with respect to $\theta$. Let $\hat{\theta} = (\hat{\theta}_1,\dots,\hat{\theta}_M)$ be the LSEs of $\theta^0$. Consider the set

$$S_\delta = \{\theta : |\beta - \beta^0| \geqslant \delta \text{ or } |\alpha - \alpha^0| \geqslant \delta,\ \beta \in [0,2\pi)\}$$

with $\alpha = \mathrm{Re}(\alpha) + \mathrm{i}\,\mathrm{Im}(\alpha)$ for any $\delta > 0$.

Observe that, to prove $\hat{\theta}$ is a strongly consistent estimator of $\theta^0$, it is enough to prove that

$$\varliminf \inf_{S_\delta} \frac{1}{n}\{Q(\theta) - Q(\theta^0)\} > 0 \quad \text{a.s.} \tag{3}$$

for all $\delta > 0$ [16] (Lemma 1), here a.s. means almost surely. Now observe that

$$\frac{1}{n}\{Q(\theta) - Q(\theta^0)\} = f(\theta,\theta_0) + g(\theta,\theta_0),$$

where

$$f(\theta,\theta_0) = \frac{1}{n}\sum_{t=1}^{n}\Bigg[\Bigg\{\sum_{j=1}^{M} (\alpha_{jr}^0\cos(\beta_j^0 t) - \alpha_{jr}\cos(\beta_j t) - \alpha_{jc}^0\sin(\beta_j^0 t) + \alpha_{jc}\sin(\beta_j t))\Bigg\}^2 + \Bigg\{\sum_{j=1}^{M} (\alpha_{jc}^0\cos(\beta_j^0 t) - \alpha_{jc}\cos(\beta_j t) + \alpha_{jr}^0\sin(\beta_j^0 t) - \alpha_{jr}\sin(\beta_j t))\Bigg\}^2\Bigg]$$

and

$$g(\theta,\theta_0)=\frac{2}{n}\sum_{t=1}^{n}\Bigg[\varepsilon_{tr}\Bigg\{\sum_{j=1}^{M}(\alpha_{jr}^0\cos(\beta_j^0 t)-\alpha_{jr}\cos(\beta_j t) - \alpha_{jc}^0\sin(\beta_j^0 t)+\alpha_{jc}\sin(\beta_j t))\Bigg\} + \varepsilon_{tc}\Bigg\{\sum_{j=1}^{M}(\alpha_{jc}^0\cos(\beta_j^0 t)-\alpha_{jc}\cos(\beta_j t) + \alpha_{jr}^0\sin(\beta_j^0 t)-\alpha_{jr}\sin(\beta_j t))\Bigg\}\Bigg].$$

We show that

$$\underline{\lim}\inf_{S_\delta} f(\theta,\theta_0)>0 \quad \text{a.s.} \tag{4}$$

and

$$\lim_{n\to\infty}\sup_{S_\delta} g(\theta,\theta_0)=0 \quad \text{a.s.} \tag{5}$$

which shows Eq. (3). Now observe that Eq. (4) can be shown similarly as [6] and to prove Eq. (5) we need the following lemma. Note that the lemma is stronger than some of the already existing similar type of results in the literature see for example [6,11,15].

**Lemma 1.** *Let $X_1,X_2,\ldots$ be a sequence of i.i.d random variables with zero mean and $E|X_1|^2=\sigma^2<\infty$, then*

$$\lim_{n\to\infty}\sup_{0\leqslant\gamma\leqslant 2\pi}\left|\frac{1}{n}\sum_{t=1}^{n}X_t\cos(t\gamma)\right|\to 0 \quad \text{a.s.}$$

**Proof.** Let $Z_t=X_t.I_{[|X_t|\leqslant\sqrt{t}]}$, then $P\{Z_t\neq X_t i.o.\}=0$. Thus,

$$\sup_\gamma\frac{1}{n}\sum_{t=1}^{n}X_t\cos(t\gamma)\to 0 \quad \text{a.s.}$$

$$\equiv\sup_\gamma\frac{1}{n}\sum_{t=1}^{n}Z_t\cos(t\gamma)\to 0 \quad \text{a.s.}$$

Let $U_t=Z_t-EZ_t$. Note that

$$\sup_\gamma\left|\frac{1}{n}\sum_{t=1}^{n}E(Z_t\cos(t\gamma))\right|\leqslant\frac{1}{n}\sum_{t=1}^{n}|EZ_t| =\frac{1}{n}\sum_{t=1}^{n}\left|\int_{|x|\leqslant\sqrt{t}}x\,\mathrm{d}F(x)\right|\to 0. \tag{6}$$

Thus we only need to show that

$$\sup_{0\leqslant\gamma\leqslant 2\pi}\left|\frac{1}{n}\sum_{t=1}^{n}U_t\cos(t\gamma)\right|\to 0 \quad \text{a.s.}$$

for any fixed $\gamma$ and $\varepsilon>0$ and $0<h\leqslant 1/(4\sqrt{n})$, we have

$$P\left\{\left|\frac{1}{n}\sum_{t=1}^{n}U_t\cos(t\gamma)\right|\geqslant\varepsilon\right\} \leqslant 2\exp(-hn\varepsilon)E\exp\left(h\left|\sum_{t=1}^{n}U_t\cos(t\gamma)\right|\right) \leqslant 2\exp(-hn\varepsilon)\prod_{t=1}^{n}E\exp(|hU_t\cos(t\gamma)|) \leqslant 2\exp(-hn\varepsilon)\prod_{t=1}^{n}(1+h^2\sigma^2) \leqslant 2\exp(-hn\varepsilon+nh^2\sigma^2). \tag{7}$$

Note that $|hU_t\cos(t\gamma)|\leqslant\frac{1}{4}$ and on using the fact that $\mathrm{e}^{|x|}\leqslant 2\mathrm{e}^x$ and $\mathrm{e}^x\leqslant(1+x+x^2)$ for $|x|\leqslant\frac{1}{4}$, we have Eq. (7). Choose $h=1/(4\sqrt{n})$, then for large $n$

$$P\left\{\left|\frac{1}{n}\sum_{t=1}^{n}U_t\cos(t\gamma)\right|\geqslant\varepsilon\right\}\leqslant 2\exp\left\{-\frac{\sqrt{n}\varepsilon}{4}+\frac{\sigma^2}{16}\right\} \leqslant 2\exp\left\{-\frac{\sqrt{n}\varepsilon}{5}\right\}.$$

Let $k=n^2$, choose $\gamma_1,\gamma_2,\ldots,\gamma_k$ such that for each $\gamma\in[0,2\pi]$, we have a $\gamma_j$ satisfying $|\gamma_j-\gamma|\leqslant 2\pi/n^2$. Note that

$$\left|\frac{1}{n}\sum_{t=1}^{n}U_t(\cos(t\gamma)-\cos(t\gamma_j))\right| \leqslant\frac{1}{n}\sum_{t=1}^{n}\sqrt{t}t\frac{2\pi}{n^2}\leqslant\frac{2\pi}{\sqrt{n}}\to 0.$$

Therefore for large $n$, we have

$$P\left\{\sup_{\gamma}\left|\frac{1}{n}\sum_{t=1}^{n} U_t\cos(t\gamma)\right| \geqslant 2\varepsilon\right\}$$

$$\leqslant P\left\{\max_{j\leqslant n^2}\left|\frac{1}{n}\sum_{t=1}^{n} U_t\cos(t\gamma_j)\right| \geqslant \varepsilon\right\}$$

$$\leqslant 2n^2\exp\left\{-\frac{\sqrt{n}\varepsilon}{5}\right\}.$$

Since

$$\sum_{n=1}^{\infty} n^2\exp\left\{-\frac{\varepsilon}{5}\right\} < \infty$$

$$\sup_{\gamma}\left|\frac{1}{n}\sum_{t=1}^{n} U_t\cos(t\gamma)\right| \to 0 \quad \text{a.s.}$$

Similarly it can be shown that

$$\sup_{\gamma}\left|\frac{1}{n}\sum_{t=1}^{n} U_t\sin(t\gamma)\right| \to 0 \quad \text{a.s.}$$

Thus we have Lemma 1.

Lemma 1 immediately implies Eq. (5). Therefore we have the following result.

**Theorem 1.** *If* $\{\varepsilon_{tr}\}$ *and* $\{\varepsilon_{tc}\}$ *are independent sequences of random variables with mean zero and finite variance and* $\hat{\theta}$ *is the least squares estimator of* $\theta^0$ *of the model* (1) *then* $\hat{\theta}$ *converges to* $\theta^0$ *almost surely.*

Note that Theorem 1 is a very strong result, it means as the sample size increases the LSEs will be closer and closer to the true parameter value. Therefore the biases and the mean squared errors tend to zero as sample size increases. Rao and Zhao [11] proved this result but they proved this result under the assumptions that the error random variables are complex Gaussian, but here we have shown that the same result is true for a much larger class of error distributions. It may be mentioned at this point that although for the undamped model we get the consistency result for the LSEs but the same result is not true from the damped exponential model.

### *2.2. Asymptotic normality of the LSEs*

Although Theorem 1 says about the consistency of the LSEs but it does not say anything about the rate of convergence. Without the rate of convergence we do not have the idea how the estimators are converging to the true unknown parameters. In this subsection we address that problem. We prove that after proper normalization the distribution of the LSEs converges to a Gaussian random variables. It also helps us to form the asymptotic confidence band, without which the estimators may not make much sense. Consider the following notations:

$$Q'(\theta) = (Q'_1(\theta), \ldots, Q'_M(\theta)),$$

where

$$Q'_j(\theta) = \left(\frac{\partial Q(\theta)}{\partial \alpha_{jr}}, \frac{\partial Q(\theta)}{\partial \alpha_{jc}}, \frac{\partial Q(\theta)}{\partial \beta_j}\right)$$

for $j = 1, \ldots, M$, similarly let $Q''(\theta)$ be a $3M \times 3M$ matrix given by

$$Q''(\theta) = \frac{\partial^2 Q(\theta)}{\partial\theta\partial\theta^{\mathrm{T}}}.$$

Now expanding $Q''(\theta)$ around $\theta^0$ by multivariate Taylor series, we have

$$Q'(\hat{\theta}) - Q'(\theta^0) = (\hat{\theta} - \theta^0)Q''(\bar{\theta}), \tag{8}$$

where $\bar{\theta}$ is a point between $\hat{\theta}$ and $\theta^0$. Since $Q'(\theta) = 0$, Eq. (8) implies

$$(\hat{\theta} - \theta^0) = -Q'(\theta^0)[Q''(\bar{\theta})]^{-1}. \tag{9}$$

The main idea to prove that $(\hat{\theta} - \theta^0)$ converges to a normal distribution is as follows: consider the following $3M \times 3M$ diagonal matrix $\boldsymbol{D}$:

$$\boldsymbol{D} = \begin{bmatrix} \boldsymbol{D_1} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{D_2} & \boldsymbol{0} & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{D_M} \end{bmatrix}, \tag{10}$$

where each of the $\boldsymbol{D}_j$ is a $3\times 3$ diagonal matrix and $\boldsymbol{D}_j = \mathrm{diag}\{n^{-1/2}, n^{-1/2}, n^{-3/2}\}$, for $j = 1, \ldots, n$. Now consider

$$(\hat{\theta}-\theta_0)\boldsymbol{D}^{-1} = -Q'(\theta^0)\boldsymbol{D}[\boldsymbol{D}Q''(\bar{\theta})\boldsymbol{D}]^{-1}. \quad (11)$$

First we will show that $[Q'(\theta^0)\boldsymbol{D}]$ converges in distribution to a multivariate normal distribution and then we will show that $[\boldsymbol{D}Q''(\bar{\theta})\boldsymbol{D}]$ converges to a positive definite matrix almost surely and that will ensure the asymptotic normality of $(\hat{\theta}-\theta_0)\boldsymbol{D}^{-1}$ [9].

Now the elements of $Q'_j(\theta^0)$ are as follows:

$$\frac{\partial Q(\theta^0)}{\partial \alpha_{jr}} = -2\sum_{t=1}^{n} \varepsilon_{tr}\cos(\beta_j^0 t) - 2\sum_{t=1}^{n} \varepsilon_{tc}\sin(\beta_j^0 t) \quad (12)$$

for $j=1,\dots,M$.

$$\frac{\partial Q(\theta^0)}{\partial \alpha_{jc}} = 2\sum_{t=1}^{n} \varepsilon_{tr}\sin(\beta_j^0 t) - 2\sum_{t=1}^{n} \varepsilon_{tc}\cos(\beta_j^0 t) \quad (13)$$

for $j=1,\dots,M$ and

$$\frac{\partial Q(\theta^0)}{\partial \beta_j} = 2\sum_{t=1}^{n} t\varepsilon_{tr}(\alpha_{jr}^0\sin(\beta_j^0 t) + \alpha_{jc}^0\cos(\beta_j^0 t)) + 2\sum_{t=1}^{n} t\varepsilon_{tc}(\alpha_{jc}^0)(\sin(\beta_j^0 t) - \alpha_{jr}^0\cos(\beta_j^0 t)) \quad (14)$$

for $j=1,\dots,M$.

It can be easily seen that each element of $Q'_j(\theta^0)$ satisfies the Lindeberg Feller conditions. This along with the fact that $\{\varepsilon_{tr}\}$ and $\{\varepsilon_{tc}\}$ are independent sequences of random variables implies that

$$Q'_j(\theta^0)\boldsymbol{D}_j \xrightarrow{\mathscr{L}} N_3(0,\sigma^2\Sigma_j). \quad (15)$$

for $j=1,\dots,M$, here '$\xrightarrow{\mathscr{L}}$' means converges in distribution and $N_3(0,\sigma^2\Sigma_j)$ denotes the 3-variate normal distribution with mean vector zero and dispersion matrix $\sigma^2\Sigma_j$. Here $\Sigma_j$ is

$$\Sigma_j = \begin{bmatrix} 2 & 0 & -\alpha_{jc}^0 \\ 0 & 2 & \alpha_{jr}^0 \\ -\alpha_{jc}^0 & \alpha_{jr}^0 & \frac{2}{3}|\alpha_j^0|^2 \end{bmatrix} \quad (16)$$

for $j=1,\dots,M$. Note that $\sigma^2\Sigma_j$ is the limiting covariance matrix of $[Q'_j(\theta_0)\boldsymbol{D}_j]$. Also

$$Q'(\theta^0)\boldsymbol{D} \xrightarrow{\mathscr{L}} N_{3M}(0,\sigma^2\Sigma), \quad (17)$$

where $\Sigma$ is a $3M\times 3M$ matrix as given below:

$$\Sigma = \begin{bmatrix} \Sigma_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \Sigma_2 & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \Sigma_M \end{bmatrix}. \quad (18)$$

Observe that Eqs. (15) and (17) mainly follow from the facts:

$$\lim_{n\to\infty}\frac{1}{n}\sum_{n=1}^{\infty}\cos^2(\omega t) = \lim_{n\to\infty}\frac{1}{n}\sum_{n=1}^{\infty}\sin^2(\omega t) = \frac{1}{2}, \quad (19a)$$

$$\lim_{n\to\infty}\frac{1}{n}\sum_{n=1}^{\infty}\sin(\omega t)\cos(\omega t) = 0, \quad (19b)$$

$$\lim_{n\to\infty}\frac{1}{n^2}\sum_{n=1}^{\infty}t\cos^2(\omega t) = \lim_{n\to\infty}\frac{1}{n^2}\sum_{n=1}^{\infty}t\sin^2(\omega t) = \frac{1}{4}, \quad (19c)$$

$$\lim_{n\to\infty}\frac{1}{n^3}\sum_{n=1}^{\infty}t^2\cos^2(\omega t) = \lim_{n\to\infty}\frac{1}{n^3}\sum_{n=1}^{\infty}t^2\sin^2(\omega t) = \frac{1}{6}. \quad (19d)$$

Now using Eqs. (19a) and (19d) and the fact that $\bar{\theta}$ converges to $\theta^0$ almost surely, it can be shown that

$$\lim_{n\to\infty}\boldsymbol{D}Q''(\bar{\theta})\boldsymbol{D} = \lim_{n\to\infty}\boldsymbol{D}Q''(\theta^0)\boldsymbol{D} = \Sigma. \quad (20)$$

Therefore combining Eqs. (11) and (17) and (20), we have

$$(\hat{\theta}-\theta^0)\boldsymbol{D}^{-1} \xrightarrow{\mathscr{L}} N_{3M}(0,\sigma^2\Sigma^{-1}), \quad (21)$$

where

$$\Sigma^{-1} = \begin{bmatrix} \Sigma_1^{-1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \Sigma_2^{-1} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \Sigma_M^{-1} \end{bmatrix}, \quad (22)$$

$\Sigma_j^{-1}$ is the inverse matrix of $\Sigma_j$ and

$$\Sigma_j^{-1} = \begin{bmatrix} \frac{1}{2}+\frac{3}{2}\frac{\alpha_{jc}^{0\,2}}{|\alpha_j^0|^2} & -\frac{3}{2}\frac{\alpha_{jr}^0\alpha_{jc}^0}{|\alpha_j^0|^2} & 3\frac{\alpha_{jc}^0}{|\alpha_j^0|^2} \\ -\frac{3}{2}\frac{\alpha_{jr}^0\alpha_{jc}^0}{|\alpha_j^0|^2} & \frac{1}{2}+\frac{3}{2}\frac{\alpha_{jr}^{0\,2}}{|\alpha_j^0|^2} & -3\frac{\alpha_{jr}^0}{|\alpha_j^0|^2} \\ 3\frac{\alpha_{jc}^0}{|\alpha_j^0|^2} & -3\frac{\alpha_{jr}^0}{|\alpha_j^0|^2} & \frac{6}{|\alpha_j^0|^2} \end{bmatrix} \quad (23)$$

for $j = 1, \dots, M$. Therefore we have the following theorem.

**Theorem 2.** *Under the assumptions of Theorem* 1, $(\hat{\theta} - \theta^0)\boldsymbol{D}^{-1}$ *converges in distribution to a 3M-variate normal distribution, with mean vector zero and the covariance matrix* $\sigma^2\Sigma^{-1}$, *where* $\boldsymbol{D}$ *and* $\Sigma^{-1}$ *are as given above.*

Note that Theorem 2 gives us the idea of the rate of convergence of the LSEs. For example the LSEs of the amplitudes converge to the true value at the rate of $1/\sqrt{n}$, on the other hand the LSEs of the frequencies converge at the rate of $1/n^{3/2}$. Therefore the variance of the amplitude goes to zero much slower than that of the frequency. From the expression of $\Sigma^{-1}$ it is immediate that if the two frequencies are distinct then the corresponding LSEs will be asymptotically uncorrelated and the LSEs of the corresponding amplitudes will be also asymptotically uncorrelated. It was not that immediate from the expression of [11]. It is interesting to observe that $\Sigma^{-1}$ does not depend on the frequencies, it only depends on the amplitudes. Also the asymptotic variance of the LSEs of the frequencies are inversely proportional to the square of the corresponding amplitudes.

## 3. Consistency and asymptotic normality of $\hat{\sigma}^2$

In this section first we prove the consistency of $\hat{\sigma}^2$ as an estimator of $\sigma^2$, which is given by $\hat{\sigma}^2 = (1/n)Q(\hat{\theta})$. To prove this we need the following lemmas.

**Lemma 2.** *Let* $X_1, X_2, \dots$ *be a sequence of i.i.d. random variables with zero mean and* $E|X_i|^2 = \sigma^2 < \infty$ *for* $i = 1, 2, \dots$, *then*

$$\lim_{n\to\infty} \sup_{0\leqslant\gamma\leqslant 2\pi} \left| \frac{1}{n^2} \sum_{t=1}^{n} tX_t \cos(t\gamma) \right| \to 0 \quad \text{a.s.}$$

**Proof.** The proof follows similarly as the proof of Lemma 1 and hence it is omitted. □

**Lemma 3.** *If* $\hat{\beta}$ *is the LSE of* $\beta^0$ *for the model* (1), *then as* $n \to \infty$

$$n(\hat{\beta} - \beta^0) \to 0 \quad \text{a.s.}$$

**Proof.** Expanding $Q'(\hat{\theta})$ around $\theta^0$ by multivariate Taylor series and using Theorems 1 and 2 and Lemma 2, the result can be established.

Note that

$$\hat{\sigma}^2 = \frac{1}{n} Q_n(\hat{\theta}) = T_1 + T_2 + T_3,$$

where

$$T_1 = \frac{1}{n} \sum_{t=1}^{n} (\varepsilon_{tr}^2 + \varepsilon_{tc}^2) = \frac{1}{n} \sum_{t=1}^{n} |\varepsilon_t|^2,$$

$$\begin{aligned} T_2 = \frac{1}{n} \sum_{t=1}^{n} \Bigg[ & \Bigg\{ \sum_{j=1}^{M} (\alpha_{pr}^0 \cos(\beta_p^0 t) - \hat{\alpha}_{pr} \cos(\hat{\beta}_p t) \\ & - \alpha_{pc}^0 \sin(\beta_p^0 t) + \hat{\alpha}_{pc} \sin(\hat{\beta}_p t)) \Bigg\}^2 \\ & + \Bigg\{ \sum_{j=1}^{M} (\alpha_{pc}^0 \cos(\beta_p^0 t) - \hat{\alpha}_{pc} \cos(\hat{\beta}_p t) \\ & + \alpha_{pr}^0 \sin(\beta_p^0 t) - \hat{\alpha}_{pr} \sin(\hat{\beta}_p t)) \Bigg\}^2 \Bigg], \end{aligned}$$

$$\begin{aligned} T_3 = \frac{2}{n} \sum_{t=1}^{n} \Bigg[ & \varepsilon_{tr} \Bigg\{ \sum_{j=1}^{M} (\alpha_{pr}^0 \cos(\beta_p^0 t) - \hat{\alpha}_{pr} \cos(\hat{\beta}_p t) \\ & - \alpha_{pc}^0 \sin(\beta_p^0 t) + \hat{\alpha}_{pc} \sin(\hat{\beta}_p t)) \Bigg\} \\ & + \varepsilon_{tc} \Bigg\{ \sum_{j=1}^{M} (\alpha_{pc}^0 \cos(\beta_p^0 t) - \hat{\alpha}_{pc} \cos(\hat{\beta}_p t) \\ & + \alpha_{pr}^0 \sin(\beta_p^0 t) - \hat{\alpha}_{pr} \sin(\hat{\beta}_p t)) \Bigg\} \Bigg]. \end{aligned} \qquad (24)$$

Observe that $T_1$ converges to $\sigma^2$ a.s because of strong law of large numbers, $T_3$ converges to zero almost surely by Lemma 1. With the help of Theorem 1 and Lemma 3 it can be shown that $T_2$ also converges to zero almost surely. So in Eq. (24) we have $T_1 \to \sigma^2$ a.s., $T_2$ and $T_3 \to 0$ a.s. which shows that $\hat{\sigma}^2 = (1/n)Q(\hat{\theta}) \to \sigma^2$ a.s. We state the result in the following theorem.

**Theorem 3.** *If* $\hat{\theta}$ *is the LSE of* $\theta^0$ *for the model* (1) *and* $\{\varepsilon_{tr}\}$ *and* $\{\varepsilon_{tc}\}$ *satisfy the same conditions of Theorem* 1, *then* $\hat{\sigma}^2 = (1/n)Q(\hat{\theta})$ *is a strongly consistent estimator of* $\sigma^2$.

Theorem 3 implies, under the same assumptions of Theorem 1, it is not only the amplitudes or the frequencies, it is also possible to estimate the error variance quite accurately if the sample size is large. Again it does not say anything about the rate of convergence, which will be answered in the following subsection under somewhat stronger assumptions.

### 3.1. Asymptotic normality of $\hat{\sigma}^2$

In this subsection we obtain the asymptotic distribution of $\hat{\sigma}^2$. To obtain the asymptotic distribution of $\hat{\sigma}^2$, first we prove that

$$\left\{\sqrt{n}\hat{\sigma}^2 - \frac{1}{\sqrt{n}}\sum_{t=1}^{n}|\varepsilon_t|^2\right\} \xrightarrow{\mathscr{L}} 0. \tag{25}$$

Note that $\{(1/\sqrt{n})\sum_{t=1}^{n}|\varepsilon_t|^2\}$ with the proper moment assumptions (namely $E|\varepsilon_{tr}|^4 < \infty$ and $E|\varepsilon_{tc}|^4 < \infty$) converges to normal distribution because of central limit theorem (see [9]). Therefore Eq. (25) will ensure the asymptotic normality of $\sqrt{n}\hat{\sigma}^2$. Now $\sqrt{n}\hat{\sigma}^2$ can be written as

$$\sqrt{n}\hat{\sigma}^2 = G_1 + G_2 + G_3, \tag{26}$$

where

$$G_1 = \frac{1}{\sqrt{n}}Q_n(\hat{\theta}) = \frac{1}{\sqrt{n}}\sum_{t=1}^{n}(\varepsilon_{tr}^2 + \varepsilon_{tc}^2) = \frac{1}{\sqrt{n}}\sum_{t=1}^{n}|\varepsilon_t|^2,$$

$$G_2 = \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left[\left(\sum_{j=1}^{M}(\alpha_{jr}^0\cos(\beta_j^0 t) - \hat{\alpha}_{jr}\cos(\hat{\beta}_j t) - \alpha_{jc}^0\sin(\beta_j^0 t) + \hat{\alpha}_{jc}\sin(\hat{\beta}_j t))\right)^2 + \left(\sum_{j=1}^{M}\alpha_{jc}^0\cos(\beta_j^0 t) - \hat{\alpha}_{jc}\cos(\hat{\beta}_j t) + \alpha_{jr}^0\sin(\beta_j^0 t) - \hat{\alpha}_{jr}\sin(\hat{\beta}_j t)\right)^2\right],$$

$$G_3 = \frac{2}{\sqrt{n}}\sum_{t=1}^{n}\left[\varepsilon_{tr}\left(\sum_{j=1}^{M}(\alpha_{jr}^0\cos(\beta_j^0 t) - \hat{\alpha}_{jr}\cos(\hat{\beta}_j t) - \alpha_{jc}^0\sin(\beta_j^0 t) + \hat{\alpha}_{jc}\sin(\hat{\beta}_j t))\right) + \varepsilon_{tc}\left(\sum_{j=1}^{M}(\alpha_{jc}^0\cos(\beta_j^0 t) - \hat{\alpha}_{jc}\cos(\hat{\beta}_j t) + \alpha_{jr}^0\sin(\beta_j^0 t) - \hat{\alpha}_{jr}\sin(\hat{\beta}_j t))\right)\right]. \tag{27}$$

After some calculations it can be shown that $G_2$ and $G_3$ both converge to zero in distribution, which implies Eq. (25). Therefore we have the result.

**Theorem 4.** *If $\hat{\theta}_n$ is the LSE of $\theta^0$ for the model* (1) *and $\{\varepsilon_{tr}\}$ and $\{\varepsilon_{tc}\}$ are independent sequence of random variables with mean zero and finite fourth order moments and $\sigma^* = E(\varepsilon_{tr}^4) + E(\varepsilon_{tc}^4) - \sigma^4/2$, then*

$$\sqrt{n}(\hat{\sigma}^2 - \sigma^2) \xrightarrow{\mathscr{L}} N(0,\sigma^*).$$

Now if we assume that $\{\varepsilon_{tr}\}$ and $\{\varepsilon_{tc}\}$ are independent sequence of random variables each following $N(0,\sigma^2/2)$ then from Theorem 4, it follows that

$$\sqrt{n}(\hat{\sigma}^2 - \sigma^2) \xrightarrow{\mathscr{L}} N(0,\sigma^4). \tag{28}$$

Note that the above asymptotic distribution of $(\hat{\sigma}^2 - \sigma^2)$ is the same as that obtained in [11].

Similarly as Theorem 2, Theorem 4 tells us that the rate of convergence of $\hat{\sigma}^2$ is also $1/\sqrt{n}$. Therefore, if the sample size is large the asymptotic variance of $\hat{\sigma}^2$ will be proportional to $1/n$. Moreover the distribution will help us to obtain the confidence bound of the unknown variance for the general case.

## 4. Numerical experiments

In this section we perform some numerical experiments. We consider the following model:

$$Y_t = 2.0\exp(1.0it) + 3.0\exp(2.0it) + \varepsilon_t, \quad t = 1,\ldots,n. \tag{29}$$

here $\varepsilon_t$s are i.i.d. complex valued random variables with mean zero and variance 1/2 for both the real and the imaginary parts. The real and imaginary parts are independently and identically distributed. We consider the following four error distributions: (1) Gaussian, (2) uniform, (3) double exponential and (4) truncated Cauchy and four different sample sizes namely 25, 50, 75 and 100. For each data set we obtain the LSEs of $\alpha^0 = (2.0,3.0)$, $\beta^0 = (1.0,2.0)$ and $\sigma^2 = 1.0$ and also obtain the approximate 90% confidence limits for all the parameters (see for example Rao [9, p. 470], on how to compute the confidence interval if the distribution is known). The average LSEs (Estimates), the average mean squared errors (MSEs), the average length of the confidence intervals (Length) and the coverage probabilities (Coverage) are obtained over five hundred replications. It is observed that the mean of the LSEs is very close to the true parameter value in all the cases considered, therefore the MSEs and

Table 1
Gaussian error

| Sample size | 25 | 50 | 75 | 100 |
|---|---|---|---|---|
| Parameter | 1.00 | 1.00 | 1.00 | 1.00 |
| Estimates | 0.99980 | 0.99994 | 1.00001 | 1.00000 |
| ASVA | 9.6E − 4 | 1.2E − 4 | 3.6E − 6 | 1.5*E* − 6 |
| MSE | 1.375E − 3 | 1.986E − 4 | 3.998E − 6 | 1.675E − 6 |
| Length | 0.0433 | 0.0157 | 0.0079 | 0.0056 |
| Coverage | 0.885 | 0.887 | 0.911 | 0.911 |
| Sample size | 25 | 50 | 75 | 100 |
| Parameter | 2.00 | 2.00 | 2.00 | 2.00 |
| Estimates | 1.99994 | 2.00009 | 1.99999 | 2.00000 |
| ASVA | 4.27E − 5 | 5.3E − 6 | 1.6*E* − 6 | 7.0*E* − 7 |
| MSE | 6.231E − 5 | 6.783E − 6 | 2.107E − 6 | 8.962E − 7 |
| Length | 0.0371 | 0.0098 | 0.0055 | 0.0029 |
| Coverage | 0.918 | 0.894 | 0.914 | 0.907 |

Table 2
Uniform error

| Sample size | 25 | 50 | 75 | 100 |
|---|---|---|---|---|
| Parameter | 1.00 | 1.00 | 1.00 | 1.00 |
| Estimates | 0.99989 | 0.99997 | 1.00002 | 1.00001 |
| ASVA | 9.6E − 4 | 1.2E − 4 | 3.6E − 6 | 1.5E − 6 |
| MSE | 1.401E − 3 | 1.977E − 4 | 4.001E − 6 | 1.669E − 6 |
| Length | 0.0441 | 0.0169 | 0.0077 | 0.0049 |
| Coverage | 0.911 | 0.907 | 0.894 | 0.898 |
| Sample size | 25 | 50 | 75 | 100 |
| Parameter | 2.00 | 2.00 | 2.00 | 2.00 |
| Estimates | 2.00004 | 2.00011 | 1.99994 | 2.00002 |
| ASVA | 4.27E − 5 | 5.3E − 6 | 1.6E − 6 | 7.0E − 7 |
| MSE | 6.339E − 5 | 6.890E − 6 | 2.009E − 6 | 8.999E − 7 |
| Length | 0.0389 | 0.0107 | 0.0050 | 0.0033 |
| Coverage | 0.911 | 0.884 | 0.912 | 0.900 |

the variances of the LSEs are almost equal. We also obtain the asymptotic variances (ASVA) of the LSEs for comparison. The results of the frequencies are reported for all the error distributions and for different sample sizes along with their asymptotic variances in Tables 1–4. Since the results are similar in nature for the amplitudes as well as for $\sigma^2$, they are not reported here.

Some of the points are very clear from the numerical experiments, from the experiments it is immediate that the MSEs gradually decrease and approach the asymptotic variance, for all the error distributions as the sample size increases. It verifies the consistency of the LSEs and it also shows the validity of the asymptotic results even for moderate sample sizes. It is also clear that the average length

Table 3
Double exponential error

| Sample size | 25 | 50 | 75 | 100 |
|---|---|---|---|---|
| Parameter | 1.00 | 1.00 | 1.00 | 1.00 |
| Estimates | 0.99955 | 1.00001 | 1.00000 | 0.99998 |
| ASVA | 9.6E − 4 | 1.2E − 4 | 3.6E − 6 | 1.5E − 6 |
| MSE | 1.290E − 3 | 1.878E − 4 | 4.011E − 6 | 1.601E − 6 |
| Length | 0.0410 | 0.0161 | 0.0076 | 0.0049 |
| Coverage | 0.901 | 0.913 | 0.899 | 0.900 |
| Sample size | 25 | 50 | 75 | 100 |
| Parameter | 2.00 | 2.00 | 2.00 | 2.00 |
| Estimates | 1.99986 | 2.00008 | 1.99999 | 1.99999 |
| ASVA | 4.27E − 5 | 5.3E − 6 | 1.6E − 6 | 7.0E − 7 |
| MSE | 5.139E − 5 | 6.197E − 6 | 2.207E − 6 | 8.212E − 7 |
| Length | 0.0301 | 0.0082 | 0.0058 | 0.0033 |
| Coverage | 0.888 | 0.890 | 0.890 | 0.909 |

Table 4
Truncated cauchy error

| Sample size | 25 | 50 | 75 | 100 |
|---|---|---|---|---|
| Parameter | 1.00 | 1.00 | 1.00 | 1.00 |
| Estimates | 0.99801 | 1.00154 | 1.00029 | 1.00009 |
| ASVA | 9.6E − 4 | 1.2E − 4 | 3.6E − 6 | 1.5E − 6 |
| MSE | 1.883E − 3 | 2.011E − 4 | 4.997E − 6 | 2.001E − 6 |
| Length | 0.0498 | 0.0179 | 0.0098 | 0.0068 |
| Coverage | 0.912 | 0.889 | 0.911 | 0.909 |
| Sample size | 25 | 50 | 75 | 100 |
| Parameter | 2.00 | 2.00 | 2.00 | 2.00 |
| Estimates | 2.00011 | 1.99971 | 2.00091 | 2.00018 |
| ASVA | 4.27E − 5 | 5.3E − 6 | 1.6E − 6 | 7.0E − 7 |
| MSE | 6.498E − 5 | 7.009E − 6 | 2.998E − 6 | 9.178E − 7 |
| Length | 0.0401 | 0.0187 | 0.0059 | 0.0041 |
| Coverage | 0.910 | 0.910 | 0.896 | 0.905 |

of the confidence interval decreases as sample size increases. Since the performance is similar for all the error distributions, it is clear that the asymptotic results can be used to obtain the confidence bounds of the unknown parameters even for moderate sample sizes, not only for Gaussian but for a larger class of distributions. It is also observed that in all the cases the coverage probabilities are nearly 90%.

## 5. Conclusions

In this paper we proved the strong consistency of the LSEs of the parameters of the model (1) and also obtained the asymptotic distribution, under the assumptions that the errors are i.i.d. with mean zero and finite variance. Similar results have been obtained in [11] under the assumptions that the errors are i.i.d. and Gaussian random variables. We observe that the arrangement of the parameters results a very convenient form of the asymptotic covariance matrix. With the help of the asymptotic distribution we compute the confidence intervals for the unknown parameters for different error distributions. The numerical results suggest that the asymptotic results can be used even for moderate sample sizes.

## Acknowledgements

The authors would like to thank Professor Z. D. Bai for helping them to prepare this manuscript. Part of the work was done when the first author was visiting the Electrical Communications Engineering Department of Indian Institute of Science, Bangalore. The first author wishes to thank Professor V.U. Reddy for all his cooperation and help. The authors would also like to thank two referees for their valuable comments.

## References

[1] Z.D. Bai, C.R. Rao, M. Chow, An algorithm of estimation of superimposed exponential signals, in: Proc. Tencon 1989 4th IEEE Region 10, Internat. Conf. 1989, pp. 342–347.

[2] Z.D. Bai, X.R. Chen, P.R. Krishnaiah, Y.H. Wu, L.C. Zhao, Strong consistency of maximum likelihood parameter estimation of superimposed exponential signals in noise, Technical Report 87-17, Center for Multivariate Analysis, University of Pittsburgh, 1987.

[3] R.I. Jennrich, Asymptotic properties of nonlinear least squares estimators, Ann. Math. Statist. 40 (1969) 633–643.

[4] N. Kannan, Estimation of direction of arrival in signal processing models, Ph.D. Thesis, The Pennsylvania State University, 1992.

[5] D. Kundu, Asymptotic properties of the complex valued nonlinear regression model, Commun. Statist. Theory Methods A 20 (12) (1991) 3793–3803.

[6] D. Kundu, Asymptotic theory of least squares estimator of a particular nonlinear regression model, Statist. Probab. Lett. 18 (1993) 13–17.

[7] D. Kundu, Estimating the parameters of undamped exponential signals, Technometrics 35 (2) (1993) 215–218.

[8] D. Kundu, Consistency of the undamped exponential signals model for restricted parameter space, Commun. Statist. Theory Methods A 24 (1) (1995).

[9] C.R. Rao, Linear statistical inference and it's applications, 2nd. ed., Wiley, New York, 1973.

[10] C.R. Rao, Some results in signal processing, in: S.S. Gupta, J.O. Berger (Eds.), Statistical Decision Theory and Related Topics IV, Vol. 2, Springer, New York, 1988, pp. 319–332.

[11] C.R. Rao, L.C. Zhao, Asymptotic behavior of maximum likelihood estimates of superimposed exponential signals, IEEE Trans. Signal Process. 41 (1993) 1461–1463.

[12] L.W. Ricketts, J.E. Bridges, J. Miletta, EMP Radiation and Protection Techniques, Wiley, New York, 1976.

[13] P. Sircar, Accurate parameter estimation of damped sinusoidal signals samples at nonuniform spacings, Ph.D. Thesis, Syracuse University, 1987.

[14] P. Stoica, List of references on spectral line analysis, Signal Processing 31 (3) (1993) 329–340.

[15] A.M. Walker, On the estimation of a harmonic component in a time series with stationary independent residuals, Biometrika 58 (1971) 21–36.

[16] C.F.J. Wu, Asymptotic theory of nonlinear least squares estimation, Ann. Statist. 9 (1991) 501–513.