

AN EFFICIENT ALGORITHM FOR ESTIMATING THE PARAMETERS OF SUPERIMPOSED EXPONENTIAL SIGNALS IN MULTIPLICATIVE AND ADDITIVE NOISE

JIAWEN BIAN*, HUIMING PENG*, JING XING**, ZHIHUI LIU*, HONGWEI LI*

* School of Mathematics and Physics
China University of Geosciences, Wuhan, 430074, China
e-mail: hwli@cug.edu.cn

**Department of Statistics and Applied Mathematics
Hubei University of Economics, Wuhan, 430205, China

This paper considers parameter estimation of superimposed exponential signals in multiplicative and additive noise which are all independent and identically distributed. A modified Newton–Raphson algorithm is used to estimate the frequencies of the considered model, which is further used to estimate other linear parameters. It is proved that the modified Newton–Raphson algorithm is robust and the corresponding estimators of frequencies attain the same convergence rate with Least Squares Estimators (LSEs) under the same noise conditions, but it outperforms LSEs in terms of the mean squared errors. Finally, the effectiveness of the algorithm is verified by some numerical experiments.

Keywords: superimposed exponential signals, modified Newton–Raphson algorithm, multiplicative and additive noise, least squares estimators.

1. Introduction

We consider the following model of superimposed exponential signals in multiplicative and additive noise:

$$y(t) = \sum_{k=1}^p \xi_k(t) e^{i(\omega_k t + \varphi_k)} + \varepsilon(t),$$
$$t = 1, 2, \dots, N, \quad (1)$$

where $i = \sqrt{-1}$, ω_k 's and φ_k 's are unknown frequencies lying strictly between 0 and 2π , and ω_k 's are distinct. Multiplicative noise $\{\xi_k(t)\}$ is a sequence of independent and identically distributed (*i.i.d.*) real random variables with mean $\mu_k \neq 0$ and finite variance. Additive noise $\{\varepsilon(t)\}$ is a sequence of *i.i.d.* complex random variables with mean zero and finite variance $\sigma_0^2/2$ for both the real and imaginary parts that are assumed to be independent. The multiplicative and additive noise are mutually independent. The number of components p is assumed to be known in advance. Since the frequencies are the nonlinear parameter harder to be estimated compared with the linear parameters, in this paper we mainly consider the estimation of frequencies ω_j , given

a sample of size N , namely, $y(1), y(2), \dots, y(N)$. The linear parameters are estimated thereafter.

This is an important problem in time series analysis and statistical signal processing. It has been studied intensively in the last twenty years, when the signals are only contaminated by the additive noise or signals with constant amplitude. Many iterative and non-iterative procedures were developed to estimate the frequencies of the complex model or its corresponding real model very efficiently. It is known that the Least Squares Estimators (LSEs) are statistically efficient for the estimation of linear and nonlinear parameters (Jennrich, 1969) and have the best convergence rate for the frequencies and other linear parameters, which are $O_p(N^{-3/2})$ and $O_p(N^{-1/2})$, respectively (Kundu, 1997; Kundu and Miltra, 1999). Some of the iterative methods aimed at finding LSEs are the approaches of Bressler and Macovski (1986), Hwang and Chen (1993), Kannan and Kundu (1994), as well as Li and Stoica (1996). Some non-iterative methods, including those of Tufts and Kumaresan (1982), Roy and Kailath (1989), Quinn (1994), as well as Kundu and Mitra (1995), are robust and efficient for the estimation of the small sample size condition. There

are also some high-resolution methods, for example, the correlation based hybrid method (Zhang *et al.*, 1994), the cumulant based method (Sadler *et al.*, 1995; Zhang and Wang, 2000), or the linear prediction method (Chan and So, 2004), for estimating the frequencies. However, the convergence rate of the frequency estimator for these methods is usually lower than $O_p(N^{-3/2})$.

It is observed that multiplicative noise may also occur in a variety of applications (see the works of Van Trees (1971, Chapter 1), Dwyer (1991), Besson and Castanie (1993), Swami (1994), Giannakis and Zhou (1995), Ghogho *et al.* (1999) or Prasath (2011) and the references therein) or, in other words, the received signals may be subjected to random amplitude modulation, which can be attributed, but is not limited to, the fading of communication channels, illuminance and reflectance modeling in image processing, reflection from scintillating targets, and Doppler spreading caused by changing orientations of nonpoint targets. Several methods have been suggested to estimate the parameters of superimposed exponential signals in the presence of multiplication and additive noise, such as cyclic statistics methods (e.g., Giannakis and Zhou, 1995; Zhou and Giannakis, 1995; Ghogho *et al.*, 1999) and higher order spectra methods (e.g., Dwyer, 1991; Besson and Castanie, 1993; Swami, 1994; Zhou and Giannakis, 1994).

It is known that the noise has much influence on the performance of the algorithm (Gawron *et al.*, 2012). Cyclic statistics based estimators are asymptotic equivalent to LSEs (Zhou and Giannakis, 1995); however, the higher the order of the cyclic statistics, the larger the sample size required to obtain a given level of precision for parameter estimation while the low order cyclic statistics based estimator is less accurate. It is necessary to find a computationally efficient algorithm which is also equivalent to LSEs in terms of statistical performance. It was observed that general purpose iterative algorithms such as Gauss–Newton (Hartley, 1961; Koko, 2004) and Newton–Raphson (Bloomfield, 1976; Ypma, 1995) take a long time to converge to the true parameter value or converge to the local maxima or minima (Osborne and Smyth, 1995).

Recently, Bai *et al.* (2003) as well as Nandi and Kundu (2006) proposed a seven-step iterative algorithm and a three-step iterative algorithm to estimate the frequencies of superimposed signals in additive noise, respectively. It is shown that the two estimators both attain the convergence rate of LSEs and are both consistent and computationally efficient. The greatest advantage of these two algorithms lies in that they make full use of the inner relationship between the statistics for iteration and the parameters to raise the accuracy of the estimators iteratively, thus only finite and fixed steps are needed to converge.

Kundu *et al.* (2011) proposed a modified

Newton–Raphson algorithm to estimate the frequencies of superimposed exponential signals (real model) in additive noise, which was proved to be also LSE equivalent and computationally efficient. It is interesting to observe that the algorithm can be guaranteed to converge in finite iterative steps by proper step factor modification.

It is known that Asymptotic Least Squares Estimators (ALSEs) are equivalent to LSEs in the estimation of frequencies in the aspect of asymptotic distribution, under the additive noise condition. However, ALSEs are simpler and easier to be implemented in practice (Bai *et al.*, 2003; Nandi and Kundu, 2006). But to the best of the authors' knowledge nowhere, has the Newton–Raphson based iterative procedure for the ALSEs of the frequencies for a superimposed exponential model with both multiplicative and additive noise been considered. It is known that LSEs for the frequencies of superimposed signals in multiplicative and additive noise have the same convergence rate with that in only additive noise (Ghogho *et al.*, 1999; Peng *et al.*, 2009). However, it is more difficult to obtain an LSE or an ALSE equivalent estimator in multiplicative noise than in only additive noise. Stimulated by the works of Bai *et al.* (2003), Nandi and Kundu (2006) as well as Kundu *et al.* (2011), in this paper we employ a modified Newton–Raphson iterative procedure similar to that of Kundu *et al.* (2011) to estimate the frequencies of superimposed exponential signals in the presence of both multiplicative and additive noise. Our algorithm is different from that of Kundu *et al.* (2011) in the following aspects: firstly, we use an ALSE based criterion function while Kundu *et al.* (2011) used an LSE based criterion function; secondly, we consider the estimation in the presence of both multiplicative noise and additive noise while Kundu *et al.* (2011) considered the problem in the presence of only additive noise. Moreover, the algorithm proposed in this paper estimates the frequencies simultaneously while Kundu *et al.* (2011) estimated the frequencies one by one.

We use a modified step factor and an ALSE based criterion function which will be defined in Section 3 to estimate the frequencies of the model considered. It is observed that, if the initial estimator is accurate up to the order $O_p(N^{-1})$ (here $O_p(N^{-\delta})$ means $O_p(N^{-\delta})N^\delta$ is bounded in probability), then the modified Newton–Raphson iterative procedure produces a fully efficient frequency estimator with the convergence rate of $O_p(N^{-3/2})$. We use periodogram maximizers over Fourier frequencies as the initial estimator. It is known that periodogram maximizers over Fourier frequencies do not generally provide estimators up to the order $O_p(N^{-1})$ (Rice and Rosenblatt, 1988), whereas an initial estimator of the convergence rate of $O_p(N^{-(1-\delta)})$ ($\delta > 0$) is needed for most iterative techniques to work. To overcome this problem, we use a varying sample size technique (Nandi and Kundu, 2006; Bian *et al.*, 2009), i.e., we do not use the

fixed sample size available for estimation at each step. At the first step we use a fraction of it and at the last step we use the whole data set by gradually increasing the effective sample size.

The rest of the paper is organized as follows. In Section 2, we discuss the LSEs and ALSEs for the parameters of the model (1). The modified Newton–Raphson iterative procedure is presented in Section 3. In Section 4, we present numerical experiments and finally we conclude the paper in Section 5. All proofs are provided in Appendix.

2. Estimation principle

In this section, we shall review the two estimation principles of LSEs and ALSEs for the estimation of frequencies of the model (1) as follows.

2.1. Least squares estimators. The LSEs of the unknown parameters can be obtained by minimizing the residual sum of squares, namely,

$$R_N(\boldsymbol{\mu}, \boldsymbol{\omega}) = \sum_{t=1}^N \left| y(t) - \sum_{k=1}^p \mu_k e^{i(\omega_k t + \varphi_k)} \right|^2. \quad (2)$$

Note that $R_N(\boldsymbol{\mu}, \boldsymbol{\omega})$ can also be written as follows:

$$R_N(\boldsymbol{\mu}, \boldsymbol{\omega}) = [Y - X(\boldsymbol{\omega})\boldsymbol{\mu}]^H [Y - X(\boldsymbol{\omega})\boldsymbol{\mu}], \quad (3)$$

where

$$\begin{aligned} Y &= (y(1), y(2), \dots, y(N))^T, \\ \boldsymbol{\mu} &= (\mu_1 e^{i\varphi_1}, \mu_2 e^{i\varphi_2}, \dots, \mu_p e^{i\varphi_p})^T, \\ \boldsymbol{\omega} &= (\omega_1, \omega_2, \dots, \omega_p)^T \end{aligned}$$

and $X(\boldsymbol{\omega})$ is an $N \times p$ matrix, the k -th row and l -th column element of X being $e^{ik\omega_l}$. From (3), it is clear that $\boldsymbol{\mu}$ can be separated from $\boldsymbol{\omega}$. Observe that, for a fixed $\boldsymbol{\omega}$, the LSEs of $\boldsymbol{\mu}$ can be obtained as

$$\hat{\boldsymbol{\mu}}(\boldsymbol{\omega}) = [X(\boldsymbol{\omega})^H X(\boldsymbol{\omega})]^{-1} X(\boldsymbol{\omega})^H Y. \quad (4)$$

Now, if we substitute $\hat{\boldsymbol{\mu}}$ in (3) for its estimator in (4), we obtain

$$Q_N(\boldsymbol{\omega}) = R_N(\hat{\boldsymbol{\mu}}(\boldsymbol{\omega}), \boldsymbol{\omega}) = Y^H Y - Y^H P_X Y, \quad (5)$$

where $P_X = X(\boldsymbol{\omega})[X(\boldsymbol{\omega})^H X(\boldsymbol{\omega})]^{-1} X(\boldsymbol{\omega})^H$ is the projection matrix on the space spanned by the columns of $X(\boldsymbol{\omega})$. Therefore, the LSEs of $(\boldsymbol{\mu}, \boldsymbol{\omega})$ obtained by minimizing $R_N(\boldsymbol{\mu}, \boldsymbol{\omega})$ with respect to $(\boldsymbol{\mu}, \boldsymbol{\omega})$ are the same as the LSEs of $\boldsymbol{\omega}$ obtained first by minimizing $Q_N(\boldsymbol{\omega})$ with respect to $\boldsymbol{\omega}$ and then using the estimator of $\boldsymbol{\omega}$ in (5) to obtain the LSEs of $\boldsymbol{\mu}$. It is known (Peng *et al.*, 2009) that the LSEs of the frequencies have the following limiting distribution:

$$N^{3/2}(\hat{\boldsymbol{\omega}} - \boldsymbol{\omega}) \xrightarrow{\mathcal{L}} \mathcal{N}_p(\mathbf{0}, \boldsymbol{\Sigma}), \quad (6)$$

where

$$\boldsymbol{\Sigma} = \text{diag}\{6(\sigma^2 - \sigma_1^2)/\mu_1^2, 6(\sigma^2 - \sigma_2^2)/\mu_2^2, \dots, 6(\sigma^2 - \sigma_p^2)/\mu_p^2\},$$

and $\sigma^2 = \sum_{k=1}^p \sigma_k^2 + \sigma_0^2$, where $\xrightarrow{\mathcal{L}}$ means convergence in distribution and $\mathcal{N}_p(\mathbf{0}, \boldsymbol{\Sigma})$ denotes the p -variate normal distribution with zero mean and dispersion matrix $\boldsymbol{\Sigma}$.

2.2. Approximate least squares estimators. An alternative way to estimate the frequencies is to maximize the periodogram function $f_N(\omega)$, which can be defined as follows (Bian *et al.*, 2009; Kundu *et al.*, 2011):

$$f_N(\omega) = \frac{1}{N} \left| \sum_{t=1}^N y(t) e^{-i\omega t} \right|. \quad (7)$$

The estimators obtained by finding p local maxima of $f_N(\omega)$ achieve the best possible rate and are asymptotically equivalent to LSEs in the additive noise condition. This is known as ALSEs in the literature (Nandi and Kundu, 2006). Actually, it can be seen from (5) that minimizing $Q_N(\boldsymbol{\omega})$ with respect to $\boldsymbol{\omega}$ is equivalent to maximizing $Y^H X(\boldsymbol{\omega}) [X(\boldsymbol{\omega})^H X(\boldsymbol{\omega})]^{-1} X(\boldsymbol{\omega})^H Y$. If we note $X(\boldsymbol{\omega})$ as $(X(\omega_1), X(\omega_2), \dots, X(\omega_p))$, where $X(\omega_l) = (e^{i\omega_l}, e^{2i\omega_l}, \dots, e^{Ni\omega_l})^T$, then, for large N , maximizing $Y^H X(\boldsymbol{\omega}) [X(\boldsymbol{\omega})^H X(\boldsymbol{\omega})]^{-1} X(\boldsymbol{\omega})^H Y$ for $\boldsymbol{\omega}$ can be divided into maximizing the following terms for $l = 1, 2, \dots, p$, respectively:

$$\begin{aligned} I_N(\omega_l) &\triangleq Y^H X(\omega_l) [X(\omega_l)^H X(\omega_l)]^{-1} X(\omega_l)^H Y \\ &= \frac{1}{N} Y^H X(\omega_l) X(\omega_l)^H Y. \end{aligned} \quad (8)$$

It can be seen that maximizing (7) with respect to ω and maximizing (8) with respect to ω_l are equivalent, so the p local maxima of $I_N(\boldsymbol{\omega})$ can be seen as the ALSEs of frequencies for the model considered.

It is known that the optimal estimators can be obtained by maximizing (7) with infinitely dividing and searching the parameter space. However, it is so time consuming that it is hard to be implemented in practice (Nandi and Kundu, 2006). However, the periodogram based frequency estimators obtained under the condition that frequencies are Fourier frequencies can be easily obtained, although they provide estimators with the convergence rate of only $O_p(N^{-1})$ (Rice and Rosenblatt, 1988). So we shall use periodogram maximizers over Fourier frequencies as the initial estimators of the modified Newton–Raphson algorithm in the next section. Here a frequency is a Fourier frequency if it is of the form $\lambda = 2\pi k/N$, for some integer $1 \leq k \leq N$.

3. Modified Newton–Raphson algorithm

In this section, we shall discuss the modified Newton–Raphson algorithm based on the function $I_N(\omega_l)$

in (8), aiming to find a consistent estimator of ω_l in the model (1). Unlike in the traditional Newton–Raphson algorithm, we use a modified step factor 1/4 as in (9). Given a consistent estimator $\tilde{\omega}_l$ of ω_l in the model (1), we compute $\hat{\omega}_l$ for $l = 1, 2, \dots, p$ as follows:

$$\hat{\omega}_l = \tilde{\omega}_l - \frac{1}{4} \frac{I'_{N_m}(\tilde{\omega}_l)}{I''_{N_m}(\tilde{\omega}_l)}, \quad (9)$$

where $I_N(\tilde{\omega}_l)$ is defined as in (8) by substituting ω_l with $\tilde{\omega}_l$. $I'_N(\tilde{\omega}_l)$ and $I''_N(\tilde{\omega}_l)$ are the first and second derivatives of $I_N(\tilde{\omega}_l)$ with respect to $\tilde{\omega}_l$, respectively. We expect to improve the estimator of ω_l step by step using (9) from any consistent initial estimator $\tilde{\omega}_l$. The motivation of the algorithm is based on the following theorem.

Theorem 1. *If, for $l = 1, 2, \dots, p$, $\tilde{\omega}_l - \omega_l = O_p(N^{-1-\delta})$, where $\delta \in (0, 1/2]$, then*

- (a) $\hat{\omega}_l - \omega_l = O_p(N^{-1-3\delta})$ if $\delta \leq \frac{1}{6}$,
- (b) $N^{\frac{3}{2}}(\hat{\omega}_l - \omega_l) \xrightarrow{\mathcal{L}} \mathcal{N}_p(\mathbf{0}, \Sigma)$ if $\delta > \frac{1}{6}$,

where

$$\Sigma = \text{diag} \left\{ \frac{3(\sigma^2 - \sigma_1^2)}{2\mu_1^2}, \frac{3(\sigma^2 - \sigma_2^2)}{2\mu_2^2}, \dots, \frac{3(\sigma^2 - \sigma_p^2)}{2\mu_p^2} \right\},$$

$$\sigma^2 = \sum_{j=1}^p \sigma_j^2 + \sigma_0^2.$$

$\mathcal{N}_p(\mathbf{0}, \Sigma)$ denotes the p -dimensional normal distribution with zero mean and covariance matrix Σ .

Proof. See Appendix. ■

Remark 1. It is noted that the consistent estimator $\tilde{\omega}_l$ in (9), which is the estimator of ω_l before the iteration, should have a convergence rate of $O_p(N^{-1-\delta})(\delta > 0)$ so that $\hat{\omega}_l$ can be more accurate than $\tilde{\omega}_l$ after the iteration. It is observed from Theorem 1 that if the initial estimator of ω_l has a convergence rate of $O_p(N^{-1-\delta})(\delta > 0)$, then the iteration of (9) is completely feasible. Moreover, the small step factor 1/4 can prevent the algorithm from diverging or converging to a local minima according to Theorem 1.

The maximizers of the periodogram over Fourier frequencies can provide estimators for the frequencies in the model (1) with the convergence rate of $O_p(N^{-1})$ (Bian et al., 2009). However, the convergence rate of $O_p(N^{-1-\delta})(\delta > 0)$ is needed for a general purpose iterative algorithm to work, so we choose it as the initial estimator and use the technique of varying sample size as proposed by Nandi and Kundu (2006) at the same time. A detailed implementation of (9) is listed in the following. The m -th step estimator $\hat{\omega}_l^{(m)}$ is computed from the $(m - 1)$ -th step estimator $\hat{\omega}_l^{(m-1)}$ by the formula

$$\hat{\omega}_l^{(m)} = \hat{\omega}_l^{(m-1)} - \frac{1}{4} \frac{I'_{N_m}(\hat{\omega}_l^{(m-1)})}{I''_{N_m}(\hat{\omega}_l^{(m-1)})}, \quad (10)$$

where $I'_{N_m}(\tilde{\omega}_l)$ and $I''_{N_m}(\tilde{\omega}_l)$ can be obtained from (9) by replacing N and $\tilde{\omega}_l$ with N_m and $\hat{\omega}_l^{(m-1)}$, respectively, and a detailed expression can be found in the Appendix. We repeatedly choose suitable N_m at each step as follows:

- (i) Find the l -th argument maximum of $I_N(\omega)$ over Fourier frequencies, i.e., $\omega_l^{(0)}$ as the initial estimator of ω_l .
- (ii) Choose $N_1 = N^{6/7}$ and compute $\hat{\omega}_l^{(1)}$ from $\hat{\omega}_l^{(0)}$. Since $\hat{\omega}_l^{(0)} - \omega_l = O_p(N^{-1}) = O_p(N_1^{-1-\frac{1}{6}})$, using part (a) of Theorem 1, we have

$$\hat{\omega}_l^{(1)} - \omega_l = O_p(N_1^{-\frac{3}{2}}) = O_p(N^{-1-\frac{2}{7}}).$$

- (iii) Since $2/7 > 1/6$, choose $N_m = N(m \geq 2)$ and repeatedly use (10) until an optimal solution is obtained. Applying part (b) of Theorem 1, we have

$$N^{\frac{3}{2}}(\hat{\omega}^{(L)} - \omega) \xrightarrow{\mathcal{L}} \mathcal{N}_p(\mathbf{0}, \Sigma),$$

where L is the iteration number.

Therefore, it is observed that, if at any step the estimator is of the order $O_p(N^{-1-\delta})$, the method provides an estimator which improves the order to $O_p(N^{-1-3\delta})$ for $0 < \delta \leq \frac{1}{6}$ and, if $\frac{1}{6} < \delta \leq \frac{1}{2}$, then it provides the efficient estimator with convergence rate of $O_p(N^{-3/2})$ which is the same with LSEs (Ghoghho et al., 1999; Peng et al., 2009). Hence this method provides an efficient frequency estimator from the relatively poor initial estimator of the periodogram maximizer. Theoretically, only two steps are needed to attain the “optimal” estimator. However, more than two steps are needed to obtain a more accurate and robust estimator in practice. It can be seen from Theorem 1 that the asymptotic variance for the frequencies of the proposed estimators is just one fourth of the LSEs (Ghoghho et al., 1999; Peng et al., 2009).

The linear parameters of μ_l and φ_l ($l = 1, 2, \dots, p$) can be estimated based on the estimation of ω_l by (4) as follows:

$$\hat{\mu}_l = \left| \frac{1}{N} \sum_{t=1}^N y(t) e^{-i\omega_l t} \right|, \quad (11)$$

$$\hat{\varphi}_l = \arg \left\{ \frac{1}{N} \sum_{t=1}^N y(t) e^{-i\omega_l t} \right\}, \quad (12)$$

where $\arg(a)$ denotes the angle of a .

Remark 2. Note that we use a subset of $N_1 = N^{6/7}$ to begin the iterative process. Actually, the subset number is not unique. There are several other ways they can be chosen so that the iterative process will converge. For example, another set of choices can be $N_1 = N^{4/5}$ or $N_1 = N^{7/8}$. Basically, they are the same in the aspect of guaranteeing the application of Theorem 1 and the convergence of the algorithm.

Remark 3. If we denote by $\hat{\mu}(\omega)$ and $\hat{\varphi}(\omega)$, the final estimators of μ and φ respectively, then according to (11) we have

$$\hat{\mu}(\omega) - \mu(\omega) = \mu'(\omega^*) \odot (\hat{\omega} - \omega) + O_p(N^{-2}),$$

where $\varphi = (\varphi_1, \varphi_2, \dots, \varphi_p)^T$ and $\mu'(\omega^*)$ is the first derivative of $\mu(\omega^*)$ with respect to ω^* which is a point between $\hat{\omega}$ and ω . Thus

$$\begin{aligned} \text{Var}(\hat{\mu}(\omega)) &\approx \text{Var}(\hat{\omega}) \odot [\mu'(\omega^*)]^2 \\ &= \frac{1}{4} \text{Var}(\hat{\omega}_{LSE}) \odot [\mu'(\omega^*)]^2 \\ &= \frac{1}{4} \text{Var}(\hat{\mu}_{LSE}), \end{aligned} \quad (13)$$

where $\hat{\omega}_{LSE}$ and $\hat{\mu}_{LSE}$ are the LSEs of ω and μ , respectively, and \odot is the Hadamard product. The above equation shows that the variance of μ is one fourth of the variance of its LSEs and it is the same for φ .

4. Numerical experiment

In this section we present some numerical results to observe how the proposed method works for finite sample size. We consider the following two models:

Model 1:

$$y_1(t) = \xi_1(t)e^{i(0.5t + \frac{\pi}{4})} + \xi_2(t)e^{i(1.5t + \frac{\pi}{3})} + \varepsilon(t).$$

Model 2:

$$y_2(t) = \xi_1(t)e^{i(0.5t + \frac{\pi}{4})} + \xi_2(t)e^{i(0.57t + \frac{\pi}{3})} + \varepsilon(t).$$

In both cases, $\{\varepsilon(t)\}$ is taken as a sequence of *i.i.d.* Gaussian complex random variables with zero means, and both the real and the imaginary parts having finite variance $\sigma_0^2/2$. $\{\xi_1(t)\}$ and $\{\xi_2(t)\}$ are both sequences of Gaussian random variables with means 2, 2 and deviations 0.4, 0.6, respectively. To assess the sensitivity of the model for different noise levels, we consider three different σ_0 's, namely, $\sigma_0 = 0.5, 1.0$ and 1.5 . To examine the consistency of the estimator, we take the sample sizes as $N = 100, 200, 300, 400, 500$ and 1000 for Model 1, and take the sample sizes as $N = 200, 300, 400, 500, 600$ and 1000 for Model 2.

It is observed that the starting sample in Model 2 is larger than that in Model 1 for the two frequencies in Model 2 are closer than that in Model 1, so a larger sample is needed for the periodogram based initial estimator in Model 2 than in Model 1. We take the first frequency in Model 2 the same as that in Model 1, and take the second frequency in Model 2 closer to the first frequency in Model 2 so as to observe the resolution and statistical performance for the proposed algorithm when two frequencies are very close. In both cases, we generate the signal using the corresponding Models 1 and 2.

The initial estimators of the frequency parameters are obtained using the maximizers of the periodogram in (7), and are then improved by the iterative process in (9). The iterative process is terminated if the absolute difference between two consecutive iterations is less than σ^2/N^{-3} . In most cases, the iteration converges in about 30 times. The process is replicated 100 times. We report the average estimates (AEs) and the corresponding Mean Squared Errors (MSEs) of ω_1, ω_2 for Models 1 and 2, as well as the corresponding asymptotic variance of LSEs (ASVLs) and the corresponding asymptotic variances (AVs) as in part (b) of Theorem 1, in Tables 1–2 and 4–5, respectively. According to Remark 3, we expect that the MSEs of μ and φ may also be smaller than the corresponding MSEs of LSEs. To study this effect, we report the results of μ_1 in case of Model 1 in Table 3. A similar trend has been observed in the case of μ_2, φ and is not reported here.

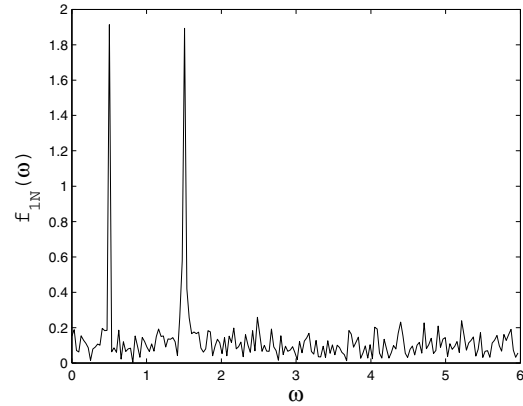


Fig. 1. Plot of the periodogram function of Model 1 with $N = 200$ and $\sigma_0 = 1.5$.

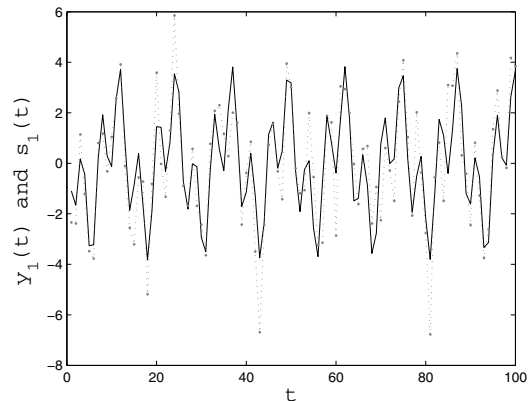


Fig. 2. Plot of the observed signal $y_1(t)$ (dashed line) and the noiseless signal $s_1(t)$ (solid line) using Model 1 with $N = 200$ and $\sigma_0 = 1.5$.

The following observations are very clear from the numerical results of Tables 1–5. It is observed that the AEs are very close to the true parameter values for most

Table 1. Average estimates (AE), mean squared errors (MSE), asymptotic variances of the modified Newton–Raphson algorithm (ASV) and asymptotic variances of the LSEs (ASVL) of the frequency ω_1 in Model 1.

	$N = 100$	$N = 200$	$N = 300$	$N = 400$	$N = 500$	$N = 1000$
$\sigma_0 = 0.5$						
AE	0.498908	0.499731	0.499893	0.499962	0.500015	0.500000
MSE	7.91543e-7	1.02935e-7	2.66455e-8	1.23431e-8	5.54595e-9	5.06825e-10
ASV	2.28750e-7	2.85936e-8	8.47222e-9	3.57422e-9	1.83000e-9	2.28750e-10
ASVL	9.15000e-7	1.14375e-7	3.38889e-8	1.42969e-8	7.32000e-9	9.15000e-10
$\sigma_0 = 1$						
AE	0.498794	0.499721	0.499897	0.499989	0.499988	0.500006
MSE	1.56290e-6	2.11714e-7	6.99356e-8	2.51921e-8	1.44951e-8	1.17149e-9
ASV	5.10003e-7	6.37500e-8	1.88889e-8	7.96875e-9	4.08000e-9	5.10000e-10
ASVL	2.04000e-6	2.55000e-7	7.55556e-8	3.18750e-8	1.632000e-8	2.04000e-9
$\sigma_0 = 1.5$						
AE	0.498966	0.499772	0.499865	0.499973	0.500012	0.499992
MSE	3.41193e-6	4.59768e-7	1.26025e-7	5.06914e-8	2.64905e-8	2.49135e-9
ASV	9.78750e-7	1.22344e-7	3.62450e-8	1.52930e-8	7.83000e-9	9.78750e-10
ASVL	3.91500e-6	4.89375e-7	1.44500e-7	6.11719e-8	3.13200e-8	3.91500e-9

Table 2. Average estimates (AE), mean squared errors (MSE), asymptotic variances of the modified Newton–Raphson algorithm (ASV) and asymptotic variances of the LSEs (ASVL) of the frequency ω_2 in Model 1.

	$N = 100$	$N = 200$	$N = 300$	$N = 400$	$N = 500$	$N = 1000$
$\sigma_0 = 0.5$						
AE	1.501048	1.500245	1.500088	1.500050	1.500016	1.500000
MSE	5.80705e-7	7.01068e-8	1.88964e-8	9.01031e-9	3.68061e-9	3.68990e-10
ASV	1.53750e-7	1.92188e-8	5.69444e-9	2.40234e-9	1.23000e-9	1.53750e-10
ASVL	6.15000e-7	7.68750e-8	2.27778e-8	9.60938e-9	4.92000e-9	6.15000e-10
$\sigma_0 = 1$						
AE	1.501014	1.500229	1.500144	1.500053	1.500023	1.500003
MSE	1.53017e-6	1.72904e-7	6.14974e-8	2.15401e-8	1.26851e-8	1.15089e-9
ASV	4.35000e-7	5.437501e-8	1.61111e-8	6.79688e-9	3.48000e-9	4.35000e-10
ASVL	1.74000e-6	2.17500e-7	6.44444e-8	2.71875e-8	1.39200e-8	1.74000e-9
$\sigma_0 = 1.5$						
AE	1.500659	1.500233	1.500092	1.500011	1.500008	1.500004
MSE	3.51063e-6	4.21503e-7	1.21324e-7	4.71154e-8	2.59425e-8	1.87135e-9
ASV	9.03750e-7	1.12969e-7	3.34722e-8	1.41211e-8	7.23000e-9	9.03750e-10
ASVL	3.61500e-6	4.51875e-7	1.33889e-7	5.64844e-8	2.89200e-8	3.61500e-9

of the cases. The biases and the MSEs decrease as the noise variance decreases or as the sample size increases. Therefore, the consistency of the proposed estimators is verified. Meanwhile, the MSEs of the proposed estimators are usually smaller than the asymptotic variances of the LSEs, which is observed in the case of the frequencies as well as other linear parameters. This verifies Remark 3 and the effectiveness of the proposed algorithm.

Comparing Table 1 with Table 4, it is observed that the performance of AEs and MSEs is similar for the estimation of the first frequency parameter while the estimation of the first frequency in Model 2 is less accurate and with little higher MSEs. It is not surprising that the two frequencies in Model 2 are closer than those in Model 1, and it is known that the estimation for dense frequencies is a more complex problem than that of the sparse frequencies condition.

For illustration purposes, we plot the periodogram function of the observed signals in Fig. 1 using Model 1 and in Fig. 5 using Model 2, respectively, when the additive noise deviation is 1.5 and the sample size is 200. It is known that the number of peaks in the periodogram function plot roughly gives an estimate of the number of frequencies. But it depends on the magnitude of the amplitude associated with each effective frequency and the noise variance. From Fig. 1, it is quite clear that there are two peaks while the two peaks in Fig. 5 are not very obvious as the two frequencies in Model 2 are very close. So the parameter estimation for Model 2 seems to be more difficult than that for Model 1.

We also plot the observed signal $y_1(t)$ of Model 1 along with the corresponding noiseless signal $s_1(t)$ (without multiplicative and additive noise) in Fig. 2 to see the influence of the noise on the exhibition of the

Table 3. Average estimates (AE), mean squared errors (MSE), asymptotic variances of the modified Newton–Raphson algorithm (ASV) and asymptotic variances of the LSEs (ASVL) of the mean of the first amplitude μ_1 in Model 1.

	$N = 100$	$N = 200$	$N = 300$	$N = 400$	$N = 500$	$N = 1000$
$\sigma_0 = 0.5$						
AE	2.001248	1.999835	1.999092	2.002730	2.002699	2.000524
MSE	2.56489e-3	1.20031e-3	8.88457e-4	6.89340e-4	5.66325e-2	1.67834e-4
ASV	7.12500e-4	3.56250e-4	2.37500e-4	1.78125e-4	1.42500e-4	7.12500e-5
ASVL	2.85000e-3	1.42500e-3	9.50000e-4	7.12500e-4	5.70000e-4	2.85000e-4
$\sigma_0 = 1$						
AE	2.004960	2.008402	1.997179	2.001021	1.997989	2.000600
MSE	5.54634e-3	2.72138e-3	2.09832e-2	1.49263e-3	1.18801e-3	5.52628e-4
ASV	1.65000e-3	8.25000e-4	5.50000e-4	4.12500e-4	1.32000e-3	1.65000e-4
ASVL	6.60000e-3	3.30000e-3	2.20000e-3	1.65000e-3	3.30000e-4	6.60000e-4
$\sigma_0 = 1.5$						
AE	2.015056	2.006714	2.000242	1.997307	2.007195	2.007386
MSE	1.06417e-2	6.25935e-3	4.02462e-3	3.17906e-3	1.88648e-3	1.18345e-3
ASV	3.21250e-3	1.60625e-3	1.07083e-3	8.03125e-4	6.42500e-4	3.21250e-4
ASVL	1.28500e-2	6.42500e-3	4.28333e-3	3.21250e-3	2.57000e-3	1.28500e-3

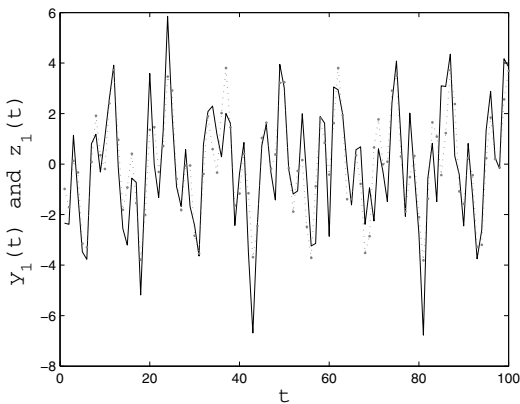


Fig. 3. Observed signals $y_1(t)$ (solid line) and the estimated signals $z_1(t)$ (dashed line) of Model 1 with $N = 200$ and $\sigma_0 = 1.5$.

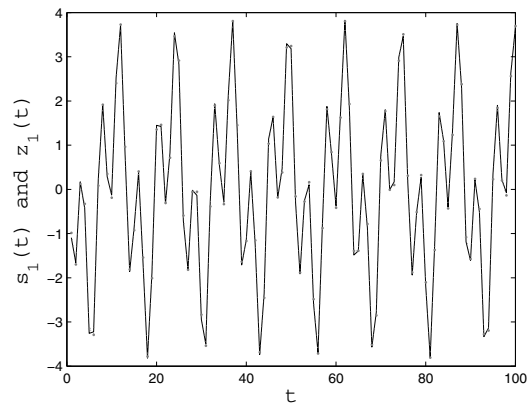


Fig. 4. Noiseless signals $s_1(t)$ (solid line) and the estimated signals $z_1(t)$ (dashed line) of Model 1 with $N = 200$ and $\sigma_0 = 1.5$.

signal. The noiseless signals are plotted by the true value of the parameters, i.e., ω , μ and φ , which are estimated by the modified Newton–Raphson algorithm, (11) and (12), respectively. It can be seen from Fig. 2 that the multiplicative and additive noise has much influence on the shape of the signal. To assess the performance of the proposed algorithm, we plot the estimated signal $z_1(t)$ for Model 1 along with the corresponding observed signal $y_1(t)$ as well as the estimated signal $z_1(t)$ along with the noiseless signal $s_1(t)$ when the additive noise variance is 1.5 and the sample size is 200, in Figs. 3 and 4, respectively. It can be seen from Figs. 3 and 4 that the estimated signal fits quite well the noiseless signal and not as well as with the observed signal, which verifies the effectiveness of the proposed algorithm and also further shows the influence of the multiplicative and additive noise on the signal. To illustrate the effectiveness

of estimation for Model 2, we plot the estimated signal $z_2(t)$ along with the corresponding noiseless signal $s_2(t)$ of Model 2 in Fig. 6. It is observed that the estimated signal fits quite well the noiseless signal, which further verifies the effectiveness of the proposed algorithm when two frequencies are very close to each other.

Finally, we must point out that the resolution of the algorithm is limited as the resolution of the algorithm is restricted to the periodogram based initial estimator, so this algorithm will fail when the distance of two frequencies is less than the frequency resolution. However, it is observed from Model 2 that the proposed algorithm works quite well when the distance of two frequencies is large and approaches the frequency resolution of the model considered. A better estimation can be expected to be achieved with a higher-resolution initial estimator or with a larger sample size when the

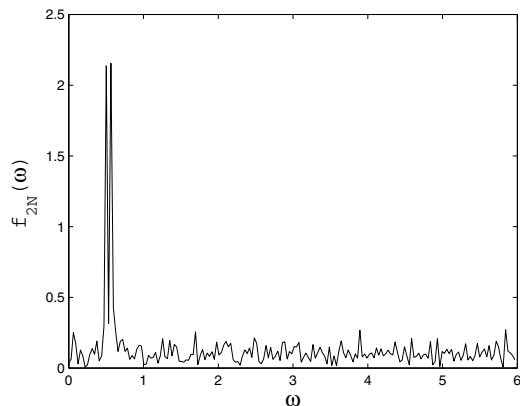


Fig. 5. Plot of the periodogram function of Model 2 with $N = 200$ and $\sigma_0 = 1.5$.

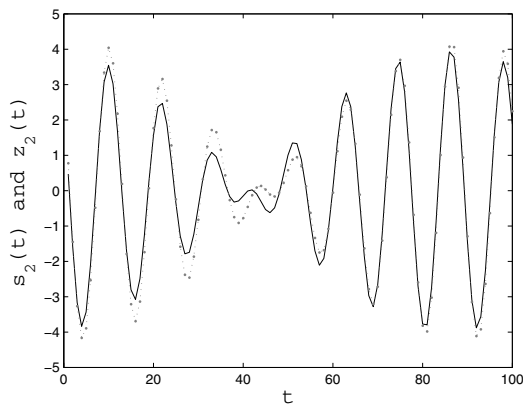


Fig. 6. Noiseless signals $s_2(t)$ (solid line) and the estimated signals $z_2(t)$ (dashed line) of Model 2 with $N = 200$ and $\sigma_0 = 1.5$.

frequencies are too close.

Remark 4. Comparing the experimental results with those of Kundu *et al.* (2011), more iterations are needed for the algorithm to work. However, lower computation cost is required since the computation of the inverse matrix as in the work of Kundu *et al.* (2011) is avoided. Moreover, the statistical performance for ALSE based parameter estimation in this paper does not seem to decrease compared with LSE based parameter estimation of Kundu *et al.* (2011).

5. Conclusions

In this paper, we mainly considered the estimation of the frequencies of a superimposed exponential signal model. We generalized the modified Newton–Raphson algorithm (Kundu *et al.*, 2011) from the additive noise condition to the multiplicative and additive noise condition. Unlike in the work of Kundu *et al.* (2011), our algorithm is based on the ALSE criterion which is computationally simpler than

the LSE criterion based algorithm of Kundu *et al.* (2011), since the computation of an inverse matrix as in the work of Kundu *et al.* (2011) is avoided. Moreover, the proposed algorithm can be used to estimate the frequencies of the model considered simultaneously while Kundu *et al.* (2011) estimated the frequencies one by one.

The asymptotic distribution of the proposed estimators is obtained. It is observed from the asymptotic distribution matrix that the proposed estimators attain the same convergence rate with the LSEs and outperform them in terms of asymptotic variance. Simulation experiments are performed to verify the effectiveness of the algorithm, especially for the estimation of the model with close space frequencies, which is expected to be improved with a higher-resolution initial estimator or larger sample size. Finally, it can also be seen that the modified Newton–Raphson algorithm converges quickly from the initial estimator of periodogram maximizers over Fourier frequencies, it naturally saves computational time and can be used for online implementation.

Acknowledgment

This work was supported in part by the National Natural Science Foundation of China under Grants 61071188, 11126274 and 61102103, the Natural Science Foundation of Hubei Province, China, under Grants 2011CDB333 and 2010CDB04205, and the Special Fund for Basic Scientific Research of Central Colleges, China University of Geosciences (Wuhan) (Nos. CUGL100239 and CUGL100236).

References

- Bai, Z.D., Rao, C.R., Chow M. and Kundu, D. (2003). An efficient algorithm for estimating the parameters of superimposed exponential signals, *Journal of Statistical Planning and Inference* **110**(1–2): 23–34.
- Besson, B. and Castanie, F. (1993). On estimating the frequency of a sinusoid in auto-regressive multiplicative noise, *Signal Processing* **30**(1): 65–83.
- Bian, J., Li, H. and Peng, H. (2009). An efficient and fast algorithm for estimating the frequencies of superimposed exponential signals in zero-mean multiplicative and additive noise, *Journal of Statistical Computation and Simulation* **74**(12): 1407–1423.
- Bian, J., Li, H. and Peng, H. (2009). An efficient and fast algorithm for estimating the frequencies of superimposed exponential signals in multiplicative and additive noise, *Journal of Information and Computational Science* **6**(4): 1785–1797.
- Bloomfield, P. (1976). *Fourier Analysis of Time Series: An Introduction*, Wiley, New York, NY.
- Bressler, Y. and MaCovski, A. (1986). Exact maximum likelihood parameters estimation of superimposed

Table 4. Average estimates (AE), the mean squared errors (MSE), asymptotic variances of the modified Newton–Raphson algorithm (ASV) and asymptotic variances of the LSEs (ASVL) of the mean of the first amplitude ω_1 in Model 2.

	$N = 200$	$N = 300$	$N = 400$	$N = 500$	$N = 600$	$N = 1000$
$\sigma_0 = 0.5$						
AE	0.498721	0.499778	0.499820	0.499922	0.499988	0.500005
MSE	8.81628e-8	3.76099e-8	1.37593e-8	8.58659e-9	4.99198e-9	7.43070e-10
ASV	2.85938e-8	8.47222e-9	3.57422e-9	1.83000e-9	1.05903e-9	2.28750e-10
ASVL	1.14375e-7	3.38889e-8	1.42969e-8	7.32000e-9	4.23611e-9	9.15000e-10
$\sigma_0 = 1$						
AE	0.498598	0.499769	0.499803	0.499885	0.499933	0.499981
MSE	2.15727e-7	6.68830e-8	2.61861e-8	2.15055e-8	1.11643e-8	1.57380e-9
ASV	6.37500e-8	1.88889e-8	7.96875e-9	4.08000e-9	2.36111e-9	5.10000e-10
ASVL	2.55000e-7	7.55556e-8	3.18750e-8	1.63200e-8	9.44444e-9	2.04000e-9
$\sigma_0 = 1.5$						
AE	0.498543	0.499758	0.499789	0.499883	0.499906	0.500029
MSE	3.62122e-7	1.18962e-7	6.88733e-8	3.20848e-8	1.92016e-8	3.18865e-9
ASV	1.22344e-7	3.62450e-8	1.52930e-8	7.83000e-9	4.53125e-9	9.78750e-10
ASVL	4.89375e-7	1.45000e-7	6.11719e-8	3.13200e-8	1.81250e-8	3.51900e-9

exponential signals in noise, *IEEE Transactions on Signal Processing* **34**(5): 1081–1089.

Chan, K.W. and So, H.C. (2004). Accurate frequency estimation for real harmonic sinusoids, *IEEE Signal Processing Letters* **11**(7): 609–612.

Dwyer, R.F. (1991). Fourth-order spectra of Gaussian amplitude modulated sinusoids, *Journal of the Acoustical Society of America* **90**(2): 918–926.

Fuller, W.A. (1996). *Introduction to Statistical Time Series*, 2nd Edn., Wiley, New York, NY.

Gawron, P., Klamka, J. and Winiarczyk, R. (2012). Noise effects in the quantum search algorithm from the viewpoint of computational complexity, *International Journal of Applied Mathematics and Computer Science* **22**(2): 493–499, DOI: 10.2478/v10006-012-0037-2.

Ghogho, M., Swami, A. and Garell, B. (1999). Performance analysis of cyclic statistics for the estimation of harmonics in multiplicative and additive noise, *IEEE Transactions on Signal Processing* **47**(12): 3235–3249.

Ghogho, M., Swami, A. and Nandi, A.K. (1999). Non-linear least squares estimation for harmonics in multiplicative and additive noise, *Signal Processing* **78**(1): 43–60.

Giannakis, G.B. and Zhou, G. (1995). Harmonics in multiplicative and additive noise: Parameter estimation using cyclic statistics, *IEEE Transactions on Signal Processing* **43**(9): 2217–2221.

Hartley, H.O. (1961). The modified Gauss–Newton method for the fitting of non-linear regression functions by least squares, *Technometrics* **3**(2): 269–280.

Hwang, J.K. and Chen, Y.C. (1993). A combined detection-estimation algorithm for the harmonic-retrieval problem, *Signal Processing* **30**(2): 177–197.

Jennrich, R.I. (1969). Asymptotic properties of non-linear least squares estimators, *The Annals of Mathematical Statistics* **40**(2): 633–643.

Kannan, N. and Kundu, D. (1994). On modified EVLP and ML methods for estimating superimposed exponential signals, *Signal Processing* **39**(3): 223–233.

Koko, J. (2004). Newton’s iteration with a conjugate gradient based decomposition method for an elliptic PDE with a nonlinear boundary condition, *International Journal of Applied Mathematics and Computer Science* **14**(1): 13–18.

Kundu, D. (1997). Asymptotic theory of the least squares estimators of sinusoidal signals, *Statistics* **30**(3): 221–238.

Kundu, D., Bai, Z., Nandi, S. and Bai, L. (2011). Super efficient frequency estimation, *Journal of Statistical Planning and Inference* **141**(8): 2576–2588.

Kundu, D. and Mitra, A. (1995). Consistent method of estimating superimposed exponential signals, *Scandinavian Journal of Statistics* **22**(1): 73–82.

Kundu, D. and Mitra, A. (1999). On asymptotic behavior of the least squares estimators and the confidence intervals of the superimposed exponential signals, *Signal Processing* **72**(2): 129–139.

Li, J. and Stoica, P. (1996). Efficient mixed-spectrum estimation with applications to target feature extraction, *IEEE Transactions on Signal Processing* **44**(2): 281–295.

Mangulis, V. (1965). *Handbook of Series for Scientists and Engineers*, Academic Press, New York, NY.

Nandi, S. and Kundu, D. (2006). An efficient and fast algorithm for estimating the parameters of sinusoidal signals, *Sankhya* **68**(2): 283–306.

Osborne, M.R. and Smyth, G.K. (1995). A modified Prony algorithm for fitting sum of exponential functions, *SIAM Journal on Scientific and Statistical Computing* **16**(1): 119–138.

Peng, H., Li, H. and Bian, J. (2009). Asymptotic behavior of least squares estimators for harmonics in multiplicative and additive noise, *Journal of Information and Computational Science* **6**(4): 1847–1860.

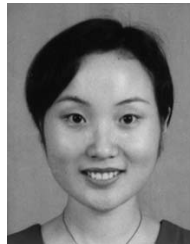
- Prasath, V.B.S. (2011). A well-posed multiscale regularization scheme for digital image denoise, *International Journal of Applied Mathematics and Computer Science* **21**(4): 769–777, DOI: 10.2478/v10006-011-0061-7.
- Quinn, B.G. (1994). Estimating frequency by interpolation using Fourier coefficients, *IEEE Transactions on Signal Processing* **42**(5): 1264–1268.
- Rice, J. A. and Rosenblatt, M. (1988). On frequency estimation, *Biometrika* **75**(3): 477–484.
- Roy, R. and Kailath, T. (1989). ESPRIT: Estimation of signal parameters via rotational invariance techniques, *IEEE Transactions on Signal Processing* **37**(7): 984–995.
- Sadler, B., Giannakis, G. and Shamsunder, S. (1995). Noise subspace techniques in non-Gaussian noise using cumulants, *IEEE Transactions on Aerospace and Electronic Systems* **31**(3): 1009–1018.
- Swami, A. (1994). Multiplicative noise models: Parameter estimation using cumulants, *Signal Processing* **36**(3): 355–373.
- Tufts, D.W. and Kumaresan, R. (1982). Estimation of frequencies of multiple sinusoids: Making linear prediction perform like maximum likelihood, *Proceedings of IEEE* **70**(9): 975–989.
- Van Trees, H.L. (1971). *Detection, Estimation and Modulation Theory, Part III: Radar-Sonar Signal Processing and Gaussian Signals in Noise*, Wiley, New York, NY.
- Ypma, T.J. (1995). Historical development of the Newton–Raphson method, *SIAM Review* **37**(4): 531–551.
- Zhang, X.D., Liang, Y.C. and Li, Y.D. (1994). A hybrid approach to harmonic retrieval in non-Gaussian ARMA noise, *IEEE Transactions on Information Theory* **40**(7): 1220–1226.
- Zhang, Y. and Wang, S.X. (2000). Harmonic retrieval in colored non-Gaussian noise using cumulants, *IEEE Transactions on Signal Processing* **48**(3): 982–987.
- Zhou, G. and Giannakis, G.B. (1994). On estimating random amplitude modulated harmonics using higher-order spectra, *IEEE Journal of Oceanic Engineering* **19**(4): 529–539.
- Zhou, G. and Giannakis, G.B. (1995). Harmonics in multiplicative and additive noise: Performance analysis of cyclic estimators, *IEEE Transactions on Signal Processing* **43**(6): 1445–1460.



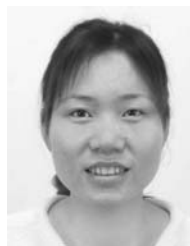
Jiawen Bian received his M.Sc. degree in applied mathematics in 2004 from Hubei University and a Ph.D. degree in earth exploration and information technology in 2010 from the China University of Geosciences. He is currently a lecturer at the School of Mathematics and Physics, China University of Geosciences, Wuhan. His research interests include statistical signal processing and time series analysis.



Huiming Peng was born in Hunan, China. He received his M.Sc. degree in probability and statistics in 2005 from Wuhan University and Ph.D. degree in earth exploration and information technology in 2011 from the China University of Geosciences. His research interests include statistical signal processing and bioinformatics.



Jing Xing received the M.Sc. degree from the School of Mathematics and Statistics, Huazhong University of Sciences and Technology, Wuhan, China, in 2008. She is currently pursuing the Ph.D. degree at the Institute of Geophysics and Geomatics, China University of Geosciences, Wuhan. She is a lecturer at the Department of Mathematics and Statistics, Hubei University of Economics, Wuhan. Her research interest is statistical signal processing.



Zhihui Liu received the M.Sc. degree from the School of Mathematics and Statistics, Wuhan University, China, in 2005. She is currently pursuing the Ph.D. degree at the Institute of Geophysics and Geomatics, China University of Geosciences, Wuhan. She is a lecturer at the School of Mathematics and Physics, China University of Geosciences. Her research interests include statistical signal processing and time series analysis.



Hongwei Li received the Ph.D. degree in applied mathematics from Peking University, Beijing, China, in 1996. From 1996 to 1998, he was a postdoctoral fellow at the Institute of Information Science, Beijing Jiaotong University. Since 1999, he has been a professor at the School of Mathematics and Physics, China University of Geosciences, Wuhan. His research interests include statistical signal processing, blind signal processing, multidimensional signal processing, pattern recognition and time series analysis.

Appendix

Proof of Theorem 1

We need the following two lemmas for the proof of Theorem 1.

Lemma A1. For all $k \in \mathbb{N}$ and $\omega \in \mathbb{R}$, we have

$$\sum_{t=1}^N t^k e^{i\omega t} = \sum_{t=1}^N t^k e^{i\omega t} = \begin{cases} O(N^k) & \text{if } \omega \neq 0, \\ O(N^{k+1}) & \text{if } \omega = 0. \end{cases}$$

Proof. See the work of Mangulis (1965). ■

Lemma A2. If $\{e(t)\}$ is a sequence of i.i.d. random variables with zero mean and finite variance, then

$$\sum_{t=1}^N e(t) t^k e^{i\theta t} = O_p(N^{k+\frac{1}{2}}), \quad \forall \theta \in \mathbb{R}. \quad a.s.,$$

Table 5. Average estimates (AE), mean squared errors (MSE), asymptotic variances of the modified Newton–Raphson algorithm (ASV) and asymptotic variances of the LSEs (ASVL) of the mean of the first amplitude ω_2 in Model 2.

	$N = 200$	$N = 300$	$N = 400$	$N = 500$	$N = 600$	$N = 1000$
$\sigma_0 = 0.5$						
AE	0.571359	0.570112	0.569981	0.570083	0.570008	0.570006
MSE	7.20544e-8	1.85488e-8	1.08902e-8	4.13682e-9	3.10130e-9	5.64307e-10
ASV	1.92188e-8	5.69444e-9	2.40234e-9	1.23000e-9	7.11806e-10	1.53750e-10
ASVL	7.68750e-8	2.27778e-8	9.60938e-9	4.92000e-9	2.84722e-9	6.15000e-10
$\sigma_0 = 1$						
AE	0.571430	0.570204	0.569986	0.570092	0.570016	0.570013
MSE	1.55689e-7	6.06032e-8	3.76675e-8	1.60800e-8	9.20212e-9	1.39253e-9
ASV	5.43750e-8	1.61111e-8	6.79688e-9	3.48000e-9	2.01389e-9	4.35000e-10
ASVL	2.17500e-7	6.44444e-8	2.71875e-8	1.39200e-8	8.05556e-9	1.74000e-9
$\sigma_0 = 1.5$						
AE	0.571463	0.570213	0.570772	0.570099	0.570030	0.570022
MSE	2.26359e-7	1.60292e-7	6.46088e-8	4.08008e-8	2.214789e-8	3.13654e-9
ASV	1.12969e-7	3.34722e-8	1.41211e-8	7.23000e-9	4.184028e-9	9.03750e-10
ASVL	4.51875e-7	1.45000e-7	5.64844e-8	3.13200e-8	1.673611e-8	3.61500e-9

where a.s. denotes ‘almost surely’.

Proof. See the work of Bai *et al.* (2003). ■

Since $\{\xi_k(t)\}$ is a sequence of *i.i.d.* random variables with mean u_k and variance σ_k^2 , if we write $\xi_k(t) - \mu_k$ as $\eta_k(t)$, then $\{\eta_k(t)\}$ is a sequence of *i.i.d.* random variables with zero mean and variance σ_k^2 , so we have $\xi_k(t) = u_k + \eta_k(t)$. If we write

$$\begin{aligned} X &\triangleq X(\tilde{\omega}_l) = (e^{i\tilde{\omega}_l}, e^{2i\tilde{\omega}_l}, \dots, e^{Ni\tilde{\omega}_l})^T, \\ Y &\triangleq (y(1), y(2), \dots, y(N))^T, \\ D &\triangleq \text{diag}\{1, 2, \dots, N\}, \end{aligned}$$

then

$$\dot{X} \triangleq \frac{dX}{d\tilde{\omega}_l} = iDX, \quad \ddot{X} \triangleq \frac{d^2X}{d\tilde{\omega}_l^2} = -D^2X, \quad (\text{A1})$$

where ‘T’ denotes the transpose, $I'_N(\tilde{\omega}_l)$ and $I''_N(\tilde{\omega}_l)$ denote the first and second derivatives of $I_N(\tilde{\omega}_l)$, respectively. Here $\tilde{\omega}_l$ is the estimator of ω_l satisfying $\tilde{\omega}_l - \omega_l = O_p(N^{-1-\delta})(\delta > 0)$. Now, we compute $I'_N(\tilde{\omega}_l)$ and $I''_N(\tilde{\omega}_l)$, respectively, as follows.

Since $I_N(\tilde{\omega}_l) = \frac{1}{N}Y^H X X^H Y$ and due to (A1), we have

$$I'_N(\tilde{\omega}_l) = \frac{1}{N}(Y^H \dot{X} X^H Y + Y^H X \dot{X}^H Y) \quad (\text{A2})$$

$$= \frac{i}{N}(Y^H D X X^H Y - Y^H X X^H D Y) \quad (\text{A3})$$

and

$$\begin{aligned} I''_N(\tilde{\omega}_l) &= \frac{1}{N}(Y^H \ddot{X} X^H Y + 2Y^H \dot{X} \dot{X}^H Y \\ &\quad + Y^H X \ddot{X}^H Y) \\ &= \frac{1}{N}(-Y^H D^2 X X^H Y + 2Y^H D X X^H D Y \\ &\quad - Y^H X X^H D^2 Y), \end{aligned} \quad (\text{A4})$$

where ‘H’ denotes the operation of transpose and conjugation. By Lemmas 1 and 2, we have

$$\begin{aligned} \frac{1}{N}Y^H X &= \mu_l e^{-i\varphi_l} + O_p\left(\frac{1}{N^{1/2}}\right), \\ \frac{1}{N^2}Y^H D X &= \frac{1}{2}\mu_l e^{-i\varphi_l} + O_p\left(\frac{1}{N^{1/2}}\right), \\ \frac{1}{N^3}Y^H D^2 X &= \frac{1}{3}\mu_l e^{-i\varphi_l} + O_p\left(\frac{1}{N^{1/2}}\right). \end{aligned} \quad (\text{A5})$$

Using (A4) and (A5), we get

$$\begin{aligned} \frac{1}{N^3}I''_N(\tilde{\omega}_l) &= -\frac{1}{3}\mu_l^2 + \frac{1}{2}\mu_l^2 - \frac{1}{3}\mu_l^2 + O_p\left(\frac{1}{N^{1/2}}\right) \\ &= -\frac{1}{6}\mu_l^2 + O_p\left(\frac{1}{N^{1/2}}\right). \end{aligned} \quad (\text{A6})$$

From (A2), we have

$$\begin{aligned} \frac{1}{N^3}I'_N(\tilde{\omega}_l) &= \frac{i}{N^4}(Y^H D X X^H Y - Y^H X X^H D Y) \\ &= -\frac{2}{N^4}\text{Im}(Y^H D X X^H Y), \end{aligned} \quad (\text{A7})$$

where $\text{Im}(a)$ denotes the image part of a . Now we compute $Y^H D X X^H Y$ as follows:

$$Y^H D X X^H Y = \left[\sum_{t=1}^N \bar{y}(t) t e^{i\tilde{\omega}_l t} \right] \left[\sum_{t=1}^N y(t) e^{-i\tilde{\omega}_l t} \right], \quad (\text{A8})$$

where $\bar{y}(t)$ denotes the conjugation of $y(t)$ and

$$\begin{aligned} \sum_{t=1}^N \bar{y}(t)te^{i\tilde{\omega}_l t} &= \sum_{k=1}^p \sum_{t=1}^N \mu_k e^{-i\varphi_k t} t e^{i(\tilde{\omega}_l - \omega_k)t} \\ &+ \sum_{k=1}^p \sum_{t=1}^N \eta_k(t) e^{-i\varphi_k t} t e^{i(\tilde{\omega}_l - \omega_k)t} \\ &+ \sum_{t=1}^N \bar{\varepsilon}(t) t e^{i\tilde{\omega}_l t} \\ &\triangleq A_1 + A_2 + A_3. \end{aligned} \tag{A9}$$

We compute A_1, A_2 and A_3 as follows:

$$A_1 = \sum_{k=1}^p \mu_k e^{-i\varphi_k} \sum_{t=1}^N t e^{i(\tilde{\omega}_l - \omega_k)t} \triangleq \sum_{k=1}^p A_{1k}(N). \tag{A10}$$

If $k = l$, using the Taylor series of $e^{i(\tilde{\omega}_l - \omega_l)t}$ up to the fourth order, we have

$$\begin{aligned} A_{1l}(N) &= \mu_l e^{-i\varphi_l} \sum_{t=1}^N t e^{i(\tilde{\omega}_l - \omega_l)t} \\ &= \mu_l e^{-i\varphi_l} \left[\sum_{t=1}^N t + i(\tilde{\omega}_l - \omega_l) \sum_{t=1}^N t^2 \right. \\ &\quad - \frac{1}{2}(\tilde{\omega}_l - \omega_l)^2 \sum_{t=1}^N t^3 \\ &\quad - \frac{1}{6}i(\tilde{\omega}_l - \omega_l)^3 \sum_{t=1}^N t^4 \\ &\quad \left. + \frac{1}{24}(\tilde{\omega}_l - \omega_l)^4 \sum_{t=1}^N t^5 e^{i\theta_1(\tilde{\omega}_l - \omega_l)t} \right], \end{aligned} \tag{A11}$$

where $0 < \theta_1 < 1$. If $k \neq l$, using Lemma 1, we have

$$A_{1k}(N) = \mu_k e^{-i\varphi_k} \sum_{t=1}^N t e^{i(\tilde{\omega}_l - \omega_k)t} = O_p(N). \tag{A12}$$

Using the Taylor series of $e^{i(\tilde{\omega}_l - \omega_l)t}$ up to the first order and Lemma 2, we have

$$\begin{aligned} A_2 &= \sum_{k=1}^p \sum_{t=1}^N \eta_k(t) e^{-i\varphi_k t} t e^{i(\omega_l - \omega_k)t} e^{i(\tilde{\omega}_l - \omega_l)t} \\ &= \sum_{k=1}^p \sum_{t=1}^N \eta_k(t) e^{-i\varphi_k t} t e^{i(\omega_l - \omega_k)t} \end{aligned}$$

$$\begin{aligned} &+ i(\tilde{\omega}_l - \omega_l) \sum_{k=1}^p \sum_{t=1}^N \eta_k(t) e^{-i\varphi_k t} t^2 \\ &\times e^{i(\omega_l - \omega_k)t} e^{i\theta_2(\tilde{\omega}_l - \omega_l)t} \\ &= \sum_{k=1}^p \sum_{t=1}^N \eta_k(t) e^{-i\varphi_k t} t e^{i(\omega_l - \omega_k)t} + O_p(N^{3/2-\delta}), \end{aligned} \tag{A13}$$

where $0 < \theta_2 < 1$. Similarly, we have

$$A_3 = \sum_{t=1}^N \varepsilon(t) t e^{i\tilde{\omega}_l t} + O_p(N^{3/2-\delta}), \tag{A14}$$

and

$$\begin{aligned} \sum_{t=1}^N y(t) e^{-i\tilde{\omega}_l t} &= \sum_{k=1}^p \sum_{t=1}^N \mu_k e^{i\varphi_k} e^{i(\omega_k - \tilde{\omega}_l)t} \\ &+ \sum_{k=1}^p \sum_{t=1}^N \eta(k) e^{i\varphi_k} e^{i(\omega_k - \tilde{\omega}_l)t} \\ &+ \sum_{t=1}^N \varepsilon(t) e^{-i\tilde{\omega}_l t} \\ &= \mu_l N e^{i\varphi_l} + O_p(N^{1-\delta}). \end{aligned} \tag{A15}$$

Combining (A7)–(A15), we have

$$\begin{aligned} \frac{1}{N^3} I'_N(\tilde{\omega}_l) &= -\frac{2}{N^4} \text{Im}(Y^H D X X^H Y) \\ &= -\frac{2\mu_l^2}{N^3} \left[(\tilde{\omega}_l - \omega_l) \sum_{t=1}^N t^2 \right. \\ &\quad - \frac{1}{6}(\tilde{\omega}_l - \omega_l)^3 \sum_{t=1}^N t^4 \left. \right] + O_p(N^{-2}) \\ &\quad + O_p(N^{-1-4\delta}) + O_p(N^{-3/2-\delta}) \\ &\quad - \frac{2\mu_l}{N^3} \left\{ \sum_{k=1}^p \sum_{t=1}^N \eta_k(t) t \sin [(\omega_l - \omega_k)t - \varphi_k] \right. \\ &\quad + \sum_{t=1}^N [\text{Re}(\varepsilon(t)) t \sin(\omega_l t) \\ &\quad \left. + \text{Im}(\varepsilon(t)) t \cos(\omega_l t) \right\}, \end{aligned} \tag{A16}$$

where $\text{Re}(a)$ denotes the real part of a . Finally, combing (A6) and (A16), we have

$$\begin{aligned} \hat{\omega}_l &= \tilde{\omega}_l - \frac{1}{4} \frac{I'_N(\tilde{\omega}_l)}{I''_N(\tilde{\omega}_l)} = \tilde{\omega}_l - \frac{1}{4} \frac{\frac{1}{N^3} I'_N(\tilde{\omega}_l)}{\frac{1}{N^3} I''_N(\tilde{\omega}_l)} \end{aligned}$$

$$\begin{aligned}
 &= \tilde{\omega}_l - \frac{1}{4} \frac{-\frac{2\mu_l^2}{N^3} [(\tilde{\omega}_l - \omega_l) \sum_{t=1}^N t^2 - \frac{1}{6}(\tilde{\omega}_l - \omega_l)^3 \sum_{t=1}^N t^4]}{-\frac{1}{6}\mu_l^2 + O_p\left(\frac{1}{N^{1/2}}\right)} \\
 &\quad - \frac{3}{\mu_l N^3} \left\{ \sum_{k=1}^p \sum_{t=1}^N \eta_k(t) t \sin[(\omega_l - \omega_k)t - \varphi_k] \right. \\
 &\quad \left. + \sum_{t=1}^N [\operatorname{Re}(\varepsilon(t))t \sin(\omega_l t) + \operatorname{Im}(\varepsilon(t))t \cos(\omega_l t)] \right\} \\
 &\quad + O_p(N^{-1-4\delta}) + O_p(N^{-3/2-\delta}) \\
 &= \omega_l + (\tilde{\omega}_l - \omega_l) O_p(N^{-2\delta}) \\
 &\quad - \frac{3}{\mu_l N^3} \left\{ \sum_{k=1}^p \sum_{t=1}^N \eta_k(t) t \sin [(\omega_l - \omega_k)t - \varphi_k] \right. \\
 &\quad \left. + \sum_{t=1}^N [\operatorname{Re}(\varepsilon(t))t \sin(\omega_l t) + \operatorname{Im}(\varepsilon(t))t \cos(\omega_l t)] \right\} \\
 &\quad + O_p(N^{-1-4\delta}) + O_p(N^{-3/2-\delta}). \tag{A17}
 \end{aligned}$$

If $\delta \leq \frac{1}{6}$, then $\hat{\omega}_l - \omega_l = O_p(N^{-1-3\delta})$. If $\delta > \frac{1}{6}$, then using the central limit theorem (Fuller, 1996) we have

$$\begin{aligned}
 &\hat{\omega}_l - \omega_l \\
 &\triangleq -\frac{3}{\mu_l N^{3/2}} \left\{ \sum_{k=1}^p \sum_{t=1}^N \eta_k(t) t \sin [(\omega_l - \omega_k)t - \varphi_k] \right. \\
 &\quad \left. + \sum_{t=1}^N [\operatorname{Re}(\varepsilon(t))t \sin(\omega_l t) + \operatorname{Im}(\varepsilon(t))t \cos(\omega_l t)] \right\} \\
 &\quad + O_p(N^{-1-4\delta}) + O_p(N^{-3/2-\delta}) \\
 &\xrightarrow{\mathcal{L}} \mathcal{N}_p(\mathbf{0}, \mathbf{\Sigma}),
 \end{aligned}$$

where $\mathcal{N}_p(\mathbf{0}, \mathbf{\Sigma})$ is defined as in Theorem 1.

Received: 19 October 2011

Revised: 12 May 2012