Statistical Analysis of Non Linear Least Squares Estimation for Harmonic Signals in Multiplicative and Additive Noise

HUIMING PENG, SHAOQUAN YU, JIAWEN BIAN, YUJIE ZHANG, AND HONGWEI LI

School of Mathematics and Physics, China University of Geosciences, Wuhan, China

In this paper we consider the problem of parameter estimation for the multicomponent harmonic signals in multiplicative and additive noise. The nonlinear least squares (NLLS) estimators, NLLS\textsubscript{1} and NLLS\textsubscript{2} proposed by Ghogho et al. (1999b) to estimate the coherent model parameters for single-component harmonic signal, are generalized to the multicomponent harmonic signals for the cases of nonzero- and zero-mean multiplicative noise, respectively. By statistical analysis, some asymptotic results of the NLLS estimators are derived, including the strong consistency, the strong convergence rate and the asymptotic normality. Furthermore, the NLLS\textsubscript{1} and NLLS\textsubscript{2} based estimators are proposed to estimate the noncoherent model parameters for the cases of nonzero- and zero-mean multiplicative noise, respectively, meanwhile the strong consistency and the asymptotic normality of the NLLS-based estimators are also derived. Finally some numerical experiments are performed to see how the asymptotic results work for finite sample sizes.

Keywords  Asymptotic normality; Harmonic signal; Multiplicative noise; Non-linear least squares; Strong consistency; Strong convergence rate

Mathematics Subject Classification  Primary 62F12; Secondary 62E20

1. Introduction

In this paper, we consider the following model of the multicomponent harmonic signals in multiplicative and additive noise:

\[ x(t) = \sum_{k=1}^{M} s_k(t)e^{j(\omega_k t + \phi_k)} + n(t), \quad \text{for } t = 1, 2, \ldots, T, \]
where \( j = \sqrt{-1} \), \( M \), \( \omega_0^k \)'s, \( \phi_0^k \)'s, \( s_k(t) \)'s, and \( n(t) \) are the number of signals, the frequencies, the phases, the multiplicative noise and the additive noise, respectively. In addition, we make the following assumptions:

**Assumption 1.** (i) The frequencies \( \omega_0^k \)'s are distinct in \((0, \pi)\) and the phases \( \phi_0^k \)'s are deterministic constants in \([0, \pi)\); (ii) The multiplicative noise \( s_k(t) \) is a sequence of i.i.d. real-valued random variables with mean \( \mu_{s_k}^0 \geq 0 \) and variance \( \sigma_{s_k}^2 \); (iii) The additive noise \( n(t) \) is a sequence of i.i.d. complex-valued random variables with mean zero, and both the real and imaginary parts have variance \( \sigma_0^2/2 \) and they are independent; (iv) The noise \( s_k(t) \)'s and \( n(t) \) are mutually independent.

**Assumption 2.** Let \( n_r(t) \) and \( n_i(t) \) denote the real and imaginary parts of \( n(t) \), respectively. We assume that all the noise have finite fourth-order moments, i.e., \( \mathbb{E}[n_r^4(t)] < \infty \), \( \mathbb{E}[n_i^4(t)] < \infty \), and \( \mathbb{E}[s_k^4(t)] < \infty \), for \( k = 1, 2, \ldots, M \).

In general, \( \omega^0 = (\omega_0^1, \ldots, \omega_0^M) \), \( \phi^0 = (\phi_0^1, \ldots, \phi_0^M) \), \( \mu^0 = (\mu_{s_1}^0, \ldots, \mu_{s_M}^0) \), \( \sigma_{s_k}^2 \)'s and \( \sigma_0^2 \) are unknown. It is desired to estimate these unknown parameters. Here we assume that \( M \) is known.

This is an important problem in statistical signal processing. In the last 20 years, a lot of iterative and noniterative procedures were developed to estimate the frequencies efficiently when the signals are only contaminated by additive noise. We may refer to Bressler and Macovski (1986), Hwang and Chen (1993), Kannan and Kundu (1994), Li and Stoica (1996), Chan et al. (2010), So et al. (2010), and Sun and So (2012) for the iterative methods aimed at finding the nonlinear least squares (NLLS) estimators, and refer to Tufts and Kumaresan (1982), Roy and Kailath (1989), Quinn (1994), Kundu and Mitra (1995), and Bai et al. (2003) for the noniterative methods. There are also many other high-resolution methods such as Shi and Fairman (1994), Zhang et al. (1994), Sadler et al. (1995), Zhang and Wang (2000), and Chan and So (2004). The performances of these methods depend on the successful detection of the number of signals. Thus several methods have been suggested for this purpose such as Zhao et al. (1986), Rao (1988), Quinn (1989), Wang (1993), Kavalieris and Hannan (1994), Kundu and Mitra (2000), and Yang and Li (2007a). For more information on parameter estimation of harmonics in additive noise, the readers may refer to Quinn and Hannan (2001) and the references therein.

However, multiplicative noise may also occur in a variety of applications. For example, in Doppler-radar processing (Giannakis and Zhou, 1995), knowledge of the frequency from a pulse train reflected from a moving object yields the target’s velocity, and it is more appropriate to model the harmonic signals as having random rather than constant amplitude if the point-target assumption is no longer valid, or, when the target scintillates; in underwater acoustic applications (Giannakis and Zhou, 1995), the multiplicative noise describes the effects on acoustic waves due to the fluctuations caused by the medium (diffraction, internal waves, or microstructures) changing orientation and interference from scatterers of the target. Several methods have been suggested to estimate the frequencies of harmonic signals in multiplication and additive noise under the assumption that the number of signal is known, such as second-order statistics (Besson and Castanie, 1993), higher-order statistics (Swami, 1994), cyclic statistics (Giannakis and Zhou, 1995; Zhou and Giannakis, 1995; Li and Cheng, 1998; Ghogho et al., 1999a), and three-step iterative approach (Bian et al., 2009 and 2011). However, in practice, the number of signals can be unknown and estimating the number of signals will be the first step. Without multiplicative noise, there
are some methods to determine the number of signals, e.g., Akaike information criterion (AIC) (Akaike, 1974) and minimum description length (MDL) (Schwarz, 1978). But the performance of AIC and MDL will degrade in the presence of unknown multiplicative noise. For this purpose, a few methods have been proposed to estimate the number of signals of harmonics in multiplicative and additive noise. For example, Li and Cheng (1998) proposed cyclic statistics based on the statistical properties of sample cyclic-moments, and obtained the strong consistency and strong convergence rate of the proposed estimator; Yang and Li (2007b) proposed enhanced matrix based on the eigenvalue properties of covariance matrix of the constructed enhanced matrix.

It is well known that the NLLS estimation has been seen as the most intuitive method and it has played a very important role in harmonic parameter estimation. As early as 1970s, the statisticians Walker (1971, 1973) and Hannan (1973) preliminarily studied the NLLS method in harmonic parameter estimation. Later, much more statisticians like Rao and Zhao (1993), Giannakis and Zhou (1995), Stoica et al. (1997), Kundu and Mitra (1999), Besson and Stoica (1999), Ghogho et al. (1999b), Cohen and Francos (2002) did a lot of works and moved forward this field further. For example, Kundu and Mitra (1999) considered the multicomponent harmonics model in additive noise and derived some fantastic statistical results for the NLLS estimators including the strong consistency, the strong convergence rate and the asymptotic normality and so on; Ghogho et al. (1999b) considered the single-component harmonics model in multiplicative and additive noise and proposed two specific NLLS estimators, i.e. NLLS$_1$ and NLLS$_2$ that consist of matching the data and the squared data respectively, moreover the expressions for the asymptotic covariances of the two NLLS estimators were derived. In this paper, we generalize the NLLS estimators proposed by Ghogho et al. (1999b) to the case of multi-component harmonics in multiplicative and additive noise. We separately address the cases of nonzero- and zero-mean multiplicative noise by using NLLS$_1$ and NLLS$_2$, respectively, to estimate the coherent model parameters. By following the proof techniques proposed by Kundu and Mitra (1999), the asymptotic covariances of the two NLLS estimators in the case of multicomponent multiplicative noise are derived, which are consistent with the corresponding results (see Ghogho et al., 1999b) if it is reduced to the single-component model, moreover the strong consistency and the strong convergence rate are also derived. Furthermore, the NLLS-based estimators are proposed to estimate the noncoherent model parameters, and the strong consistency, the strong convergence rate and the asymptotic normality for the proposed estimators are also derived. Finally, the simulation results verify the efficiency of all the estimators.

The rest of the paper is organized as follows. In Sec. 2, we describe the parameter estimation based on NLLS$_1$ and NLLS$_2$ for multicomponent model (1) in the cases of nonzero- and zero-mean multiplicative noise, respectively. The statistical analysis and the theoretical results for the estimators are provided in Sec. 3. In Sec. 4, we present some numerical experiments, and finally we conclude the paper in Sec. 5. All proofs are provided in the Appendices.

2. Parameter Estimation Based on NLLS

2.1. Nonzero-mean Multiplicative Noise

In this subsection, we consider $\mu_{0k} \neq 0$, for $k = 1, 2, \ldots, M$. First, the data-based NLLS$_1$ proposed by Ghogho et al. (1999b) is employed to estimate the coherent model parameters including the frequencies, the phases, and the multiplicative noise means in this case.
Then, the NLLS$_1$-based estimators are proposed to estimate the noncoherent model parameters, i.e., the noise parameters including the multiplicative noise variances, the total noise variance, and the additive noise variance.

In this case, the coherent and noncoherent components in data $x(t) = y(t) + \varepsilon(t)$ are given, respectively, by

$$y(t) = \sum_{k=1}^{M} \mu^0_{sk} e^{j(\omega^0_k t + \phi^0_k)}$$  \hspace{1cm} (2)

$$\varepsilon(t) = \sum_{k=1}^{M} (s_k(t) - \mu^0_{sk}) e^{j(\omega^0_k t + \phi^0_k)} + n(t) \hspace{1cm} (3)$$

Notice that the noncoherent component $\varepsilon(t)$ is zero-mean and can be regarded as additive noise. Hence, the coherent model parameters including the frequencies, the phases, and the multiplicative noise means in this case can be estimated by least squares estimation directly based on data $x(t)$ referred as NLLS$_1$ (see Ghogho et al., 1999b). We denote the parameter vector $\theta = (\theta_1, \theta_2, \ldots, \theta_M)$, where $\theta_k = (\omega_k, \phi_k, \mu_{sk})$ for $k = 1, 2, \ldots, M$. Similarly, $\theta^0$ and $\theta^0_k$s are also defined. Now the NLLS$_1$ of true parameter vector $\theta^0$, i.e., $\hat{\theta}$, can be obtained by

$$\hat{\theta} = \arg \min_{\theta} Q(\theta),$$  \hspace{1cm} (4)

where

$$Q(\theta) = \frac{1}{T} \sum_{t=1}^{T} |x(t) - \sum_{k=1}^{M} \mu_{sk} e^{j(\omega_k t + \phi_k)}|^2$$

Based on the NLLS$_1$ defined above, the following estimators are proposed to estimate the multiplicative noise variance $\sigma^2_{sk}$, for $k = 1, 2, \ldots, M$, and the total noise variance $\sigma^2 = \sigma^2_0 + \sum_{k=1}^{M} \sigma^2_{sk}$, respectively,

$$\hat{\sigma}^2_{sk} = \text{Re} \left[ \frac{1}{T} \sum_{t=1}^{T} x^2(t) e^{-j(2\omega_k t + 2\phi_k)} \right] - \hat{\mu}^2_{sk},$$  \hspace{1cm} (5)

$$\hat{\sigma}^2 = \frac{1}{T} Q(\hat{\theta}),$$ \hspace{1cm} (6)

where the symbol “Re” denotes the real part of a complex value. Finally, combining the estimations of the multiplicative noise variances and the total noise variance, the additive noise variance $\sigma^2_0$ can be estimated intuitively by

$$\hat{\sigma}^2_0 = \hat{\sigma}^2 - \sum_{k=1}^{M} \hat{\sigma}^2_{sk}$$ \hspace{1cm} (7)

### 2.2. Zero-mean Multiplicative Noise

In this subsection, we consider $\mu^0_{sk} = 0$, for $k = 1, 2, \ldots, M$. First, the squared-data-based NLLS$_2$ proposed by Ghogho et al. (1999b) is employed to estimate the coherent model.
parameters, including the frequencies, the phases, and the multiplicative noise variances in this case. Then, the NLLS-based estimators are proposed to estimate the noncoherent model parameters, i.e., the remaining noise parameters including the total noise variance and the additive noise variance.

In this case, the coherent harmonic signals in \( x(t) \) vanish. Fortunately, spectral analysis of the squared data allows one to recover the harmonic signals (see Ghogho et al., 1999b; Giannakis and Zhou, 1995). The coherent and noncoherent components in squared data

\[
x^2(t) = y(t) + \varepsilon(t)
\]

are given, respectively, by

\[
y(t) = \sum_{k=1}^{M} p_k^0 e^{i(2\omega_0^k t + 2\phi_0^k)},
\]

\[
\varepsilon(t) = \sum_{k=1}^{M} \left( s_k^2(t) - p_k^0 \right) e^{i(2\omega_0^k t + 2\phi_0^k)} + 2n(t) \sum_{k=1}^{M} s_k(t) e^{i(\omega_0^k t + \phi_0^k)}
\]

\[
+ \sum_{k \neq l} s_k(t)s_l(t) e^{i(\omega_0^k t + \phi_0^k + \omega_0^l t + \phi_0^l)} + n^2(t),
\]

where \( p_k^0 = (\sigma_0^2)^k \), for \( k = 1, 2, \ldots, M \). Notice that the noncoherent component \( \varepsilon(t) \) is zero-mean and can be regarded as additive noise. Hence, the coherent model parameters including the frequencies, the phases, and the multiplicative noise variances in this case can be estimated by least squares estimation based on the squared data \( x^2(t) \) referred as NLLS\( _2 \) (see Ghogho et al., 1999b). We denote the parameter vector \( \theta = (\theta_1, \theta_2, \ldots, \theta_M) \), where \( \theta_k = (\omega_k, \phi_k, p_k) \) for \( k = 1, 2, \ldots, M \). Similarly, \( \theta^0 \) and \( \theta_0^k \)'s are also defined. Now the NLLS\( _2 \) of true parameter vector \( \theta^0 \), i.e., \( \hat{\theta} \), can be obtained by

\[
\hat{\theta} = \arg\min_{\theta} Q(\theta),
\]

where

\[
Q(\theta) = \sum_{t=1}^{T} \left| x^2(t) - \sum_{k=1}^{M} p_k e^{i(2\omega_0^k t + 2\phi_0^k)} \right|^2.
\]

Notice that \( x(t) \) is a sequence of independent complex-valued random variables with mean zero under Assumption 1. Hence, the total noise variance \( \sigma^2 = \sigma_0^2 + \sum_{k=1}^{M} \sigma_{sk}^2 \), i.e., the variance of \( x(t) \), can be estimated directly by

\[
\hat{\sigma}^2 = \frac{1}{T} \sum_{t=1}^{T} |x(t)|^2
\]

Finally, combining the estimations of the multiplicative noise variances and the total noise variance, the additive noise variance \( \sigma_0^2 \) can be estimated intuitively by

\[
\hat{\sigma}_0^2 = \hat{\sigma}^2 - \sum_{k=1}^{p} \hat{\sigma}_{sk}^2.
\]
3. Statistical Analysis

In the case of multi component harmonics in additive noise, the asymptotic performance of the NLLS estimators was studied by Kundu and Mitra (1999). In this section, by following the proof techniques proposed by Kundu and Mitra (1999), we address the case of multi-component harmonics in multiplicative and additive noise and derive some statistical results including the strong consistency, the strong convergence rate, and the asymptotic normality for the estimators described in Sec. 2 corresponding to the cases of nonzero- and zero-mean multiplicative noise.

3.1. Nonzero-mean Multiplicative Noise

We first state the statistical results of the NLLS for the coherent model parameters including the frequencies, the phases, and the multiplicative noise means as follows.

Theorem 1. Under Assumption 1, \( \hat{\theta} \) is a strongly consistent estimator of \( \theta^0 \).

Theorem 2. Under Assumption 1, \( N(\hat{\omega}_k - \omega^0_k) \rightarrow 0 \), a.s. for \( k = 1, 2, \ldots, M \).

Theorem 3. Under Assumption 1, \( (\hat{\theta} - \theta^0)D \) converges in distribution to a \( 3M \)-variate normal distribution with mean vector zero and the covariance matrix \( \Sigma = \text{diag}\{\Sigma_1, \Sigma_2, \ldots, \Sigma_M\} \), which is a \( M \)-dimensional block diagonal matrix whose block element is given by, for \( k = 1, 2, \ldots, p, \)

\[
\Sigma_k = \begin{bmatrix}
\sigma^2 - \sigma^2_{sk} & 6 & 0 \\
-3 & -3 & 0 \\
0 & 1 & \frac{1}{2}(\sigma^2 + \sigma^2_k)
\end{bmatrix},
\]

where \( D = \text{diag}\{D_1, D_2, \ldots, D_M\} \) is a \( M \)-dimensional block diagonal matrix whose block element is given by \( D_k = \text{diag}\{T^{3/2}, T^{1/2}, T^{1/2}\} \), for \( k = 1, 2, \ldots, M \).

Corollary 1. Under the same assumption of Theorem 3, the NLLS\(_1\)'s, i.e., \( \hat{\omega}_k, \hat{\phi}_k \) and \( \hat{\mu}_{sk} \), for \( k = 1, 2, \ldots, M \), are asymptotically consistent and normal, with asymptotic variances given by

\[
\text{Var}(\hat{\omega}_k) = \frac{6}{T^3} \frac{\sigma^2 - \sigma^2_{sk}}{\left(\mu^0_{sk}\right)^2}, \text{Var}(\hat{\phi}_k) = \frac{2}{T} \frac{\sigma^2 - \sigma^2_{sk}}{\left(\mu^0_{sk}\right)^2}, \text{Var}(\hat{\mu}_{sk}) = \frac{1}{2T} \left(\sigma^2 + \sigma^2_{sk}\right)
\]

The expressions of the variances in Corollary 1 follow from the asymptotic covariance matrix (ACM) in Theorem 3. By comparing the ACM in Theorem 3 with the corresponding result of NLLS\(_1\) obtained by Ghogho et al. (1999b) corresponding to the case of single-component harmonics in multiplicative and additive noise, it shows the consistency if the multicomponent model in (1) is reduced to the single-component case. Note that Theorem 1 is a very strong result, it means that the NLLS\(_1\)'s will be closer and closer to the true parameter values as the sample size increases. Therefore the biases and the mean squared errors (MSEs) tend to zero as sample size increases. Meanwhile, Theorem 2 gives the strong convergence rate of the frequency estimator that is also an important issue in statistics. Furthermore, the obtained asymptotic distributions in Theorem 3 will help us
to obtain the confidence bounds for these coherent parameters. From the expression of ACM in Theorem 3, it is immediate that the corresponding NLLS’s are asymptotically uncorrelated for distinct harmonic components, the NLLS’s of the multiplicative noise mean are asymptotically uncorrelated with both of the NLLS’s of the frequency and the phase for the same harmonic component, and the ACM does not depend on the frequencies and the phases but only depends on the noise. Note that Theorem 3 also gives us the idea of the weak convergence rate of the NLLS’s, such as it is \( O_p(T^{-3/2}) \) for the frequencies but \( O_p(T^{-1/2}) \) for all of the phases and the multiplicative noise means, which indicates that the estimation accuracy for the nonlinear parameters is better than the one for the linear parameters.

Now we state the statistical results of the NLLS-based estimators for the noncoherent model parameters, i.e., the noise variances including the multiplicative noise variances, the total noise variance, and the additive noise variance.

**Theorem 4.** Under Assumptions 1 and 2, \( \hat{\sigma}_{sk}^2 \)'s are strongly consistent estimators of \( \sigma_{sk}^2 \)'s, and it also can be obtained that \( \sqrt{T}(\hat{\sigma}_{sk}^2 - \sigma_{sk}^2) \) converges in distribution to a normal distribution with mean zero and variance \( \sigma_{sk}^* \), which is given by

\[
\sigma_{sk}^* = \text{Var}[s_k^2(1)] + 2(\sigma^2 + \sigma_{sk}^2)(\mu_{sk}^0)^2
\]

**Theorem 5.** Under Assumptions 1 and 2, \( \hat{\sigma}^2 \) is a strongly consistent estimator of \( \sigma^2 \), and it also can be obtained that \( \sqrt{T}(\hat{\sigma}^2 - \sigma^2) \) converges in distribution to a normal distribution with mean zero and variance \( \sigma^* \), which is given by

\[
\sigma^* = \text{Var}[|n(1)|^2] - \sigma_0^4 + \sum_{k=1}^M \left[ \text{Var}[|s_k(1) - \mu_{sk}^0|^2] - \sigma_{sk}^4 \right] + \sigma^4
\]

**Theorem 6.** Under Assumptions 1 and 2, \( \hat{\sigma}_0^2 \) is a strongly consistent estimator of \( \sigma_0^2 \), and it also can be obtained that \( \sqrt{T}(\hat{\sigma}_0^2 - \sigma_0^2) \) converges in distribution to a normal distribution with mean zero and variance \( \sigma_0^* \), which is given by

\[
\sigma_0^* = \sum_{k=1}^M \text{Var}[|s_k(1) - \mu_{sk}^0|^2] + \sum_{k=1}^M \text{Var}[s_k^2(1)] - \sum_{k=1}^M \sigma_{sk}^4
\]

\[
+ 2\sum_{k=1}^p \left( \sigma^2 + \sigma_{sk}^2 \right)(\mu_{sk}^0)^2 + \sigma^4 - \sigma_0^4 + \text{Var}[|n(1)|^2]
\]

**Corollary 2.** If we further assume that all the noise distributions are Gaussian, then the expressions of the asymptotic variances in Theorems 4–6 can be reduced to

\[
\sigma_{sk}^* = 2\sigma_{sk}^4 + 2(\sigma^2 + 3\sigma_{sk}^2)(\mu_{sk}^0)^2
\]

\[
\sigma^* = \sigma^4 + \sum_{k=1}^M \sigma_{sk}^4
\]
\[ \sigma_0^* = \sigma^4 + 3 \sum_{k=1}^{M} \sigma_{sk}^4 + 2 \sum_{k=1}^{M} (\sigma^2 + \sigma_{sk}^2)(\mu_{sk}^2)^2 \]

**Corollary 3.** Under Assumptions 1 and 2, the estimators of all the noise variances, i.e., \( \hat{\sigma}_{sk}^2 \)'s, \( \hat{\sigma}^2 \), and \( \hat{\sigma}_0^2 \), are asymptotically consistent and normal, with asymptotic variances given by

\[
\begin{align*}
\text{Var}(\hat{\sigma}_{sk}^2) &= \frac{\sigma_{sk}^*}{T}, \\
\text{Var}(\hat{\sigma}^2) &= \frac{\sigma^*}{T}, \\
\text{Var}(\hat{\sigma}_0^2) &= \frac{\sigma_{0*}^*}{T}
\end{align*}
\]

where \( \sigma_{sk}^* \)'s, \( \sigma^* \), and \( \sigma_{0*}^* \) are defined in Theorems 4–6, respectively.

The simple formulas in Corollary 2 follow from the complicated ones in Theorems 4–6 under the additional Gaussian assumption. Theorems 4–6 imply that, under Assumptions 1 and 2, it is possible to estimate the parameters quite accurately not only for the frequencies, the phases and the multiplicative noise means but also for the multiplicative noise variances, the total noise variance and the additive noise variance when the sample size is large enough. Similarly as Theorem 3, Theorems 4–6 also tell us that the weak convergence rate of all the noise variances is \( O_P(T^{-1/2}) \). Moreover the obtained asymptotic distributions will help us to obtain the confidence bounds for the unknown noise variances.

### 3.2. Zero-mean Multiplicative Noise

In this subsection, we state the results for the case of zero-mean multiplicative noise along the same line as the statements for the case of nonzero-mean multiplicative noise. For the purpose of simplification or compactness of the paper, we simplify some statements by comparison with the case of nonzero-mean multiplicative noise, and we also omit some notations on the results, the readers can easily get the similarities and the differences between the cases of nonzero- and zero-mean multiplicative noise by simple comparison.

We first state the statistical results of the NLLS \( L^2 \) for the coherent model parameters including the frequencies, the phases, and the multiplicative noise variances as follows.

**Theorem 7.** The same result (Theorem 1) holds.

**Theorem 8.** The same result (Theorem 2) holds.

**Theorem 9.** The similar result (Theorem 3) holds with the only difference lied in the covariance matrix whose block element is given by, for \( k = 1, 2, \ldots, M \),

\[
\sum_k = \begin{bmatrix}
A_k & \left( \begin{array}{cc} 6 & -3 \\ -3 & 2 \end{array} \right) & 0 \\
0 & 0 & B_k/2
\end{bmatrix},
\]

where

\[
A_k = \text{Var}[|n(1)|^2] + \sum_{i=1}^{M} \mathbb{E}[s_i^2(1)] + \sigma_0^4 + 4\sigma_0^2 \sum_{i=1}^{M} \sigma_{si}^2 + \sum_{i \neq j} \sigma_{si}^2 \sigma_{sj}^2 - \text{Var}[s_k^2(1)]
\]
\[ B_k = \text{Var}[|n(1)|^2] + \sum_{i=1}^{M} E[s_i^2(1)] + \sigma_0^4 + 4\sigma_0^2 \sum_{i=1}^{M} \sigma_i^2 + \sum_{i \neq j} \sigma_i^2 \sigma_j^2 + \text{Var}[s_k^2(1)] \]

**Corollary 4.** If we further assume that all the noise distributions are Gaussian, then the expressions of \( A_k \) and \( B_k \) in Theorem 9 can be reduced to
\[
A_k = 2\sigma^4 - \sum_{i \neq j} \sigma_i^2 \sigma_j^2 - 2\sigma_{sk}^4, \quad B_k = 2\sigma^4 - \sum_{i \neq j} \sigma_i^2 \sigma_j^2 + 2\sigma_{sk}^4
\]

**Corollary 5.** Under the same assumption of Theorem 9, the LSEs, i.e., \( \hat{\omega}_k, \hat{\phi}_k, \) and \( \hat{\sigma}_{sk}^2 \), for \( k = 1, 2, \ldots, M \), are asymptotically consistent and normal, with asymptotic variances given by
\[
\text{Var}(\hat{\omega}_k) = \frac{3}{2T^3} \frac{A_k}{\sigma_{sk}^4}, \quad \text{Var}(\hat{\phi}_k) = \frac{1}{2T} \frac{A_k}{\sigma_{sk}^4}, \quad \text{Var}(\hat{\sigma}_{sk}^2) = \frac{1}{2T} B_k
\]

Note that, by comparing the ACM in Theorem 9 with the corresponding result of NLLS 2 obtained by Ghogho et al. (1999b) corresponding to the case of single-component harmonics in multiplicative and additive noise, it shows the consistency if the multicomponent model in (1) is reduced to the single-component case.

Now we state the statistical results of the NLLS 2-based estimators for the noncoherent model parameters, i.e., the remaining noise variances including the total noise variance and the additive noise variance.

**Theorem 10.** The same result (Theorem 5) holds if we replace \( \mu_0^0 \)’s by zero.

**Theorem 11.** The similar result (Theorem 6) holds with the only difference lied in the variance that is given by
\[
\sigma_0^* = \frac{M + 2}{2} \text{Var}[|n(1)|^2] + \frac{M + 3}{2} \sum_{k=1}^{M} \text{Var}[s_k^2(1)] + \frac{M}{2} \sigma_0^4
\]
\[
+ 2(M + 1)\sigma_0^2 \sum_{k=1}^{M} \sigma_{sk}^2 + \frac{M + 2}{2} \sum_{k \neq l} \sigma_{sk}^2 \sigma_{sl}^2
\]

**Corollary 6.** If we further assume that all the noise distributions are Gaussian, then the expressions of the asymptotic variances in Theorems 10–11 can be reduced to
\[
\sigma^* = \sigma^4 + \sum_{k=1}^{M} \sigma_{sk}^4, \quad \sigma_0^* = (M + 1)\sigma^4 + 2 \sum_{k=1}^{M} \sigma_{sk}^4 - \frac{M}{2} \sum_{k \neq l} \sigma_{sk}^2 \sigma_{sl}^2
\]

**Corollary 7.** Under Assumptions 1 and 2, the estimators of both the total noise variance and the additive noise variance, i.e., \( \hat{\sigma}^2 \) and \( \hat{\sigma}_0^2 \), are asymptotically consistent and normal,
with asymptotic variances given by
\[ \text{Var}(\hat{\sigma}^2) = \frac{\sigma^*}{T}, \quad \text{Var}(\hat{\sigma}_0^2) = \frac{\sigma_0^*}{T}, \]
where \( \sigma^* \) and \( \sigma_0^* \) are defined in Theorem 10 and Theorem 11, respectively.

4. Numerical Results

In this section, we present some numerical experiment results to see how the NLLS’s or the NLLS’s-based estimators work for finite sample sizes, and whether the asymptotic results can be used for small sample sizes. We consider the following common model:

\[ x(t) = s_1(t) e^{i(1.0t+2.0)} + s_2(t) e^{i(2.0t+3.0)} + n(t) \]  

(13)

In this case, both the multiplicative noise \( s_1(t) \) and \( s_2(t) \) are i.i.d. real Gaussian random variables whose variances are 1.0 and 2.0, respectively. The additive noise \( \varepsilon(t) \) is i.i.d. complex Gaussian random variables with mean zero, and both of the real and imaginary parts have variance 0.5. All the noises are mutually independent.

In the case of nonzero-mean multiplicative noise, the multiplicative noise means in model (13) are taken as 1.0 and 2.0, respectively. We consider six different sample sizes, namely \( T = 50, 100, 150, 200, 250, \) and 300. For each data set generated from the above model, the coherent model parameters including the frequencies \( (\omega_1^0, \omega_2^0) = (1.0, 2.0) \), the phases \( (\phi_1^0, \phi_2^0) = (2.0, 3.0) \) and the multiplicative noise means \( (\mu_1^0, \mu_2^0) = (1.0, 2.0) \), can be estimated by NLLS1 proposed in Eq. (4), after that the noncoherent model parameters including the multiplicative noise variances \( (\sigma_{s_1}^2, \sigma_{s_2}^2) = (1.0, 2.0) \), the total noise variance \( \sigma^2 = 4.0 \) and the additive noise variance \( \sigma_0^2 = 1.0 \) can be estimated by the NLLS1-based estimators proposed in Eqs (5–7) respectively. We replicate the process 500 times and calculate the average estimates (AE) and the MSEs for all the parameters. The corresponding asymptotic variances (AVAR) are also reported for the purpose of comparison. We also calculate the approximate 90% confidence limits for all the parameters and obtain the expected confidence interval length (length) using the true parameter values (see Kundu and Mitra, 1999). The coverage percentages (coverage) are also obtained over 500 replications. It worth noting that, due to the highly nonlinearity of the objective function of the NLLS1, we employed a hybrid stochastic searching algorithm to solve this nonlinear programming problem. The genetic algorithm was first used as a global method to search the entire parameter space, then followed by the Nelder–Mead Simplex algorithm to search for the local minimum in the vicinity of output from the genetic algorithm. The corresponding optimization procedures were implemented using the MATLAB gatool and fminsearch functions, respectively. Multiple settings for the genetic algorithm were tested, and the following were selected, Generations = 500; PopulationSize = 120 (20 \times \text{number of parameters}); EliteCount = 8; CrossoverFraction = 0.5; MutationFcn = \{ @mutationuniform, 0.5 \}. All the computations are performed in MATLAB 7.5 (R2007b). All the numerical results are presented in Tables 1–4 corresponding to the frequencies, the phases, the multiplicative noise means, and the noise variances, respectively.

The following observations are very clear from the numerical experiments. First, it can be observed that the AEs of all the parameters are very close to the true parameter values in all the considered cases, meanwhile the MSEs of all the parameters gradually decrease and approach the AVarS as the sample size increases, which verifies the consistency of the proposed estimators and also shows the validity of the asymptotic results even for moderate sample sizes. Next, it is also clear that in all cases the Lengths decrease as
Table 1
The results of the frequencies

| Sample size | 50     | 100    | 150    | 200    | 250    | 300    |
|-------------|--------|--------|--------|--------|--------|--------|
| $\omega_0^1$ | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| AE          | 1.0112 | 1.0008 | 1.0010 | 1.0007 | 1.0001 | 1.0000 |
| MSE         | $2.124e-4$ | $2.187e-5$ | $6.148e-6$ | $2.868e-6$ | $1.557e-6$ | $7.243e-7$ |
| AVAR        | $1.440e-4$ | $1.800e-5$ | $5.333e-6$ | $2.250e-6$ | $1.152e-6$ | $6.667e-7$ |
| Length      | 0.0394  | 0.0139  | 0.0076  | 0.0050  | 0.0035  | 0.0027  |
| Coverage    | 0.8680  | 0.8720  | 0.8780  | 0.8840  | 0.8780  | 0.8920  |

When the sample size increases, and the coverage percentages are nearly 90%, which means that the asymptotic results can be used to obtain the confidence bounds of the unknown parameters even for moderate sample sizes. Then, it can be observed as expected that the estimations of the frequencies are more accurate than other parameters for almost all of the considered sample sizes. Finally, it can be seen from Table 4 that the NLLS$_1$-based estimators proposed for the estimations of noise variances also work well.

In the case of zero-mean multiplicative noise in model (13), after squaring the data, the coherent model parameters including the frequencies, the phases, and the multiplicative noise variances, can be estimated by NLLS$_2$ proposed in Eq. (10), after that the noncoherent model parameters including the total noise variance and the additive noise variance can be estimated by the NLLS$_2$-based estimators proposed in Eqs (11) and (12), respectively. The procedures of implementation are the same as the case of non-zero mean multiplicative

Table 2
The results of the phases

| Sample size | 50     | 100    | 150    | 200    | 250    | 300    |
|-------------|--------|--------|--------|--------|--------|--------|
| $\phi_1^0$  | 2.0000 | 2.0000 | 2.0000 | 2.0000 | 2.0000 | 2.0000 |
| AE          | 2.0412 | 2.0304 | 2.0114 | 2.0112 | 2.0078 | 2.0009 |
| MSE         | 0.1801 | 0.0745 | 0.0456 | 0.0333 | 0.0262 | 0.0213 |
| AVAR        | 0.1200 | 0.0600 | 0.0400 | 0.0300 | 0.0240 | 0.0200 |
| Length      | 1.1362 | 0.8034 | 0.6560 | 0.5681 | 0.5081 | 0.4639 |
| Coverage    | 0.8660 | 0.8860 | 0.8740 | 0.8820 | 0.8840 | 0.8940 |
| $\phi_2^0$  | 3.0000 | 3.0000 | 3.0000 | 3.0000 | 3.0000 | 3.0000 |
| AE          | 3.0141 | 2.9821 | 2.9895 | 3.0710 | 2.9921 | 2.9988 |
| MSE         | 0.0266 | 0.0150 | 0.0103 | 0.0061 | 0.0059 | 0.0040 |
| AVAR        | 0.0200 | 0.0100 | 0.0067 | 0.0050 | 0.0040 | 0.0033 |
| Length      | 0.4639 | 0.3280 | 0.2678 | 0.2319 | 0.2074 | 0.1894 |
| Coverage    | 0.8800 | 0.8820 | 0.8880 | 0.8900 | 0.8920 | 0.8980 |
### Table 3
The results of the multiplicative noise means

| Sample size | 50   | 100  | 150  | 200  | 250  | 300  |
|-------------|------|------|------|------|------|------|
| $\mu_{11}$  | 1.000| 1.000| 1.000| 1.000| 1.000| 1.000|
| AE          | 1.110| 1.094| 1.051| 1.031| 1.010| 1.005|
| MSE         | 0.061| 0.030| 0.019| 0.015| 0.012| 0.009|
| AVAR        | 0.050| 0.025| 0.017| 0.012| 0.010| 0.008|
| Length      | 0.733| 0.519| 0.423| 0.367| 0.328| 0.299|
| Coverage    | 0.886| 0.914| 0.876| 0.902| 0.898| 0.902|
| $\mu_{12}$  | 2.000| 2.000| 2.000| 2.000| 2.000| 2.000|
| AE          | 2.105| 2.088| 2.067| 2.014| 2.008| 2.005|
| MSE         | 0.069| 0.030| 0.020| 0.015| 0.013| 0.010|
| AVAR        | 0.060| 0.030| 0.020| 0.015| 0.012| 0.010|
| Length      | 0.803| 0.568| 0.464| 0.401| 0.359| 0.328|
| Coverage    | 0.912| 0.878| 0.908| 0.898| 0.888| 0.906|

### Table 4
The results of the noise variances

| Sample size | 50   | 100  | 150  | 200  | 250  | 300  |
|-------------|------|------|------|------|------|------|
| $\sigma_{11}^2$ | 1.000| 1.000| 1.000| 1.000| 1.000| 1.000|
| AE          | 1.211| 1.068| 1.125| 1.036| 1.013| 1.008|
| MSE         | 0.425| 0.281| 0.164| 0.101| 0.074| 0.058|
| AVAR        | 0.320| 0.160| 0.107| 0.080| 0.064| 0.053|
| Length      | 1.855| 1.312| 1.071| 0.928| 0.889| 0.794|
| Coverage    | 0.926| 0.910| 0.912| 0.829| 0.886| 0.894|
| $\sigma_{12}^2$ | 2.000| 2.000| 2.000| 2.000| 2.000| 2.000|
| AE          | 2.166| 2.093| 2.085| 2.051| 2.021| 2.012|
| MSE         | 1.340| 1.105| 0.617| 0.490| 0.411| 0.321|
| AVAR        | 1.760| 0.880| 0.586| 0.440| 0.352| 0.293|
| Length      | 4.351| 3.079| 2.512| 2.175| 1.946| 1.776|
| Coverage    | 0.956| 0.924| 0.920| 0.918| 0.912| 0.908|
| $\sigma^2$  | 4.000| 4.000| 4.000| 4.000| 4.000| 4.000|
| AE          | 4.103| 4.066| 4.030| 4.016| 4.008| 4.004|
| MSE         | 0.514| 0.262| 0.166| 0.125| 0.094| 0.073|
| AVAR        | 0.420| 0.210| 0.140| 0.105| 0.084| 0.070|
| Length      | 2.125| 1.503| 1.227| 1.063| 0.951| 0.868|
| Coverage    | 0.884| 0.892| 0.908| 0.912| 0.894| 0.896|
| $\sigma_0^2$ | 1.000| 1.000| 1.000| 1.000| 1.000| 1.000|
| AE          | 0.725| 0.905| 0.833| 0.927| 0.974| 0.982|
| MSE         | 1.616| 1.413| 0.957| 0.724| 0.551| 0.433|
| AVAR        | 2.500| 1.250| 0.833| 0.625| 0.500| 0.417|
| Length      | 5.186| 3.667| 2.994| 2.593| 2.319| 2.177|
| Coverage    | 0.968| 0.946| 0.928| 0.922| 0.912| 0.904|
Figure 1. MSE, AVAR, and ACRB of frequency versus ISNR1; $\hat{\omega}_1$ in solid lines; $\hat{\omega}_2$ in dotted lines; $N = 200$.

Figure 2. MSE, AVAR, and ACRB of phase versus ISNR1; $\hat{\phi}_1$ in solid lines; $\hat{\phi}_2$ in dotted lines; $N = 200$. 
noise mentioned above. Due to the space limitations, the simulation results are just provided in Tables S1–S3 in the supplementary corresponding to the frequencies, the phases, and the noise variances, respectively. It can be observed from Tables S1–S3 that the similar results as the case of nonzero mean multiplicative noise mentioned above hold, which verify the efficiency of the proposed estimators in the case of zero mean multiplicative noise.

In addition, we carry out the following simulations to evaluate the performance of the proposed hybrid stochastic searching algorithm. The MSE is employed for the performance measure. For comparison, the MSE, the A\VAR, and the asymptotic Cramer–Rao bound (ACRB) are included. Notice that the Fisher information matrix is reduced to zero in the case of zero-mean multiplicative noise, hence the ACRB cannot be obtained in this case (see Ghogho et al., 1999b). Thus, we restrict our attention to the case of nonzero-mean multiplicative noise, in which the ACRB has been provided by Francos and Friedlander (1995) and Mao and Bao (1996). It is well known that the signal-to-noise ratio (SNR) is usually valuable for the performance evaluation, and it has been reported from Ghogho et al. (1999b) that the intrinsic SNR referred as ISNR is meaningful for the harmonics in multiplicative and additive noise, hence the ISNR is included in these simulations. In this case, the ISNR of the first harmonic component referred as ISNR1 is defined as 

\[
\text{ISNR} = 10\log_{10}\left(\frac{\mu_0^2}{\sigma_1^2}\right) \text{ (dB)},
\]

and the ISNR2 is similarly defined. We carry out the simulation by the same procedures from the model (13) with the same parameters except varying \(\sigma_1^2\), and then the MSE, A\VAR, and ACRB versus ISNR1 for different coherent parameters can be evaluated. All the results in terms of ISNR1 are shown in Figs. 1–3 corresponding to frequency, phase, and multiplicative noise mean, respectively. The following observations are very clear from the numerical experiments. First, it can...

![Graph showing MSE, A\VAR, and ACRB of multiplicative noise mean versus ISNR1; \(\hat{\mu}_1\) in solid lines; \(\hat{\mu}_2\) in dotted lines; \(N = 200\).](image-url)
be observed that the MSEs are close to the corresponding AVARs as the ISNR1 increases for all the cases. Next, it is interesting to observe that for the cases of both frequency and phase, i.e., nonlinear parameters, the AVARs are almost always less than the corresponding ACRBs, and only one of two frequencies or phases are close to the corresponding ACRB, which also hold for MSEs and ACRBs. Finally, for the case of multiplicative noise mean, i.e., linear parameter, one of two can close to the corresponding ACRB as the ISNR1 increases. On the other hand, we have also carried out the similar simulation in terms of ISNR2, and we do not show the data due to the space limitations, from which the similar results hold.

5. Conclusions

This study focused on the parameter estimation of multicomponent harmonic signals in multiplicative and additive noise, in which the nonzero- and zero-mean multiplicative noise cases were addressed separately based on NLLS. First, the NLLS₁ and NLLS₂ proposed by Ghogho et al. (1999b) were generalized to estimate the coherent model parameters for the multiple harmonics corresponding to the cases of nonzero- and zero-mean multiplicative noise, respectively. Then, the NLLS₁- and NLLS₂-based estimators were proposed to estimate the noncoherent model parameters. Next, by following the techniques of performance analysis proposed by Kundu and Mitra (1999) corresponding to the harmonics only in additive noise, some statistical results for the proposed estimators were theoretically proved, including the strong consistency, the strong convergence rate, and the asymptotic normality. Finally, the numerical results suggested that the asymptotic results can be used even for moderate sample sizes, which practically shows the efficiency of the proposed estimators.

Acknowledgment

The authors would like to thank the anonymous referees for their valuable comments and helpful suggestions.

Funding

The research was supported by the National Nature Science Foundation of China under grants 61071188 and 11126274 and 61203287, the Nature Science Foundation of Hubei Province of China under grants 2009CDB077 and 2011CDB333, and the Fundamental Research Funds for the Central Universities, China University of Geosciences (Wuhan) under grants CUGL100236 and CUGL100239.

Supplementary Data

Supplemental data for this article can be accessed at www.tandfonline.com/lsta.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Trans. Automat. Control* 19(6):716–723.
Bai, Z. D., Rao, C. R., Chow, M., Kundu, D. (2003). An efficient algorithm for estimating the parameters of superimposed exponential signals. *J. Stat. Plan. Infer.* 110(1-2):23–34.
Besson, O., Castanie, F. (1993). On estimating the frequency of a sinusoid in auto-regressive multiplicative noise. *Signal Process.* 30(1):65–83.
Besson, O., Stoica, P. (1999). Nonlinear least-squares approach to frequency estimation and detection for sinusoidal signals with arbitrary envelope. *Digit. Signal Process.* 9(1):45–56.

Bian, J. W., Li, H. W., Peng, H. M. (2009). An efficient and fast algorithm for estimating the frequencies of superimposed exponential signals in zero-mean multiplicative and additive noises. *J. Stat. Comput. Sim.* 79(12):1407–1423.

Bian, J. W., Li, H. W., Peng, H. M. (2011). An efficient and fast algorithm for estimating the frequencies of 2-D superimposed exponential signals in zero-mean multiplicative and additive noise. *J. Stat. Plan. Infer.* 141(3):1277–1289.

Bressler, Y., Macovski, A. (1986). Exact maximum likelihood parameters estimation of superimposed exponential signals in noise. *IEEE Trans. Signal Process.* 34(5):1081–1089.

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

Chan, F. K. W., So, H. C., Lau, W. H., Chan, C. F. (2010). Efficient approach for sinusoidal frequency estimation of gapped data. *IEEE Signal Process. Lett.* 17(6):611–614.

Cohen, G., Francos, J. M. (2002). Least squares estimation of 2-D sinusoids in colored noise: asymptotic analysis. *IEEE Trans. Inf. Theory* 48(8):2243–2252.

Francos, J. M., Friedlander, B. (1995). Bounds for estimation of multicomponent signals with random amplitude and deterministic phase. *IEEE Trans. signal Process.* 43(5):1161–1172.

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

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

Giannakis, G. B., Zhou, G. T. (1995). Harmonics in multiplicative and additive noise: parameter estimation using cyclic statistics. *IEEE Trans. Signal Process.* 43(9):2217–2221.

Hannan, E. J. (1973). The estimation of frequency. *J. Appl. Probab.* 10(3):510–519.

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

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

Kavalieris, L., Hannan, E. J. (1994). Determining the number of terms in a trigonometric regression. *J. Time Ser. Anal.* 15(6):613–625.

Kundu, D., Mitra, A. (1995). Consistent method of estimating superimposed exponential signals. *Scand. J. Statist.* 22(1):73–82.

Kundu, D., Mitra, A. (1999). On asymptotic behavior of the least squares estimators and the confidence intervals of the superimposed exponential signals. *Signal Process.* 72(2):29–139.

Kundu, D., Mitra, A. (2000). Detecting the number of signals for an undamped exponential model using cross-validation approach. *Signal Process.* 80(3):525–534.

Li, H. W., Cheng, Q. S. (1998). Cyclic statistics based approach to harmonic retrieval in multiplicative and additive noise. *Acta Electron. Sinica* 26(7):105–112.

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

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

Mao, Y. C., Bao, Z. (1996). Cramer–Rao bound for harmonics in Gaussian multiplicative and complex Gaussian additive noise. In *proceedings of ICSP*, pp.497–500.

Nandi, S. (2001). *Analyzing some non-stationary signal processing models.* Ph.D. Thesis, Indian Institute of Technology, Kanpur.

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

Quinn, B. G. (1994). Estimating frequency by interpolation using Fourier coefficients. *IEEE Trans. Signal Process.* 42(5):1264–1268.

Quinn, B. G., Hannan, E. J. (2001). *The Estimation and Tracking of Frequency.* New York: Cambridge University Press.
Appendix A: The Proofs of Theorems 1–6 in Case of Nonzero-mean Multiplicative Noise

To prove the Theorems, we need the following lemmas firstly.

Lemma 1. Let $\theta$ and $Q(\theta)$ be same as (4) and define

$$S_{H,H} = \{ \theta : |\omega_k - \omega^0_k| > \delta \text{ or } |\phi_k - \phi^0_k| > \delta \text{ or } |\mu_{sk} - \mu^0_{sk}| > \delta, \text{ for any } k = 1, 2, \ldots, M, \text{ and } \mu_{si} \leq H, \text{ for all } i = 1, 2, \ldots, M \}.$$
If for any \( \delta > 0 \) and for some \( 0 < H < \infty \),
\[
\liminf_{N \to \infty} \inf_{\theta \in S_{\delta,H}} \frac{1}{T} \left[ Q(\theta) - Q(\theta^0) \right] > 0 \text{ a.s.}
\]
then the LSEs, \( \hat{\theta} \), is a strongly consistent estimator of \( \theta^0 \).

**Proof.** The proof can be obtained by contradiction along the same line as Lemma 1 of Wu (1981).

**Lemma 2.** Let \( X(t) \) be a sequence of i.i.d. random variables with mean zero and \( E|X(1)|^2 < \infty \), then for \( k = 0, 1, 2, \ldots \),
\[
\lim_{T \to \infty} \sup_{\gamma} \left| \frac{1}{T^{k+1}} \sum_{t=1}^{T} t^k X(t)e^{\gamma t} \right| = 0 \text{ a.s.}
\]

**Proof.** The result can be obtained by combining Lemma 1 of Kundu and Mitra (1999) and Lemma 2.5 of Nandi (2001, p. 27).

**Lemma 3.** (Mangulis, 1965) For \( \gamma \neq 0 \) and \( k = 0, 1, 2, \ldots \),
\[
\lim_{T \to \infty} \left| \frac{1}{T^{k+1}} \sum_{t=1}^{T} t^k e^{\gamma t} \right| = 0
\]

**Proof of Theorem 1.** Note that the set \( S_{\delta,H} \), defined in Lemma 1, can be written as
\[
S_{\delta,H} = U_1 \cup U_2 \cup \ldots \cup U_M \cup \Phi_1 \cup \Phi_2 \cup \ldots \cup \Phi_M \cup W_1 \cup W_2 \cup \ldots \cup W_M
\]

where, for \( k = 1, 2, \ldots, M \),
\[
U_{sk} = \{ \theta : |\mu_{sk} - \mu_{sk}^0| > \delta, \mu_{si} \leq H, \text{ for all } i = 1, 2, \ldots, M \},
\]
\[
\Phi_k = \{ \theta : |\phi_k - \phi_k^0| > \delta, \mu_{si} \leq H, \text{ for all } i = 1, 2, \ldots, M \},
\]
\[
W_k = \{ \theta : |\omega_k - \omega_k^0| > \delta, \mu_{si} \leq H, \text{ for all } i = 1, 2, \ldots, M \}
\]

By some calculations, the following expression can be obtained
\[
\frac{1}{T} \left[ Q(\theta) - Q(\theta^0) \right] = f_T(\theta, \theta^0) + g_T(\theta, \theta^0), \quad (A.1)
\]

where
\[
f_T(\theta, \theta^0) = \frac{1}{T} \sum_{t=1}^{T} \left\{ \sum_{k=1}^{M} \left[ \mu_{sk}^0 \cos (\omega_k^0 t + \phi_k^0) - \mu_{sk} \cos(\omega_k t + \phi_k) \right] \right\}^2
\]
\[
+ \frac{1}{T} \sum_{t=1}^{T} \left\{ \sum_{k=1}^{M} \left[ \mu_{sk}^0 \sin (\omega_k^0 t + \phi_k^0) - \mu_{sk} \sin(\omega_k t + \phi_k) \right] \right\}^2,
\]
Statistical Analysis of Non-Linear Least Squares

\[ g_T(\theta, \theta^0) = \frac{2}{T} \sum_{t=1}^{T} \sum_{k=1}^{M} \varepsilon_r(t) \left[ \mu_{sk}^0 \cos(\omega_k^0 t + \phi_k^0) - \mu_{sk} \cos(\omega_k t + \phi_k) \right] + \frac{2}{T} \sum_{t=1}^{T} \sum_{k=1}^{M} \varepsilon_c(t) \left[ \mu_{sk}^0 \sin(\omega_k^0 t + \phi_k^0) - \mu_{sk} \sin(\omega_k t + \phi_k) \right], \]

where, \( \varepsilon_r(t) \) and \( \varepsilon_c(t) \) denote the real and imaginary parts, respectively, of \( \varepsilon(t) \) that is the noncoherent component in data \( x(t) \) and defined in (4). So, for any \( \delta > 0 \) and a fixed \( 0 < H < \infty \), we have, for \( j = 1, 2, \ldots, M \),

\[ \lim_{T \to \infty} \inf_{\theta \in U_{sk}} f_T(\theta, \theta^0) \geq \lim_{T \to \infty} \inf_{|\mu_{sk} - \mu_{sk}^0| > \delta} \frac{1}{T} \sum_{t=1}^{T} \left[ \mu_{sk}^0 \cos(\omega_k^0 t + \phi_k^0) - \mu_{sk} \cos(\omega_k t + \phi_k) \right]^2 + \lim_{T \to \infty} \inf_{|\mu_{sk} - \mu_{sk}^0| > \delta} \frac{1}{T} \sum_{t=1}^{T} \left[ \mu_{sk}^0 \sin(\omega_k^0 t + \phi_k^0) - \mu_{sk} \sin(\omega_k t + \phi_k) \right]^2 \]

\[ = \lim_{T \to \infty} \inf_{|\mu_{sk} - \mu_{sk}^0| > \delta} \frac{1}{T} \sum_{t=1}^{T} \left[ \left( \mu_{sk}^0 - \mu_{sk} \right) \cos(\omega_k^0 t + \phi_k^0) \right]^2 + \left[ \left( \mu_{sk}^0 - \mu_{sk} \right) \sin(\omega_k^0 t + \phi_k^0) \right]^2 \]

\[ = \lim_{T \to \infty} \inf_{|\mu_{sk} - \mu_{sk}^0| > \delta} \left( \mu_{sk}^0 - \mu_{sk} \right)^2 \]

\[ = \delta^2 > 0 \quad \text{a.s.} \quad (A.2) \]

Similarly, it can be proved that, for \( k = 1, 2, \ldots, M \),

\[ \lim_{T \to \infty} \inf_{\theta \in U_k} f_T(\theta, \theta^0) > 0 \quad \text{a.s.} \quad (A.3) \]

\[ \lim_{T \to \infty} \inf_{\theta \in W_k} f_T(\theta, \theta^0) > 0 \quad \text{a.s.} \quad (A.4) \]

From (A.2)–(A.4), it can be observed that

\[ \lim_{T \to \infty} \inf_{\theta \in S_{\delta, H}} f_T(\theta, \theta^0) > 0 \quad \text{a.s.} \quad (A.5) \]

Using Lemma 2, it can be easily shown that

\[ \lim_{T \to \infty} \sup_{\theta \in S_{\delta, H}} g_T(\theta, \theta^0) = 0 \quad \text{a.s.} \quad (A.6) \]

Therefore, combining (A.1), (A.5), (A.6), and Lemma 1, the result of Theorem 1 follows.

**Proof of Theorem 2.** We use the following notation: \( Q'(\theta) = (Q'_1(\theta), Q'_2(\theta), \ldots, Q'_M(\theta)) \), where \( Q'_k(\theta) = \left( \frac{\partial Q(\theta)}{\partial \omega_k}, \frac{\partial Q(\theta)}{\partial \phi_k}, \frac{\partial Q(\theta)}{\partial \mu_{sk}} \right) \), for \( k = 1, 2, \ldots, M \). Similarly, let \( Q''(\theta) \) be a \( M \times M \) block matrix whose block element is given by \( Q''_{kl}(\theta) = \frac{\partial^2 Q(\theta)}{\partial \omega_k \partial \omega_l} \), for \( k, l = 1, 2, \ldots, M \).
Expanding \( Q'(\hat{\theta}) \) around \( \theta^0 \) by using multivariate Taylor series expansion up to the first-order term, it can be obtained that
\[
Q'(\hat{\theta}) - Q'(\theta^0) = (\hat{\theta} - \theta^0)Q''(\bar{\theta}),
\] (A.7)
where \( \bar{\theta} \) is a point between \( \hat{\theta} \) and \( \theta^0 \). Since \( Q'(\hat{\theta}) = 0 \), (A.7) implies that
\[
(\hat{\theta} - \theta^0)D\sqrt{T} = -\left[ Q'(\theta^0)D^{-1} \right] \left[ D^{-1}Q''(\bar{\theta})D^{-1} \right]^{-1},
\] (A.8)
where \( D \) is defined in Theorem 3.

By some regular but complicated calculations, we can obtain the expressions of the elements of \( Q'(\theta^0) \) and \( Q''(\theta^0) \) in details. Note that due to space limitations, we do not show these expressions here and just provide them in the supplementary information. Now by combining Lemma 2, it can be shown that
\[
\lim_{T \to \infty} Q'(\theta^0)D^{-1}\sqrt{T} = 0 \quad \text{a.s.}
\] (A.9)

By combining Lemma 3 and the fact that \( \bar{\theta} \) converges to \( \theta^0 \) almost surely, it also can be shown that
\[
\lim_{T \to \infty} D^{-1}Q''(\bar{\theta})D^{-1} = \lim_{T \to \infty} D^{-1}Q''(\theta^0)D^{-1} = \Sigma^{(1)} \quad \text{a.s.}
\] (A.10)
where \( \Sigma^{(1)} = \text{diag}\{\Sigma_1^{(1)}, \Sigma_2^{(1)}, \ldots, \Sigma_M^{(1)}\} \) is a \( M \)-dimensional block diagonal matrix whose block element is given as follows, for \( k = 1, 2, \ldots, M \),
\[
\Sigma_k^{(1)} = \begin{bmatrix}
\frac{1}{2}(\mu_{sk}^0)^2 & 2 & 3 \\
2 & 3 & 6 \\
3 & 6 & 0 \\
0 & 2 & 0
\end{bmatrix}
\]
Notice that \( \Sigma^{(1)} \) is a positive definite matrix.

Therefore, by combining (A.8)–(A.10), it can be obtained that \( (\hat{\theta} - \theta^0)D\sqrt{T} \) converges to zero vector almost surely, which implies the result of Theorem 2. □

Proof of Theorem 3. Referring to the proof of Theorem 2, it can be obtained from (A.8) that
\[
(\hat{\theta} - \theta^0)D = -\left[ Q'(\theta^0)D^{-1} \right] \left[ D^{-1}Q''(\bar{\theta})D^{-1} \right]^{-1}
\] (A.11)

On the one hand, by combining the expressions of the elements of \( Q'(\theta^0) \) presented in the proof of Theorem 2 and Lindberg–Feller central limit theorem (see Rao, 1973, p. 128), it can be shown that
\[
Q'(\theta^0)D^{-1} \to \mathcal{N}_{3M}(0, \Sigma^{(2)}),
\] (A.12)
where \( \mathcal{N}_3(0, \Sigma^{(2)}) \) denotes the \( 3M \)-variate normal distribution with mean vector zero and covariance matrix \( \Sigma^{(2)} \), herein the covariance matrix \( \Sigma^{(2)} = \text{diag}\{\Sigma_1^{(2)}, \Sigma_2^{(2)}, \ldots, \Sigma_M^{(2)}\} \) is a
and finite variance \( \text{Var}(\sigma_{sk}) \) from (A.14), it also can be shown that \( M \) and Lemma 2, it also can be shown that \( \text{Var}(\sigma_{sk}) \) for space limitations. We observe that the following form,

\[
\sigma_{sk} \text{ variables with mean } \mu_{sk}, \text{ which implies that the first term on the right side of (A.14) converges in distribution to a normal distribution with mean } \theta_0. \]

Therefore, by combining (A.11), (A.12), and (A.10), it shows the asymptotic normality of \((\hat{\theta} - \theta^0)D\) (see Rao, 1973, p. 122), i.e., \((\hat{\theta} - \theta^0)D\) converges in distribution to a \(3M\)-variate normal distribution with mean vector zero and the covariance matrix \( \Sigma = \text{diag}(\Sigma_1, \Sigma_2, \ldots, \Sigma_M) \), which is a \(3M\)-dimensional block diagonal matrix whose block element is given by

\[
\Sigma_k = (\Sigma_k^{(1)})^{-1} \Sigma_k^{(2)} (\Sigma_k^{(1)})^{-1} = \begin{bmatrix}
\frac{\sigma^2 - \sigma_{sk}^2}{(\mu_{sk}^0)^2} & (\mu_{sk}^0)^2 & 0 \\
(\mu_{sk}^0)^2 & \frac{\sigma^2 + \sigma_{sk}^2}{2}
\end{bmatrix}
\]

for \(k = 1, 2, \ldots, M.\) The result of Theorem 3 follows. \(\square\)

**Proof of Theorem 4.** To obtain the strong consistency of \(\hat{\sigma}_{sk}^2\) defined in (5), we firstly transform it into the following form,

\[
\hat{\sigma}_{sk}^2 = \frac{1}{T} \sum_{t=1}^{T} [s_{sk}^2(t) - (\mu_{sk}^0)^2] + [(\mu_{sk}^0)^2 - \hat{\mu}_{sk}^2] + R_T(\hat{\theta}_k, \theta^0).
\]

where the expression of \(R_T(\hat{\theta}_k, \theta^0)\) is provided in the supplementary information due to space limitations. We observe that \([s_{sk}^2(t) - (\mu_{sk}^0)^2]\) is a sequence of i.i.d. real random variables with mean \(\sigma_{sk}^2\), which implies that \(\frac{1}{T} \sum_{t=1}^{T} [s_{sk}^2(t) - (\mu_{sk}^0)^2] \rightarrow \sigma_{sk}^2\), a.s. because of Kolmogorov strong law of large numbers (see Rao, 1973, p. 115). It can be shown that \((\mu_{sk}^0)^2 - \hat{\mu}_{sk}^2 \rightarrow 0\), a.s. because of the strong convergence of \(\hat{\mu}_{sk}\). Expanding \(R_{MN}(\hat{\theta}_k, \theta^0)\) around the true value \(\theta^0\) by Taylor series expansion and combining Theorem 1, Theorem 2, and Lemma 2, it also can be shown that \(R_T(\hat{\theta}_k, \theta^0) \rightarrow 0\), a.s. Therefore, combining (A.13) and the results mentioned above, it can be obtained that \(\hat{\sigma}_{sk}^2 \rightarrow \sigma_{sk}^2\), a.s., which proves the result of strong consistency in Theorem 4. \(\square\)

To obtain the asymptotic distribution of \(\hat{\sigma}_{sk}^2\), we firstly transform (A.13) into the following form,

\[
\sqrt{T}(\hat{\sigma}_{sk}^2 - \sigma_{sk}^2) = \sqrt{T} \left\{ \frac{1}{T} \sum_{t=1}^{T} [s_{sk}^2(t) - (\mu_{sk}^0)^2] - \sigma_{sk}^2 \right\} + \sqrt{T} [(\mu_{sk}^0)^2 - \hat{\mu}_{sk}^2] + \sqrt{T} R_T(\hat{\theta}_k, \theta^0)
\]

(A.14)

We observe that \([s_{sk}^2(t) - (\mu_{sk}^0)^2]\) is a sequence of i.i.d. real random variables with mean \(\sigma_{sk}^2\) and finite variance \(\text{Var}[s_{sk}^2(t)]\) under Assumptions 1 and 2, which implies that the first term on the right side of (A.14) converges in distribution to a normal distribution with mean
zero and variance \( \text{Var}[s_k^2(1)] \), because of Lindberg–Levy central limit theorem (see Rao, 1973, p. 127). Combining the strong consistency and asymptotic normality of \( \hat{\mu}_{sk} \) and the limit properties of sequence of random variables (see Rao, 1973, p. 122), it can be shown that the second term converges in distribution to a normal distribution with mean zero and variance \( 2(\sigma^2 + \sigma_{sk}^2)(\mu_{sk}^0)^2 \). After some calculations, it also can be shown by using Lemma 3 that the third term converges to zero in distribution. Therefore, combining (A.14) and the results mentioned above, it can be obtained that \( \sqrt{T}(\hat{\sigma}_{sk}^2 - \sigma_{sk}^2) \) converges in distribution to a normal distribution with mean zero and variance \( \sigma^2 \). This proves the result of asymptotic normality in Theorem 4.

**Proof of Theorem 5.** To obtain the strong consistency of \( \hat{\sigma}^2 \) defined in (6), we firstly transform it into the following form by referring to the proof of Theorem 1,

\[
\hat{\sigma}^2 = \frac{1}{T} Q(\hat{\theta}) = \frac{1}{T} \sum_{t=1}^{T} |\varepsilon(t)|^2 + f_T(\hat{\theta}, \theta^0) + g_T(\hat{\theta}, \theta^0),
\]

where \( \varepsilon(t) \) is defined in (3). We note that \( |\varepsilon(t)|^2 \) is a sequence of independent real random variables. After some regular calculations, it can be shown that

\[
\text{E}[|\varepsilon(t)|^2] = \sigma_0^2 + \sum_{k=1}^{p} \sigma_{sk}^2 = \sigma^2
\]

\[
\text{Var}[|\varepsilon(t)|^2] = \text{Var}[|n(1)|^2] + \sum_{k=1}^{M} \text{Var}[|s_k(1) - \mu_{sk}^0|^2] + 2\sigma_0^2 \sum_{k=1}^{M} \sigma_{sk}^2
\]

\[
+ 2 \sum_{k \neq l} \sigma_{sk}^2 \sigma_{sl}^2 \cos^2 (\omega_k^0 t + \phi_k^0 - \omega_l^0 t - \phi_l^0)
\]

By Lemma 3, it can be easily seen that \( \sum_{t=1}^{\infty} \frac{\text{Var}[|\varepsilon(t)|^2]}{t^2} < \infty \). Therefore, under Assumptions 1 and 2, \( \frac{1}{T} \sum_{t=1}^{T} |\varepsilon(t)|^2 \to \sigma^2 \), a.s. because of Kolmogorov strong law of large numbers (see Rao, 1973, p. 114). On the other hand, it can be shown that

\[
\frac{1}{T} \sum_{t=1}^{T} \left\{ \sum_{k=1}^{M} \left[ \mu_{sk}^0 \cos (\omega_k^0 t + \phi_k^0) - \hat{\mu}_{sk} \cos (\hat{\omega}_k t + \hat{\phi}_k) \right] \right\}^2
\]

\[
\leq \frac{M}{T} \sum_{k=1}^{M} \sum_{i=1}^{T} \left[ (\mu_{sk}^0 - \hat{\mu}_{sk})^2 \cos^2 (\omega_k^0 t + \phi_k^0) + \hat{\mu}_{si}^2 \left[ \cos (\omega_k^0 t + \phi_k^0) - \cos (\hat{\omega}_k t + \hat{\phi}_k) \right]^2 \right]
\]

\[
\leq 2M \sum_{k=1}^{M} (\mu_{sk}^0 - \hat{\mu}_{sk})^2 + \frac{2M}{T} \sum_{k=1}^{M} \sum_{i=1}^{T} \hat{\mu}_{sk}^2 (\omega_k^0 t + \phi_k^0 - \hat{\omega}_k t - \hat{\phi}_k)^2
\]

\[\to 0, \text{ a.s.}\]

where the first inequality follows from Cauchy–Schwartz inequality, the second inequality follows from the fact \( 2x^2 + 2y^2 - (x + y)^2 \geq 0 \), the third inequality follows from Taylor
series expansion, and the last step follows from Theorems 1 and 2. Along the same line, it also can be shown that

\[
1 \sum_{t=1}^{T} \left\{ \sum_{k=1}^{M} \left[ \mu_{sk}^0 \sin (\omega_k^0 t + \phi_k^0) - \hat{\mu}_{sk} \sin (\hat{\omega}_kt + \hat{\phi}_k) \right] \right\}^2 \to 0, \text{ a.s.}
\]

So, it follows that \( f_T(\hat{\theta}, \theta^0) \to 0, \text{ a.s.} \). Moreover, by Lemma 2, it can be easily shown that \( g_T(\hat{\theta}, \theta^0) \to 0, \text{ a.s.} \). Therefore, combining (A.15) and the results mentioned above, it can be obtained that \( \hat{\sigma}^2 \to \sigma^2, \text{ a.s.} \), which proves the result of strong consistency in Theorem 5.

To obtain the asymptotic distribution of \( \hat{\sigma}^2 \), we firstly transform (A.15) into the following form,

\[
\sqrt{T}(\hat{\sigma}^2 - \sigma^2) = \sqrt{T} \left\{ \frac{1}{T} \sum_{t=1}^{T} |\varepsilon(t)|^2 - \sigma^2 \right\} + \sqrt{T} f_T(\hat{\theta}, \theta^0) + \sqrt{T} g_T(\hat{\theta}, \theta^0) \quad (A.16)
\]

We observe that \( |\varepsilon(t)|^2 \) is a sequence of independent real random variables. Referring to the above proof, it can be obtained by using Lemma 3 that

\[
\mathbb{E}[|\varepsilon(t)|^2] = \sigma^2 \quad \text{and} \quad \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \text{Var}[|\varepsilon(t)|^2] = \sigma^*\quad,
\]

where \( \sigma^* \) is defined in Theorem 5. Using Lindberg–Feller central limit theorem (see Rao, 1973, p. 128), it can be shown that the first term on the right side of (A.16) converges in distribution to a normal distribution with mean zero and variance \( \sigma^* \). After some calculations and using Lemma 3, it also can be shown that both the second and third terms converge to zero in distribution. Therefore, combining (A.16) and the results mentioned above, it can be obtained that \( \sqrt{MN}(\hat{\sigma}^2 - \sigma^2) \) converges in distribution to a normal distribution with mean zero and variance \( \sigma^* \), which proves the result of asymptotic normality in Theorem 5.

**Proof of Theorem 6.** The proof can be naturally obtained by combining (7), the strong consistency and the asymptotic normal distributions of \( \hat{\sigma}^2 \) and \( \hat{\sigma}_{sk}^2 \)’s.

**Appendix B: The Proofs of Theorems 7–11 in Case of Zero-Mean Multiplicative Noise**

**Proof of Theorem 7.** It can be proved similarly as Theorem 1.

**Proof of Theorem 8.** It can be proved similarly as Theorem 2. It is worth noting that the major difference lies in the positive definite limit matrix \( \Sigma^{(1)} \) in (A.10), due to the difference of \( Q'(\theta_0) \) between the cases of nonzero mean and zero mean multiplicative noise that have been provided in the supplementary for the purpose of comparison. In this case, \( \Sigma^{(1)} \) is also a \( M \)-dimensional block diagonal matrix whose block element is given as follows, for \( k = 1, 2, \ldots, M \),

\[
\Sigma^{(1)}_k = \begin{bmatrix}
\frac{4}{5} (\sigma_{sk})^4 & \frac{2}{3} \frac{3}{6} & 0 \\
0 & 2
\end{bmatrix}
\]

\( \square \)
Proof of Theorem 9. It can be proved similarly as Theorem 3. It is worth noting that the major differences lie in the covariance matrix $\Sigma^{(2)}$ in (A.12), due to the difference of $Q'(\theta^0)$ between the cases of nonzero mean and zero mean multiplicative noise that have been provided in the supplementary for the purpose of comparison. In this case, $\Sigma^{(2)}$ is also a $M$-dimensional block diagonal matrix whose block element is given as follows, for $k = 1, 2, \ldots, M$,

$$
\Sigma^{(2)}_k = \begin{bmatrix}
\frac{4}{7}\sigma^4_{sk} A_k \begin{pmatrix} 2 & 3 \\ 3 & 6 \end{pmatrix} 0 \\
0 & 2B_k
\end{bmatrix},
$$

where the expressions of $A_k$’s and $B_k$’s are provided in Theorem 9. Therefore, by combining $\Sigma^{(2)}$ and $\Sigma^{(1)}$ given in the proof of Theorem 8, it can be obtained that the covariance matrix $\Sigma$ in this case is also a $M$-dimensional block diagonal matrix whose block element is given by, for $k = 1, 2, \ldots, M$,

$$
\Sigma_k = \left( \Sigma^{(1)}_k \right)^{-1} \Sigma^{(2)}_k \left( \Sigma^{(1)}_k \right)^{-1} = \begin{bmatrix}
\frac{A_k}{4\sigma^4_{sk}} \begin{pmatrix} 6 & -3 \\ -3 & 2 \end{pmatrix} 0 \\
0 & B_k
\end{bmatrix}
$$

□

Proof of Theorem 10. Notice that $|x(t)|^2$ is a sequence of independent real random variables. After some regular calculations it can be shown that

$$
\begin{align*}
E[|x(t)|^2] &= \sigma_0^2 + \sum_{k=1}^{M} \sigma^2_{sk} = \sigma^2 \\
\text{Var}[|x(t)|^2] &= \text{Var}[|n(1)|^2] + \sum_{k=1}^{M} \text{Var}[|s_k(1)|^2] + 2\sigma_0^2 \sum_{k=1}^{M} \sigma^2_{sk} \\
&\quad + 2 \sum_{k \neq l} \sigma^2_{sk} \sigma^2_{sl} \cos^2(\omega^0_k t + \phi^0_k - \omega^0_l t - \phi^0_l)
\end{align*}
$$

By Lemma 3, it can be easily seen that

$$
\sum_{t=1}^{\infty} \frac{\text{Var}[|x(t)|^2]}{t^2} < \infty, \quad \text{and} \quad \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \text{Var}[|x(t)|^2] = \sigma^*.
$$

where

$$
\sigma^* = \text{Var}[|n(1)|^2] - \sigma_0^4 + \sum_{k=1}^{M} \left[ \text{Var}[|s_k(1)|^2] - \sigma^4_{sk} \right] + \sigma^4
$$

Therefore, the results of Theorem 10, i.e., the strong convergence and asymptotic normality of $\hat{\sigma}$, follow from Kolmogorov strong law of large numbers (see Rao, 1973, p. 114) and Lindberg–Feller central limit theorem (see Rao, 1973, p. 128), respectively. □

Proof of Theorem 11. It can be proved similarly as Theorem 6. □