ESTIMATING PARAMETERS IN MULTICHANNEL FUNDAMENTAL FREQUENCY WITH HARMONICS MODEL

SWAGATA NANDI[1] AND DEBASIS KUNDU[2]

Abstract. In this paper, we introduce a special multichannel model in the class of multichannel sinusoidal model. In multichannel sinusoidal model, the inherent frequencies from distinct channels are same with different amplitudes. The underline assumption here is that there is a fundamental frequency which is same in each channel and the effective frequencies are harmonics of this fundamental frequency. We name this model as multichannel fundamental frequency with harmonics model. It is assumed that the errors in individual channel are independently and identically distributed whereas the signal from different channels are correlated. We propose generalized least squares estimators which become the maximum likelihood estimators also, when the error distribution of the different channels follows a multivariate Gaussian distribution. The proposed estimators are strongly consistent and asymptotically normally distributed. We have provided the implementation of the generalized least squares estimators in practice. Special attention has been taken when the number of channels is two and both have equal number of components. Simulation experiments have been carried out to observe the performances of the proposed estimators. Real data sets have been analyzed using a two-channel fundamental frequency model.

1. Introduction

The problem of estimation of parameters in harmonic signals, whose frequencies are integer multiples of an inherent fundamental frequency is an important problem in many areas of science and technology. It is required in wide range of applications, for example, music classification, compression of audio and voiced speech [19], [10], biomedical research for human circadian system ([2], [3]), modeling male voice sound [12], and so on. The problem of estimation of the fundamental frequency along with the other parameters are of interest. But many times, the frequencies of sinusoidal components from different sources are close to each other and in such a situation a multichannel set-up is more useful. Apart from that

2000 *Mathematics Subject Classification.* 62J02; 62E20; 62C05.

Key words and phrases. Multichannel model; Fundamental frequency model; Least squares estimator; Generalized least squares estimators; consistency; asymptotic normality.

if these frequencies are harmonics of a fundamental frequency, a multichannel fundamental frequency model is more appropriate to use in practice.

The problem of finding the unknown parameters of sinusoidal signals from multichannel outputs has received a considerable amount of attention in recent times. Starting from Sakai [17], researchers were interested in different forms of multichannel sinusoidal signals. Tzagkarakis and Mouchtaris [18] proposed a multichannel version of sinusoids plus noise model and applied to spot microphone signals of a music recording. Griffin et al. [6] applied compressed sensing to multichannel audio coding. Zhou, So and Christensen [23] addressed the problem of parametric modeling of multichannel damped sinusoidal signals utilizing the shift invariance property of the signal subspace under the assumption of white Gaussian noise sequences. Chan, So and Sun [4] addressed the parameter estimation of exponentially damped sinusoids in white noise using multichannel measurements. Zhou et al. [22] discussed a robust estimation method of the parameters of multichannel sinusoidal signals. Nandi and Kundu [14] considered a multichannel sinusoidal model with single component under channel-wise correlated error and discussed the problem of estimation of the unknown parameters and the asymptotic properties of the proposed estimators.

In this paper we are going to introduce a multichannel sinusoidal model with a fundamental frequency. Different channels may have different number of harmonics. A M-channel model with $p_1, \ldots, p_M$ numbers of harmonics can be written as

$$\begin{pmatrix} y_1(t) \\ \vdots \\ y_M(t) \end{pmatrix} = \begin{pmatrix} \mu_1(t; \boldsymbol{\alpha}_1^0, \lambda^0) \\ \vdots \\ \mu_M(t; \boldsymbol{\alpha}_M^0, \lambda^0) \end{pmatrix} + \begin{pmatrix} e_1(t) \\ \vdots \\ e_M(t) \end{pmatrix}; t = 1, \ldots, n. \tag{1}$$

Here

$$\mu_m(t; \boldsymbol{\alpha}_m^0, \lambda^0) = \sum_{j=1}^{p_m} \left[A_{mj}^0 \cos(j\lambda^0 t) + B_{mj}^0 \sin(j\lambda^0 t) \right]; \qquad m = 1, \ldots, M,$$

$0 < \lambda^0 < \pi/p_M$ is the fundamental frequency and $\boldsymbol{\alpha}_m^0 = (A_{m1}^0, B_{m1}^0, \ldots, A_{mp_m}^0, B_{mp_m}^0)^\top$; for $m = 1, \ldots, M$ are the amplitude parameters associated with the m-th channel. Let us use the following notations; $\boldsymbol{\alpha}^0 = (\boldsymbol{\alpha}_1^{0\top}, \ldots, \boldsymbol{\alpha}_M^{0\top})^\top$ and $\boldsymbol{\xi}^0 = (\boldsymbol{\alpha}^{0\top}, \lambda^0)^\top$. Without loss of generality, it is assumed that $p_1 \leq \ldots \leq p_M$ and $p_M << n$. Note that the restriction on the fundamental frequency namely $0 < \lambda^0 < \pi/p_M$ is a very natural restriction. Even in a single channel multicomponent sinusoidal frequency model which has the following mean function

$$\mu(t) = \sum_{j=1}^{K} \left\{ A_j \cos(\omega_j t) + B_j \sin(\omega_j t) \right\}; \quad t = 1, \ldots, n,$$

has the following restrictions on the frequencies: $0 < \omega_1 \neq \omega_2 \cdots \neq \omega_K < \pi$, due to the periodic nature of the function, see for example Hannan [8]. Hence, for the fundamental frequency with harmonics model, the above restriction is natural as it was originally assumed by Quinn and Thomson [15], and since then by all the other authors also who have considered this model, see for example Chapter 5 of Nandi and Kundu [13].

The model (1) can be written in a vector form based on the following notations, $\boldsymbol{y}(t) = (y_1(t), \ldots, y_M(t))^\top$, $\boldsymbol{\mu}^0(t) = \Big(\mu_1(t; \boldsymbol{\alpha}_1^0, \lambda^0), \ldots, \mu_M(t; \boldsymbol{\alpha}_M^0, \lambda^0)\Big)^\top$ and $\boldsymbol{e}(t) = (e_1(t), \ldots, e_M(t))^\top$, as follows:

$$\boldsymbol{y}(t) = \boldsymbol{\mu}^0(t) + \boldsymbol{e}(t); \quad t = 1, \ldots n.$$

The problem is to estimate the unknown parameters namely $\{\boldsymbol{\alpha}_m^0; m = 1, \ldots, M\}$ and λ^0, based on the sample $\{\boldsymbol{y}(t); t = 1, \ldots, n\}$ under suitable assumptions on $\{\boldsymbol{e}(t); t = 1, \ldots, n\}$.

Nandi and Kundu [14], had considered multichannel one component sinusoidal model, whereas in this manuscript we have considered multichannel fundamental frequency with harmonics model. As it has been observed in one channel model also that even though the fundamental frequency with harmonics model has less number of non-linear parameters than the sum of sinusoidal model, it needs special attention. The asymptotic properties of the least squares estimators of the two models are quite different and one cannot be obtained from the other, see for example the Chapter 4 of Nandi and Kundu [13]. That is the main purpose of this paper.

Since, it is assumed $p_1 \leq \ldots \leq p_M$, $\boldsymbol{\mu}^0(t)$ can be expressed as

$$\boldsymbol{\mu}^0(t) = \sum_{j=1}^{p_1} \begin{pmatrix} A_{1j}^0 & B_{1j}^0 \\ A_{2j}^0 & B_{2j}^0 \\ \vdots & \vdots \\ A_{Mj}^0 & B_{Mj}^0 \end{pmatrix} \begin{pmatrix} \cos(j\lambda^0 t) \\ \sin(j\lambda^0 t) \end{pmatrix} + \sum_{j=p_1+1}^{p_2} \begin{pmatrix} 0 & 0 \\ A_{2j}^0 & B_{2j}^0 \\ \vdots & \vdots \\ A_{Mj}^0 & B_{Mj}^0 \end{pmatrix} \begin{pmatrix} \cos(j\lambda^0 t) \\ \sin(j\lambda^0 t) \end{pmatrix} + \ldots + .$$

$$\sum_{j=p_{M-1}+1}^{p_M} \begin{pmatrix} 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \\ A_{Mj}^0 & B_{Mj}^0 \end{pmatrix} \begin{pmatrix} \cos(j\lambda^0 t) \\ \sin(j\lambda^0 t) \end{pmatrix}.$$

In model (1), the M-variate random vector $\boldsymbol{y}(t)$ represents the signal from the M channels at time point t; λ^0 is the fundamental frequency and in the first channel p_1 number of harmonics of λ^0 is present, similarly, in the second channel, p_2 number of harmonics is present, and so on. A frequency $j\lambda^0$ is said to be the j^{th} harmonics of the fundamental frequency λ^0. For

$j = 1, \ldots, p_m$, A^0_{mj} and B^0_{mj} denote the amplitudes corresponding to the j^{th} harmonics in the m-th channel for $m = 1, \ldots, M$. The elements of the random vector $\boldsymbol{e}(t)$ represent the channels-wise noise parts and its structure is stated in Assumption 1.

Assumption 1. *The error term $\boldsymbol{e}(t) = (e_1(t), \ldots, e_M(t))^\top$ is a sequence of independent and identically distributed (i.i.d.) random vectors with mean vector $\mathbf{0}$ and the dispersion matrix $\boldsymbol{\Sigma} = ((\sigma_{ij}))$. It is assumed that $\boldsymbol{\Sigma}^{-1} = ((\sigma^{ij}))$ exists.*

The main aim of this paper is to propose the generalized least squares estimators (GLSEs) of the unknown parameters. It may be mentioned that under the assumption of multivariate normality of the error components, the GLSEs become the maximum likelihood estimators also. First it is assumed that the dispersion matrix $\boldsymbol{\Sigma}$ is known and we have developed the GLSEs of the unknown parameters. It is observed that the proposed estimators can be obtained by solving only one non-linear equation. We have developed the consistency and the asymptotic normality properties of the GLSEs. We have performed an extensive simulations to show the effectiveness of the proposed estimators. The performances are along the expected lines.

Since in practice, the dispersion matrix $\boldsymbol{\Sigma}$ may not be known, hence, we cannot use the GLSEs directly. In practice we propose to use a two-step procedure. In the first step we obtain a consistent estimator of $\boldsymbol{\Sigma}$ based on the least squares estimators (LSEs) of $\boldsymbol{\xi}$. To compute the LSEs of $\boldsymbol{\xi}$, one does not need to know $\boldsymbol{\Sigma}$. Now to compute the GLSEs of $\boldsymbol{\xi}$, we have used this consistent estimator of $\boldsymbol{\Sigma}$ as a plug-in estimator. It has been explained in details in the data analysis section.

The two-channel model when the number of harmonics is same has several applications in practice. If $p_1 = p_2 = p$, the mean vector of the two-channel model takes the following form.

$$\boldsymbol{\mu}^0(t) = \sum_{j=1}^{p} \begin{pmatrix} A^0_{1j} & B^0_{1j} \\ A^0_{2j} & B^0_{2j} \end{pmatrix} \begin{pmatrix} \cos(j\lambda^0 t) \\ \sin(j\lambda^0 t) \end{pmatrix}. \tag{2}$$

We have discussed this special case in details. It has several applications in different fields, see for example Christensen [5], Handel and Host-Madsen [7], Nandi and Kundu [14] and the references cited there in. Further, it is observed that in this case some of the calculations can be performed more explicitly, and the asymptotic results can be written in a more compact form.

The rest of the article is organized as follows. In Section 2, we consider the GLSEs of the model (1). We have provided the methodologies, asymptotic properties of the estimators and the implementation of the GLSEs in practice. A special case when $M = 2$, and $p_1 = p_2$ has been discussed in Section 3. In Section 4, we provide numerical experiment results based on simulations. Analyses of real datasets are discussed in Section 5 and finally in Section 6, we conclude the paper. The proofs are provided in Appendices.

2. Generalized Least Squares Estimators

2.1. **Methodology.** The GLSEs of the unknown parameters can be obtained by minimizing

$$S(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_M, \lambda) = S(\boldsymbol{\xi}) = |\boldsymbol{\Sigma}| \sum_{t=1}^{n} (\boldsymbol{y}(t) - \boldsymbol{\mu}(t))^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{y}(t) - \boldsymbol{\mu}(t)).$$

Here it is assumed that $\boldsymbol{\Sigma}$ is known. We use the notation $P = p_1 + \ldots + p_M$ and the GLSE of $\boldsymbol{\xi}$ as $\widehat{\boldsymbol{\xi}}$. Since $\boldsymbol{\Sigma}$ is assumed to be known, hence, minimizing $S(\boldsymbol{\xi})$ is equivalent to minimize $R(\boldsymbol{\xi})$, where

$$R(\boldsymbol{\xi}) = \sum_{t=1}^{n} (\boldsymbol{y}(t) - \boldsymbol{\mu}(t))^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{y}(t) - \boldsymbol{\mu}(t)). \tag{3}$$

First we will show that $\widehat{\boldsymbol{\xi}}$ can be obtained by solving only one one-dimensional optimization problem. Let us write

$$R(\boldsymbol{\xi}) = \sum_{t=1}^{n} (\boldsymbol{y}(t) - \boldsymbol{X}_t(\lambda)\boldsymbol{\alpha})^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{y}(t) - \boldsymbol{X}_t(\lambda)\boldsymbol{\alpha}). \tag{4}$$

Here $\boldsymbol{X}_t(\lambda)$ is a $M \times 2P$ matrix whose m-th row is

$$(\cos(\lambda t) \quad \sin(\lambda t) \quad \cdots \quad \cos(p_1 \lambda t) \quad \sin(p_1 \lambda t) \quad \cdots \quad \cos(p_m \lambda t) \quad \sin(p_m \lambda t) \quad 0 \quad \cdots \quad 0),$$

for $m = 1, \ldots, M$. For given λ, the GLSEs of $\boldsymbol{\alpha}$ can be obtained as

$$\widehat{\boldsymbol{\alpha}}(\lambda) = \left[\sum_{t=1}^{n} \boldsymbol{X}_t^\top(\lambda) \boldsymbol{\Sigma}^{-1} \boldsymbol{X}_t(\lambda)\right]^{-1} \left[\sum_{t=1}^{n} \boldsymbol{X}_t^\top(\lambda) \boldsymbol{\Sigma}^{-1} \boldsymbol{y}(t)\right]. \tag{5}$$

Hence, $\widehat{\alpha}$, the GLSE of λ, can be obtained by minimizing $Q(\lambda)$, where

$$Q(\lambda) = \sum_{t=1}^{n} (\boldsymbol{y}(t) - \boldsymbol{X}_t(\lambda)\widehat{\boldsymbol{\alpha}}(\lambda))^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{y}(t) - \boldsymbol{X}_t(\lambda)\widehat{\boldsymbol{\alpha}}(\lambda)). \tag{6}$$

Once, $\widehat{\lambda}$ is obtained, the GLSEs of $\boldsymbol{\alpha}$ can be obtained as $\widehat{\boldsymbol{\alpha}}(\widehat{\lambda})$. Note that the minimization of $Q(\lambda)$ can be performed by using one dimensional optimization method, for example Newton-Raphson method or bisection method may be used for this purpose. It is well known that the least squares surface has several local minima, see for example Rice and Rosenblatt [16]. Hence, special attention needs to be taken to choose the initial guesses. The Fourier frequencies or the periodogram function are usually used for this purpose. The details are explained in the data analysis section.

If $\boldsymbol{\Sigma} = \boldsymbol{I}$, the identity matrix, or we want to find the least squares estimators (LSEs) of $\boldsymbol{\xi}$, then it can be obtained by minimizing

$$R_1(\boldsymbol{\xi}) = \sum_{t=1}^{n} \left(\boldsymbol{y}(t) - \boldsymbol{\mu}(t)\right)^\top \left(\boldsymbol{y}(t) - \boldsymbol{\mu}(t)\right). \tag{7}$$

Note that (7) can be written as

$$R_1(\boldsymbol{\xi}) = \sum_{j=1}^{M} \left(\boldsymbol{Y}_j - \boldsymbol{Z}_j(\lambda)\boldsymbol{\alpha}_j\right)^\top \left(\boldsymbol{Y}_j - \boldsymbol{Z}_j(\lambda)\boldsymbol{\alpha}_j\right), \tag{8}$$

where $\boldsymbol{Y}_j = (y_j(1), \ldots, y_j(n))^\top$ and

$$\boldsymbol{Z}_j(\lambda) = \begin{bmatrix} \cos(\lambda) & \sin(\lambda) & \ldots & \cos(p_j\lambda) & \sin(p_j\lambda) \\ \cos(2\lambda) & \sin(2\lambda) & \ldots & \cos(2p_j\lambda) & \sin(2p_j\lambda) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \cos(n\lambda) & \sin(n\lambda) & \ldots & \cos(np_j\lambda) & \sin(np_j\lambda) \end{bmatrix}.$$

For fixed λ, the LSEs of $\boldsymbol{\alpha}_j$, say $\widetilde{\boldsymbol{\alpha}}_j(\lambda)$, can be obtained as

$$\widetilde{\boldsymbol{\alpha}}_j(\lambda) = \left(\boldsymbol{Z}_j(\lambda)^\top \boldsymbol{Z}_j(\lambda)\right)^{-1} \boldsymbol{Z}_j^\top(\lambda)\boldsymbol{Y}_j; \qquad j = 1, \ldots, M.$$

Then $\widehat{\lambda}$, LSE of λ^0, can be obtained by minimizing

$$Q_1(\lambda) = \sum_{j=1}^{M} \boldsymbol{Y}_j^\top \left(\boldsymbol{I} - \boldsymbol{Z}_j(\lambda)\left(\boldsymbol{Z}_j(\lambda)^\top \boldsymbol{Z}_j(\lambda)\right)^{-1} \boldsymbol{Z}_j(\lambda)^\top\right) \boldsymbol{Y}_j. \tag{9}$$

If $p_1 = \ldots = p_M = p$, then $\boldsymbol{Z}_1(\lambda) = \cdots = \boldsymbol{Z}_M(\lambda) = \boldsymbol{Z}(\lambda)$, where

$$\boldsymbol{Z}(\lambda) = \begin{bmatrix} \cos(\lambda) & \sin(\lambda) & \ldots & \cos(p\lambda) & \sin(p\lambda) \\ \cos(2\lambda) & \sin(2\lambda) & \ldots & \cos(2p\lambda) & \sin(2p\lambda) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \cos(n\lambda) & \sin(n\lambda) & \ldots & \cos(np\lambda) & \sin(np\lambda) \end{bmatrix}. \tag{10}$$

2.2. **Asymptotic Properties.** In this section we provide the consistency and the asymptotic normality properties of the GLSEs as well as for the LSEs.

Theorem 2.1. *Under Assumption 1, if $\boldsymbol{\Sigma}$ is known, then $\widehat{\boldsymbol{\xi}}$ is a strongly consistent estimator of $\boldsymbol{\xi}^0$.*

Proof of Theorem 2.1 See in Appendix A. ∎

To provide the asymptotic normality properties of the GLSEs we use the following notations. The $2P+1$ diagonal matrix $\boldsymbol{D}$ as

$$\boldsymbol{D} = \operatorname{diag}\left\{n^{1/2}, \ldots, n^{1/2}, n^{3/2}\right\}. \tag{11}$$

Let us denote the m-th row of $\boldsymbol{\Sigma}^{-1}$ as $\boldsymbol{\sigma}^m$, a $1 \times M$ vector, the $M \times 1$ vectors $\boldsymbol{A}_j = (A_{1j}, \ldots, A_{Mj})^\top$, $\boldsymbol{B}_j = (B_{1j}, \ldots, B_{Mj})^\top$, for $j = 1, \ldots, p_M$. Further, $A_{mj} = B_{mj} = 0$, for $p_m + 1 \leq j \leq p_M$.

Theorem 2.2. *Under Assumption 1,*

$$\boldsymbol{D}\left(\widehat{\boldsymbol{\xi}} - \boldsymbol{\xi}^0\right) \xrightarrow{d} N_{2P+1}\left(\mathbf{0}, 2\boldsymbol{\Pi}^{-1}\right).$$

Here, '$\xrightarrow{d}$' means convergence in distribution. The $(2P+1) \times (2P+1)$ matrix $\boldsymbol{\Pi}$ has the following form:

$$\boldsymbol{\Pi} = \begin{bmatrix} \boldsymbol{\Pi}_{11} & \boldsymbol{\Pi}_{12} & \ldots & \boldsymbol{\Pi}_{1M} & \boldsymbol{a}_1 \\ \boldsymbol{\Pi}_{12}^\top & \boldsymbol{\Pi}_{22} & \ldots & \boldsymbol{\Pi}_{2M} & \boldsymbol{a}_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{\Pi}_{1M}^\top & \boldsymbol{\Pi}_{2M}^\top & \ldots & \boldsymbol{\Pi}_{MM} & \boldsymbol{a}_M \\ \boldsymbol{a}_1^\top & \boldsymbol{a}_2^\top & \ldots & \boldsymbol{a}_M & \gamma \end{bmatrix}. \tag{12}$$

Here $\boldsymbol{\Pi}_{mm}$ is a $2p_m \times 2p_m$ diagonal matrix for $m = 1, \ldots, M$, γ is a scalar, the rest of matrices and vectors are compatible. The matrices are as follows:

$$\boldsymbol{\Pi}_{mm} = \sigma^{mm}\boldsymbol{I}_{2p_m}, \quad \boldsymbol{\Pi}_{mk} = \begin{bmatrix} \sigma^{mk}\boldsymbol{I}_{2p_m} & \mathbf{0} \end{bmatrix}; \quad 1 \leq m \leq k \leq M.$$

Let us use the following notations. The $M \times M$, diagonal matrices $\boldsymbol{I}^m$, for $m = 1, \ldots, M$ are as follows:

$$\boldsymbol{I}^1 = diag\{1, \ldots, 1\}, \quad \boldsymbol{I}^2 = diag\{0, 1, \ldots, 1\}, \quad \ldots \quad \boldsymbol{I}^M = diag\{0, \ldots, 0, 1\}.$$

The vectors $\boldsymbol{a}_1^\top = (a_{11}, b_{11}, \ldots, a_{1p_1}, b_{1p_1}), \ldots, \boldsymbol{a}_M^\top = (a_{M1}, b_{M1}, \ldots, a_{Mp_M}, b_{Mp_M})$, *where*

$$\begin{aligned}
a_{1j} &= \frac{j}{2}\boldsymbol{\sigma}^1\boldsymbol{I}^1\boldsymbol{B}_j, \quad b_{1j} = -\frac{j}{2}\boldsymbol{\sigma}^1\boldsymbol{I}^1\boldsymbol{A}_j; \qquad j = 1, \ldots, p_1 \\
a_{2j} &= \frac{j}{2}\boldsymbol{\sigma}^2\boldsymbol{I}^1\boldsymbol{B}_j, \quad b_{2j} = -\frac{j}{2}\boldsymbol{\sigma}^2\boldsymbol{I}^1\boldsymbol{A}_j; \qquad j = 1, \ldots, p_1 \\
a_{2j} &= \frac{j}{2}\boldsymbol{\sigma}^2\boldsymbol{I}^2\boldsymbol{B}_j, \quad b_{2j} = -\frac{j}{2}\boldsymbol{\sigma}^2\boldsymbol{I}^2\boldsymbol{A}_j; \qquad j = p_1 + 1, \ldots, p_2 \\
&\vdots \qquad \vdots \\
a_{Mj} &= \frac{j}{2}\boldsymbol{\sigma}^M\boldsymbol{I}^1\boldsymbol{B}_j, \quad b_{Mj} = -\frac{j}{2}\boldsymbol{\sigma}^M\boldsymbol{I}^1\boldsymbol{A}_j; \qquad j = 1, \ldots, p_1 \\
&\vdots \qquad \vdots \\
a_{Mj} &= \frac{j}{2}\boldsymbol{\sigma}^M\boldsymbol{I}^M\boldsymbol{B}_j, \quad b_{Mj} = -\frac{j}{2}\boldsymbol{\sigma}^M\boldsymbol{I}^M\boldsymbol{A}_j; \qquad j = p_1 + \ldots + p_{M-1} + 1, \ldots, p_M,
\end{aligned}$$

$$\gamma = \frac{1}{3}\left[\sum_{m=1}^{M}\sum_{j=1}^{p_m}\sigma^{mm} j^2 (A_{mj}^2 + B_{mj}^2) + 2\sum_{1 \le m < k \le M}\sum_{j=1}^{p_m}\sigma^{mk} j^2 (A_{mj}A_{kj} + B_{mj}B_{kj})\right].$$

Proof of Theorem 2.2 See in Appendix B. ∎

It is important to observe that the GLSEs of the linear parameters have the asymptotic variances which are of the order $O(1/n)$, where as the asymptotic variance of the GLSE of λ is of the order $O(1/n^3)$. This is not very surprising, and it is the case even for one channel fundamental frequency with harmonics model or multiple sinusoidal frequency model also, see for example Nandi and Kundu [12] or Kundu [9] in this respect. Most likely it is due to the presence of time t with λ always in the model. It indicates that a very efficient estimator of λ is possible.

Now we provide the consistency and asymptotic normality properties of the LSEs which can be obtained by minimizing (8). They can be obtained quite conveniently and they will be used to compute GLSE in practice.

Theorem 2.3. *If* $\widetilde{\boldsymbol{\xi}}$ *denotes the LSE of* $\boldsymbol{\xi}$, *then under Assumption 1,* $\widetilde{\boldsymbol{\xi}}$ *is a strongly consistent estimator of* $\boldsymbol{\xi}^0$.

Proof of Theorem 2.3 It can be obtained along the same line as the proof of Theorem 2.1.

Theorem 2.4. *Under the same set of assumptions as in Theorem 2.2*

$$\boldsymbol{D}\left(\widetilde{\boldsymbol{\xi}} - \boldsymbol{\xi}^0\right) \xrightarrow{d} N_{2P+1}\left(\boldsymbol{0}, 2\boldsymbol{\Gamma}^{-1}\boldsymbol{H}\boldsymbol{\Gamma}^{-1}\right).$$

The $(2P+1)\times(2P+1)$ *matrix* $\mathbf{\Gamma}$ *has the following form:*

$$\mathbf{\Gamma} = \begin{bmatrix} \mathbf{\Gamma}_{11} & \mathbf{\Gamma}_{12} & \ldots & \mathbf{\Gamma}_{1M} & \boldsymbol{b}_1 \\ \mathbf{\Gamma}_{12}^\top & \mathbf{\Gamma}_{22} & \ldots & \mathbf{\Gamma}_{2M} & \boldsymbol{b}_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{\Gamma}_{1M}^\top & \mathbf{\Gamma}_{2M}^\top & \ldots & \mathbf{\Gamma}_{MM} & \boldsymbol{b}_M \\ \boldsymbol{b}_1^\top & \boldsymbol{b}_2^\top & \ldots & \boldsymbol{b}_M & \delta \end{bmatrix}. \tag{13}$$

Here $\mathbf{\Gamma}_{mm}$ *is a* $2p_m \times 2p_m$ *diagonal matrix for* $m = 1,\ldots,M$, δ *is a scalar, the rest of the matrices and vectors are compatible. The matrices are as follows:*

$$\mathbf{\Gamma}_{mm} = \boldsymbol{I}_{2p_m}, \quad \mathbf{\Gamma}_{mk} = \begin{bmatrix} \boldsymbol{I}_{2p_m} & \mathbf{0} \end{bmatrix}; \quad 1 \le m \le k \le M.$$

We denote M *unit vectors* $\{\boldsymbol{u}_m; m = 1,\ldots,M\}$, *each of order* $M\times 1$, *where the vector* $\boldsymbol{u}_m$ *has 1 at the m-th row and zero elsewhere. The vectors* $\boldsymbol{b}_1^\top = (c_{11}, d_{11}, \ldots, c_{1p_1}, d_{1p_1}), \ldots,$ $\boldsymbol{b}_M^\top = (c_{M1}, d_{M1}, \ldots, c_{Mp_M}, d_{Mp_M})$, *where*

$$\begin{aligned} c_{1j} &= \frac{j}{2}\boldsymbol{u}_1^\top \boldsymbol{B}_j, \quad d_{1j} = -\frac{j}{2}\boldsymbol{u}_1^\top \boldsymbol{A}_j; \quad j = 1,\ldots,p_1 \\ c_{2j} &= \frac{j}{2}\boldsymbol{u}_2^\top \boldsymbol{B}_j, \quad d_{2j} = -\frac{j}{2}\boldsymbol{u}_2^\top \boldsymbol{A}_j; \quad j = 1,\ldots,p_2 \\ \vdots &\quad \vdots \\ c_{Mj} &= \frac{j}{2}\boldsymbol{u}_M^\top \boldsymbol{B}_j, \quad d_{Mj} = -\frac{j}{2}\boldsymbol{u}_M^\top \boldsymbol{A}_j; \quad j = 1,\ldots,p_M, \end{aligned}$$

and

$$\delta = \frac{1}{3}\sum_{m=1}^{M}\sum_{j=1}^{p_m} j^2 (A_{mj}^2 + B_{mj}^2).$$

The matrix $\boldsymbol{H}$ *has the same form as matrix* $\mathbf{\Pi}$, *where* σ^{mk} *will be replaced by* σ_{mk}, *for* $1 \le m, k \le M$ *in all the entries.*

Proof of Theorem 2.4 It can be obtained along the same line as the proof of Theorem 2.2. ∎

It is interesting to observe that both the LSEs and GLSEs have the same rates of convergence for the linear as well as non-linear parameters, although their asymptotic variances might be different. It is observed in our extensive simulation study, see Section 4, that the mean squared errors (MSEs) of the LSEs are more than the corresponding MSEs of the GLSEs. The difference of the performances of the LSEs and GLSEs depend on the structure of $\mathbf{\Sigma}$ as expected.

2.3. **Implementation.** As it has been mentioned before that often in practice the dispersion matrix $\boldsymbol{\Sigma}$ is not known. Hence, to compute the GLSEs of the unknown parameters, one needs a consistent estimator of $\boldsymbol{\Sigma}$. Note that once the LSE of $\boldsymbol{\xi}$ is obtained, one can obtain the estimator of σ_{mk} as follows:

$$
\begin{aligned}
\widetilde{\sigma}_{mm} &= \frac{1}{n}\sum_{t=1}^{n}\left[y_m(t) - \sum_{j=1}^{p_m}\left\{\widetilde{A}_{mj}\cos(j\widetilde{\lambda}t) + \widetilde{B}_{mj}\sin(j\widetilde{\lambda}t)\right\}\right]^2; \quad m = 1,\ldots,M, \\
\widetilde{\sigma}_{mk} &= \frac{1}{n}\sum_{t=1}^{n}\left[y_m(t) - \sum_{j=1}^{p_m}\left\{\widetilde{A}_{mj}\cos(j\widetilde{\lambda}t) + \widetilde{B}_{mj}\sin(j\widetilde{\lambda}t)\right\}\right] \\
&\qquad \times\left[y_k(t) - \sum_{j=1}^{p_k}\left\{\widetilde{A}_{kj}\cos(j\widetilde{\lambda}t) + \widetilde{B}_{kj}\sin(j\widetilde{\lambda}t)\right\}\right]; \quad 1 \le m \le k \le M.
\end{aligned}
$$

The following result provides the consistency properties of $\widetilde{\sigma}_{mk}$.

Theorem 2.5. *Under Assumption 1, $\widetilde{\sigma}_{mm}$ and $\widetilde{\sigma}_{mk}$ are strongly consistent estimator of σ_{mm} and σ_{mk}, respectively, for $1 \le m \le k \le M$.*

Proof of Theorem 2.5 To prove Theorem 2.5 let us use the term $o_{as}(1)$ means it converges to zero almost surely, and the term $o_{as}(1/n)$, means $n o_{as}(1/n)$ converges to zero almost surely. Now from Theorems 2.3 and 2.4 it follows that $\widetilde{A}_{mj} = A^0_{mj} + o_{as}(1)$, $\widetilde{B}_{mj} = B^0_{mj} + o_{as}(1)$, for $j = 1,\ldots,p_m$ and $m = 1,\ldots,M$. Also $\widetilde{\lambda} = \lambda^0 + o_{as}(1/n)$. Now using these and writing $y_m(t) = \mu_m(t) + e_m(t)$, $m = 1,\ldots,M$ in the expression of $\widetilde{\sigma}_{mk}$, the results can be obtained from the strong law of large numbers. The explicit details are avoided. ∎

Hence, in practice, the GLSEs can be obtained as a two-step process. In the first step we can obtain a consistent estimator of $\boldsymbol{\Sigma}$ based on the LSEs, and at the second step it can be used to compute the GLSEs of $\boldsymbol{\xi}$.

3. A Special Case

In this section we consider a special case when the number of channels is two, and $p_1 = p_2 = p$. In this case the parameter vector $\boldsymbol{\xi}$ is a $(2p+1)$ vector, and it can be written as $\boldsymbol{\xi} = (\boldsymbol{\alpha}_1^\top, \boldsymbol{\alpha}_2^\top, \lambda)^\top = (A_{11}, B_{11}, \ldots, A_{1p}, B_{1p}, A_{21}, B_{21}, \ldots, A_{2p}, B_{2p}, \lambda)^\top$. Then $\widetilde{\boldsymbol{\xi}}$, LSE of $\boldsymbol{\xi}$, can be obtained by minimizing

$$\begin{aligned} R_1(\boldsymbol{\xi}) &= \sum_{t=1}^{n}\Big[y_1(t) - \sum_{j=1}^{p}\Big\{A_{1j}\cos(j\lambda t) + B_{1j}\sin(j\lambda t)\Big\}\Big]^2 \\ &\qquad + \sum_{t=1}^{n}\Big[y_2(t) - \sum_{j=1}^{p}\Big\{A_{2j}\cos(j\lambda t) + B_{2j}\sin(j\lambda t)\Big\}\Big]^2. \qquad (14) \end{aligned}$$

For given λ, the LSE of $\boldsymbol{\alpha}_m$ can be obtained as

$$\widetilde{\boldsymbol{\alpha}}_m(\lambda) = (\boldsymbol{Z}^T(\lambda)\boldsymbol{Z}(\lambda))^{-1}\boldsymbol{Z}^T(\lambda)\boldsymbol{Y}_m, \quad m = 1, 2.$$

Here the vector $\boldsymbol{Y}_m$ and the matrix $\boldsymbol{Z}$ are same as defined in (8) and (10), respectively. The LSE of λ can be obtained by minimizing

$$Q_1(\lambda) = \boldsymbol{Y}_1^T\Big(\boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{Z}(\lambda)}\Big)\boldsymbol{Y}_1 + \boldsymbol{Y}_2^T\Big(\boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{Z}(\lambda)}\Big)\boldsymbol{Y}_2,$$

or equivalently by maximizing

$$Q_2(\lambda) = \boldsymbol{Y}_1^T\boldsymbol{P}_{\boldsymbol{Z}(\lambda)}\boldsymbol{Y}_1 + \boldsymbol{Y}_2^T\boldsymbol{P}_{\boldsymbol{Z}(\lambda)}\boldsymbol{Y}_2,$$

here $\boldsymbol{P}_{\boldsymbol{Z}(\lambda)} = \boldsymbol{Z}(\lambda)(\boldsymbol{Z}^\top(\lambda)\boldsymbol{Z}(\lambda))^{-1}\boldsymbol{Z}^\top(\lambda)$ is the projection matrix on the columns of $\boldsymbol{Z}(\lambda)$. Clearly, $\widetilde{\boldsymbol{\xi}}$ is a consistent estimator of $\boldsymbol{\xi}^0$. In this case the asymptotic distribution of $\widetilde{\boldsymbol{\xi}}$ can be written in the following form;

Theorem 3.1. *If $\boldsymbol{D}$ is $2p+1$ diagonal matrix with all the diagonal entries as $\sqrt{n}$, except the last one which is $n^{3/2}$, then under Assumption 1,*

$$\boldsymbol{D}(\Big(\widetilde{\boldsymbol{\xi}} - \boldsymbol{\xi}^0\Big) \xrightarrow{d} N\big(\boldsymbol{0}, 2\boldsymbol{\Gamma}^{-1}\boldsymbol{H}\boldsymbol{\Gamma}^{-1}\big).$$

Here
$\boldsymbol{\Gamma} = \begin{pmatrix} \boldsymbol{J}\otimes\boldsymbol{I}_{2p} & \boldsymbol{r} \\ \boldsymbol{r}^\top & \beta \end{pmatrix}$, $\boldsymbol{J} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, $\boldsymbol{r} = (\boldsymbol{r}_1^\top \;\; \boldsymbol{r}_2^\top)^\top$,

$$\begin{aligned} \boldsymbol{r}_1^\top &= \frac{1}{2}\Big(-B_{11}^0, A_{11}^0, -2B_{12}^0, 2A_{12}^0, \ldots, -pB_{1p}^0, pA_{1p}^0\Big), \\ \boldsymbol{r}_2^\top &= \frac{1}{2}\Big(-B_{21}^0, A_{21}^0, -2B_{22}^0, 2A_{22}^0, \ldots, -pB_{2p}^0, pA_{2p}^0\Big), \\ \beta &= \frac{1}{3}\sum_{j=1}^{p} j^2\Big(A_{1j}^{0^2} + B_{1j}^{0^2} + A_{2j}^{0^2} + B_{2j}^{0^2}\Big), \end{aligned}$$

and
$\boldsymbol{H} = \begin{pmatrix} \boldsymbol{\Sigma}\otimes\boldsymbol{I}_{2p} & \boldsymbol{s} \\ \boldsymbol{s}^T & \gamma \end{pmatrix}$, $\boldsymbol{s} = (\boldsymbol{s}_1^T \;\; \boldsymbol{s}_2^T)^T$,

$$\boldsymbol{s}_1^\top = \Big(-\sigma_1^2 B_{11}^0 - \sigma_{12}B_{21}^0, \sigma_1^2 A_{11}^0 + \sigma_{12}A_{21}^0, \ldots, -\sigma_1^2 pB_{1p}^0 - \sigma_{12}pB_{2p}^0, \sigma_1^2 pA_{1p}^0 + \sigma_{12}pA_{2p}^0\Big)_{1\times 2p}$$

$$\boldsymbol{s}_2^\top = \Big(-\sigma_{12}B_{11}^0 - \sigma_2^2 B_{21}^0, \sigma_{12}A_{11}^0 + \sigma_2^2 A_{21}^0, \ldots, -\sigma_{12}pB_{1p}^0 - \sigma_2^2 pB_{2p}^0, \sigma_{12}pA_{1p}^0 + \sigma_2^2 pA_{2p}^0\Big)_{1\times 2p}$$

$$\gamma = \frac{2}{3}\sum_{j=1}^{p} j^2\Big[\sigma_1^2\Big(A_{1j}^{0^2} + B_{1j}^{0^2}\Big) + \sigma_2^2\Big(A_{2j}^{0^2} + B_{2j}^{0^2}\Big) + 2\sigma_{12}\Big(A_{1j}^0 A_{2j}^0 + B_{1j}^0 B_{2j}^0\Big)\Big].$$

Here '$\otimes$' denotes the Kronecker product. ∎

Now let us discuss about the GLSE of $\boldsymbol{\xi}$. The GLSE of $\boldsymbol{\xi}$ can be obtained by minimizing

$$\begin{aligned} R(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \lambda) \;\; = \;\; & \sum_{t=1}^{n}\Big[\sigma_{22}\Big(y_1(t) - \mu_1(t)\Big)^2 + \sigma_{11}\Big(y_2(t) - \mu_2(t)\Big)^2 \\ & -2\sigma_{12}\Big(y_1(t) - \mu_1(t)\Big)\Big(y_2(t) - \mu_2(t)\Big)\Big]. \end{aligned} \qquad (15)$$

If we denote $\boldsymbol{\eta}_j = (A_{1j}, B_{1j}, A_{2j}, B_{2j})^\top$, $j = 1, \ldots, p$ and $\boldsymbol{\eta} = (\boldsymbol{\eta}_1^\top, \ldots, \boldsymbol{\eta}_p^\top)^\top$, and we use the following notations; $M_n(\lambda)$ is a $2p \times 2p$ matrix of the following form.

$$M_n(\lambda) = \begin{pmatrix} M(1,1,\lambda) & M(1,2,\lambda) & \cdots & M(1,p,\lambda) \\ M(2,1,\lambda) & M(2,2,\lambda) & \cdots & M(2,p,\lambda) \\ \vdots & \vdots & \vdots & \vdots \\ M(p,1,\lambda) & M(p,2,\lambda) & \cdots & M(p,p,\lambda) \end{pmatrix},$$

where

$$M(j,k,\lambda) = \frac{2}{n}\begin{pmatrix} \sum_{t=1}^n \cos(j\lambda t)\cos(k\lambda t) & \sum_{t=1}^n \cos(j\lambda t)\sin(k\lambda t) \\ \sum_{t=1}^n \sin(j\lambda t)\cos(k\lambda t) & \sum_{t=1}^n \sin(j\lambda t)\sin(k\lambda t) \end{pmatrix},$$

and the vector $W_n(\lambda)$ of order $4p$ is defined as

$$W_n(\lambda) = (W_{1n}(\lambda)^T, \ldots W_{pn}(\lambda)^T)^T, \quad W_{jn}(\lambda) = \frac{2}{n}\begin{pmatrix} \sum_{t=1}^n y_1(t)\cos(j\lambda t) \\ \sum_{t=1}^n y_1(t)\sin(j\lambda t) \\ \sum_{t=1}^n y_2(t)\cos(j\lambda t) \\ \sum_{t=1}^n y_2(t)\sin(j\lambda t) \end{pmatrix}.$$

For a given λ, the GLSEs of $\boldsymbol{\eta}$ can be written as

$$\begin{aligned} \widehat{\boldsymbol{\eta}}(\lambda) &= (\boldsymbol{\Sigma}^{-1} \otimes M_n(\lambda))^{-1}(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{I}_{2p})W_n(\lambda) \\ &= (\boldsymbol{\Sigma} \otimes (M_n(\lambda))^{-1})(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{I}_{2p})W_n(\lambda) \\ &= (\boldsymbol{I}_2 \otimes (M_n(\lambda))^{-1})W_n(\lambda). \end{aligned}$$

Note that for large n, $M(j,k,\lambda) = o(1)$ when $j \neq k = 1, \ldots, p$. Therefore for large n, $M_n(\lambda)$ is a box diagonal matrix with the j-th diagonal sub-matrix as $M(j,j,\lambda)$. Using this special structure of $M_n(\lambda)$ for large n, we deduce that

$$\begin{aligned}\widehat{\boldsymbol{\eta}}_j(\lambda) &= (\boldsymbol{\Sigma}^{-1}\otimes M(j,j,\lambda))^{-1}(\boldsymbol{\Sigma}^{-1}\otimes \boldsymbol{I}_2)W_{jn}(\lambda)\\ &= (\boldsymbol{I}_2\otimes (M(j,j,\lambda))^{-1})W_{jn}(\lambda), \quad j=1,\ldots,p.\end{aligned}$$

This boils down to the method of estimation of the linear parameters sequentially, once the fundamental frequency λ is estimated..

Additionally, we also have $M(j,j,\lambda)=\boldsymbol{I}_2+o(1/n)$ for large n. Therefore, using this in $\widehat{\boldsymbol{\eta}}_j(\lambda)$ for large n, we have

$$\begin{aligned}\widehat{\boldsymbol{\eta}}_j(\lambda) &= (\boldsymbol{I}_2\otimes \boldsymbol{I}_2)^{-1})W_{jn}(\lambda)+o(1/n)\\ &= W_{jn}(\lambda)+o(1/n), \quad j=1,\ldots,p.\end{aligned}$$

This is nothing but the approximate LSEs of the linear parameters corresponding to the j-th component. Finally, the GLSE of λ can be obtained by minimizing $R(\widehat{\boldsymbol{\alpha}}_1(\lambda),\widehat{\boldsymbol{\alpha}}_2(\lambda),\lambda)$ with respect to λ. Under the same assumptions as in Theorem 3.1, the GLSEs are consistent estimators. The asymptotic distribution of $\widehat{\boldsymbol{\xi}}$ can be written as follows.

Theorem 3.2. *Under the same assumptions as Theorem 3.1,*

$$\boldsymbol{D}\Big(\widehat{\boldsymbol{\xi}}-\boldsymbol{\xi}^0\Big)\xrightarrow{d} N\big(\boldsymbol{0},2\boldsymbol{H}_G^{-1}\big).$$

Here $\boldsymbol{H}_G$ *is a* $(4p+1)\times(4p+1)$ *matrix, and*

$$\boldsymbol{H}_G=\begin{pmatrix}\boldsymbol{\Sigma}^{-1}\otimes \boldsymbol{I}_{2p} & \boldsymbol{v}\\ \boldsymbol{v}^T & \gamma_G\end{pmatrix}, \quad \boldsymbol{v}=(\boldsymbol{v}_1^T \ \ \boldsymbol{v}_2^T)^T,$$

$$\boldsymbol{v}_1^T=\Big(v_{11},u_{11},v_{12},u_{12},\ldots,v_{1p},u_{1p}\Big)^T, \quad \boldsymbol{v}_2^T=\Big(v_{21},u_{21},v_{22},u_{22},\ldots,v_{2p},u_{2p}\Big)^T,$$

$$v_{1j}=\frac{j}{2|\boldsymbol{\Sigma}|}\big(\sigma_2^2B_{1j}^0-\sigma_{12}B_{2j}^0\big), \quad u_{1j}=-\frac{j}{2|\boldsymbol{\Sigma}|}\big(\sigma_2^2A_{1j}^0-\sigma_{12}A_{2j}^0\big),$$

$$v_{2j}=\frac{j}{2|\boldsymbol{\Sigma}|}\Big(\sigma_1^2B_{2j}^0-\sigma_{12}B_{1j}^0\Big), \quad u_{2j}=-\frac{j}{2|\boldsymbol{\Sigma}|}\big(\sigma_1^2A_{2j}^0-\sigma_{12}A_{1j}^0\big), \ j=1,\ldots,p$$

$$\gamma_G=\frac{1}{3|\boldsymbol{\Sigma}|}\sum_{j=1}^{p}j^2\Big[\sigma_2^2\Big(A_{1j}^{0^2}+B_{1j}^{0^2}\Big)+\sigma_1^2\Big(A_{2j}^{0^2}+B_{2j}^{0^2}\Big)-2\sigma_{12}\Big(A_{1j}^0A_{2j}^0+B_{1j}^0B_{2j}^0\Big)\Big].$$

It is worth mentioning that Christensen [5] also considered a two-channel model as discussed in this section and provided the maximum likelihood estimators of the unknown parameters based on the assumption that the errors are normally distributed and the matrix $\boldsymbol{\Sigma}$ is a diagonal matrix. Hence, the methodology proposed by Christensen [5], can be obtained as a special case of this manuscript. Further, the author did not consider more

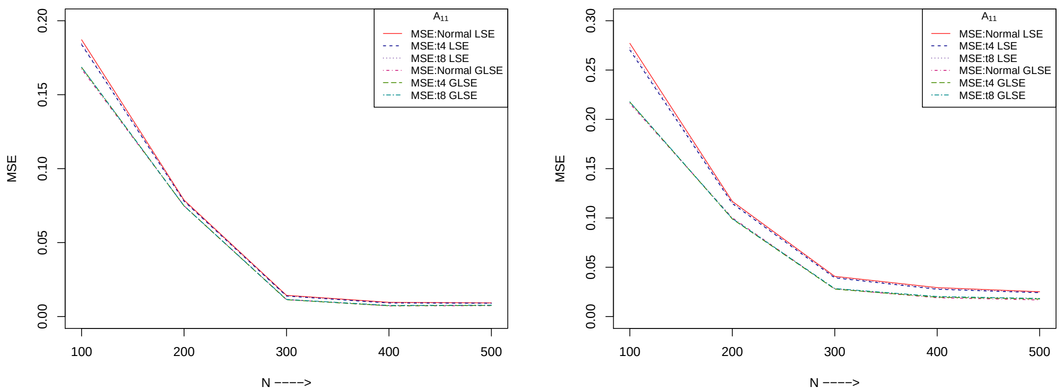

Figure 1. The MSEs of different estimates of A_{11}.

than two channels model and did not provide any asymptotic properties of the proposed estimators. These are some of the major contributions of this manuscript.

4. Numerical Experiments

In this section we present some simulation results for two channel model with $p_1 = p_2 = 2$. In Section 5 we present the analysis of a data set where $p_1 \neq p_2$. The main aim is to compare the behavior of the LSEs and GLSEs for different model parameters, for different error distributions and for different sample sizes. We consider the following model parameter:

$$\begin{aligned} &A_{11} = 4.0; \quad B_{11} = 6.5; \quad A_{12} = 5.0; \quad B_{12} = 3.0; \\ &A_{21} = 5.0; \quad B_{21} = 3.0; \quad A_{22} = 3.0; \quad B_{22} = 2.0, \quad \lambda = 0.2. \end{aligned} \tag{16}$$

The sequence of the random vectors $\{\boldsymbol{e}(t)\}$ has been considered in three different ways:

(1) a sequence of bivariate normal vectors with mean $\mathbf{0}$ and variance matrix $\boldsymbol{\Sigma}$,
(2) a sequence of bivariate t_4 distributed random vectors with the same mean and variance structure as (1),
(3) a sequence of bivariate t_8 distributed random vectors with the same mean and variance structure as (1).

Here t_ν denotes as t distribution with ν degrees of freedom.

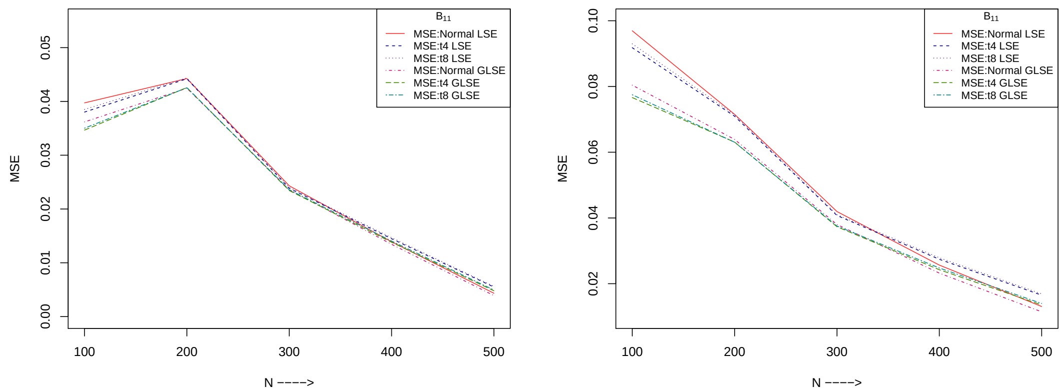

FIGURE 2. The MSEs of different estimates of B_{11}.

The variance covariance matrix considered in simulation studies are

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}; \quad \boldsymbol{\Sigma}_1 : \sigma_1^2 = 1, \sigma_2^2 = 1, \rho = 0.9; \quad \boldsymbol{\Sigma}_2 : \sigma_1^2 = 3, \sigma_2^2 = 3, \rho = 0.95.$$

The sample sizes considered in these experiments are 100, 200, 300, 400 and 500. In all the three cases of error generations, the data are generated using parameter values (16) and different $\boldsymbol{\Sigma}_k$'s and sample sizes. For every value of N, the sample size and $\boldsymbol{\Sigma}$, the error matrix, 5000 realizations are generated and estimates are obtained using both the LS and GLS methods. The average estimates and mean squared errors (MSEs) of all the parameter estimates using both the proposed methods are computed. We report the MSEs of different parameters estimated in Figures 1-9. In Figure 1, the MSEs of the LSE and GLSE of A_{11} from all the three cases considered here are plotted against the sample size. The plot at the left is for dispersion matrix $\boldsymbol{\Sigma}_1$ and the right one is for $\boldsymbol{\Sigma}_2$. Similarly, the MSEs of the other parameter estimates are plotted in Figures 2-9.

The following observations can be made from the experiment described above.

(1) The MSEs of the LSEs as well as the GLSEs in all the cases considered here decrease as the sample size increases.
(2) The spread of the MSEs in right side plots are wider than that of the left side plots. This is due to the fact that the elements of $\boldsymbol{\Sigma}_2$ are larger than the elements of $\boldsymbol{\Sigma}_1$.
(3) There are clear separation of lines for LSEs and GLSEs. The MSEs in all three cases (1), (2) and (3) considered here for an estimator using LS method are clubbing at

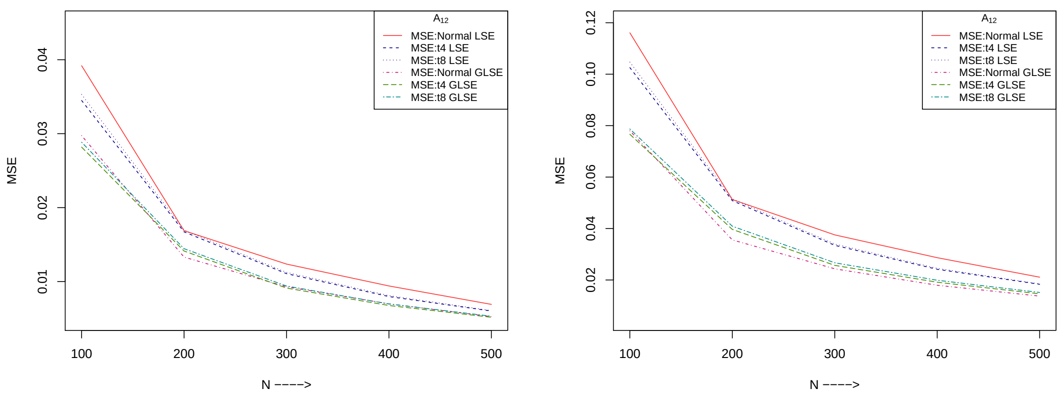

FIGURE 3. The MSEs of different estimates of A_{12}.

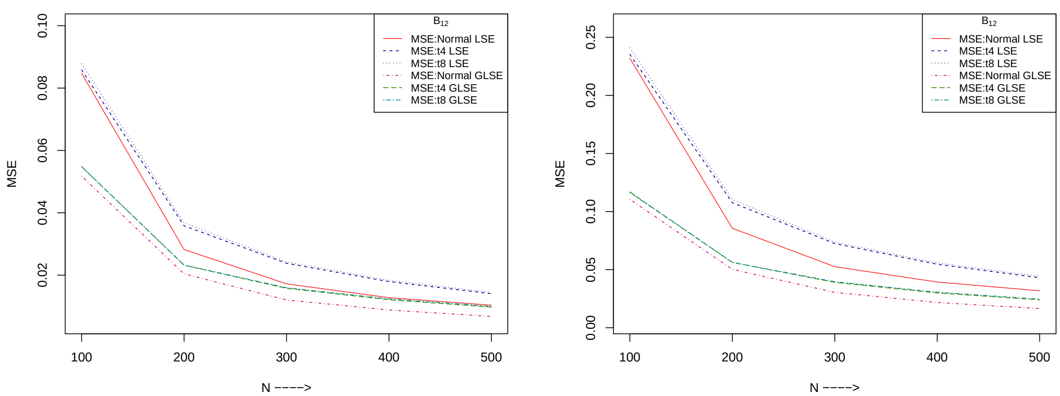

FIGURE 4. The MSEs of different estimates of B_{12}.

slightly larger values than the case when GLS method is used. The lines for GLS method are also clubbed. This has been observed in case of each parameter estimator.

(4) In all the linear parameter estimators using LS method, the MSEs is larger when error is distributed as bivariate t than the case when the error is bivariate normal. The MSEs for t_8 distribution is marginally larger than t_4 distribution. This has not been observed in case of the fundamental frequency.

With the same model parameters as given in (16) and $\sigma_1^2 = \sigma_2^2 = 1.0$, we would like to see the behavior of the MSEs with varying ρ, the correlation coefficient of the bivariate error process. The sample size N is fixed at 500 and ρ is varied from 0.1 to 0.9 and 0.98. The MSEs of the LSEs and GLSEs of the fundamental frequency λ have been calculated when

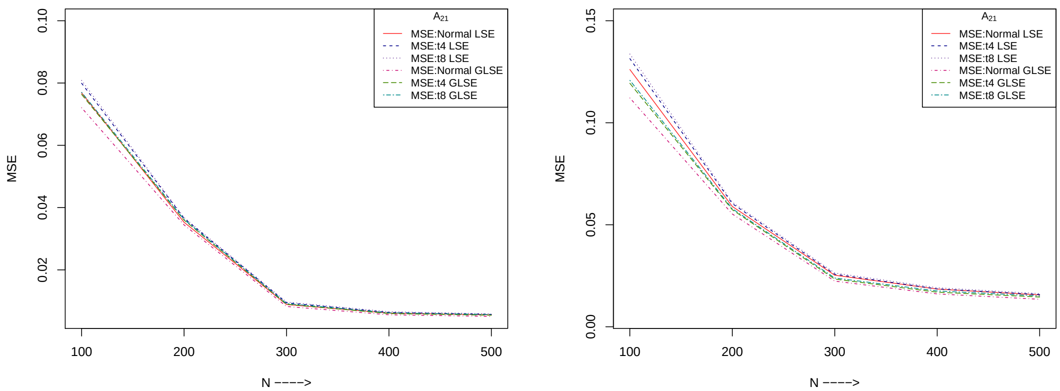

FIGURE 5. The MSEs of different estimates of A_{21}.

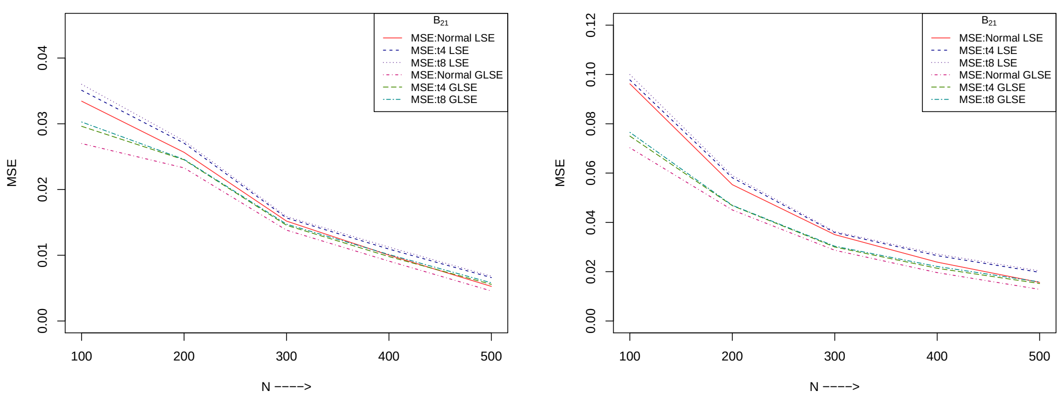

FIGURE 6. The MSEs of different estimates of B_{21}.

the error random vectors are bivariate normal, bivariate t_4 and bivariate t_8 as considered before. These six cases have been plotted in Figure 10. We observe that the MSEs of the LSEs increases as ρ increases under all the error distributions considered here whereas the MSEs of the GLSEs first increases and then decreases as ρ increases. The MSEs of the LSEs is always larger than that of the GLSEs in all the cases. In case of both the LSEs and GLSEs, the MSEs decrease with the error distribution from t_8, t_4 and then normal.

5. Data Analysis

In this Section, we analyze two short duration speech data using a two-channel fundamental frequency model, proposed in this article. The 'aaa' and 'uuu' voiced speech data, both

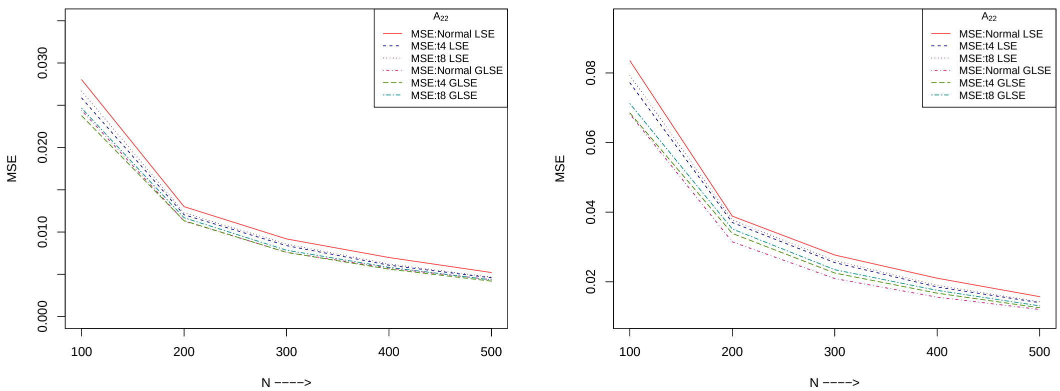

FIGURE 7. The MSEs of different estimates of A_{22}.

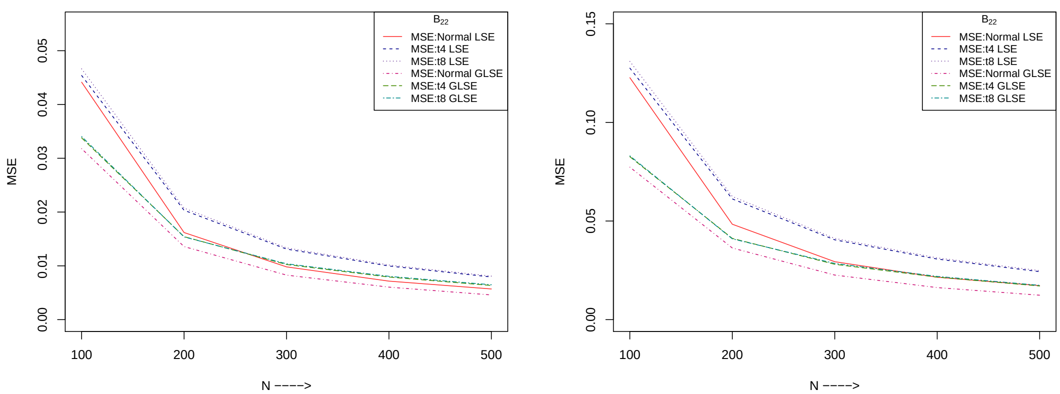

FIGURE 8. The MSEs of different estimates of B_{22}.

have 512 signal values, sampled at 10 kHz frequency were collected at the Speech Signal Processing laboratory of the Indian Institute of Technology, Kanpur. The mean corrected and scaled data are plotted in Figure 11. To have an idea about inherent frequencies, we plot the periodogram functions of both the datasets. The periodogram function of a set of n observations $\{z(t), t = 1, \ldots, n\}$ is defined for $\omega \in (0, \pi)$ as

$$I(\omega) = \frac{1}{n} \left| \sum_{t=1}^{n} z(t) e^{-i\omega t} \right|^2 .$$

The periodogram functions of 'aaa' and 'uuu' are plotted in Figure 12. The preliminary analysis of the periodograms reveals that the first significant frequency (peak) in 'aaa' data is close to 0.113 and that of 'uuu' data is close to 0.114. So they are approximately equal.

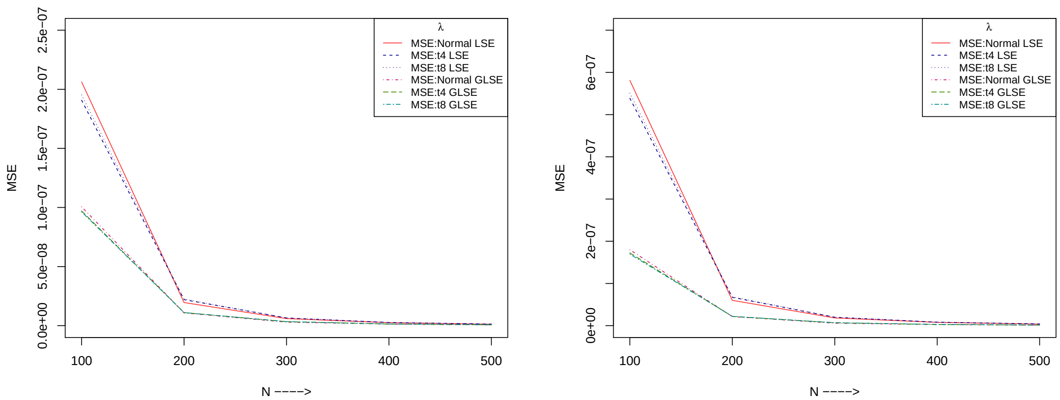

FIGURE 9. The MSEs of different estimates of λ.

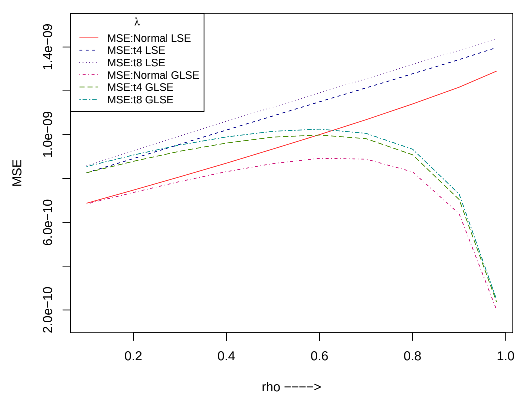

FIGURE 10. The MSEs of LSEs and GLSEs of λ under different error distribution with varying ρ.

We also observe that in the periodogram function of 'uuu' data, rest of the peaks, that is the frequencies, are at equal intervals. This means that they are the harmonics of the first frequency, λ as per our notation. In 'aaa dataset, although all the significant peaks are not at equal intervals, any significant peak is at $k\lambda$ where k is an integer and λ is the first frequency, the fundamental frequency. Therefore, these two data sets can be analyzed using two-channel fundamental frequency model with different number of harmonics. In this analysis, it is expected that amplitude estimates corresponding to some of the harmonics in 'aaa' data will be close to zero.

In this case p_1 and p_2 are unknown. A significant amount of work has been done in estimating the number of harmonics for one channel fundamental frequency model, see for

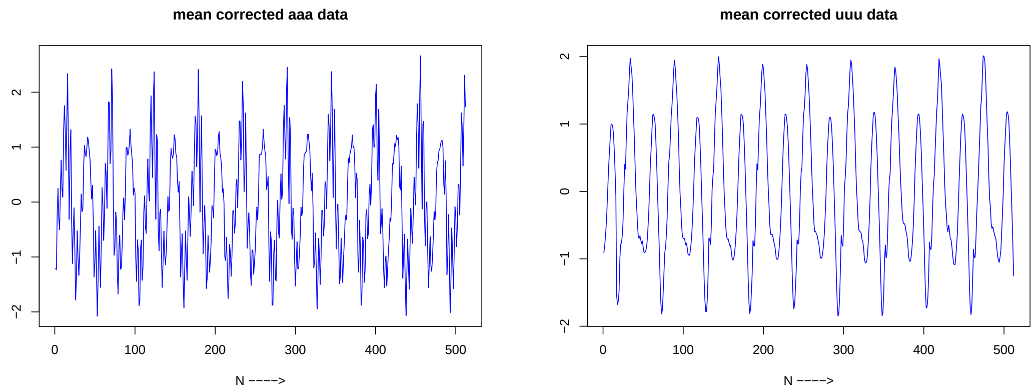

FIGURE 11. The mean corrected and scaled 'aaa' and 'uuu' data.

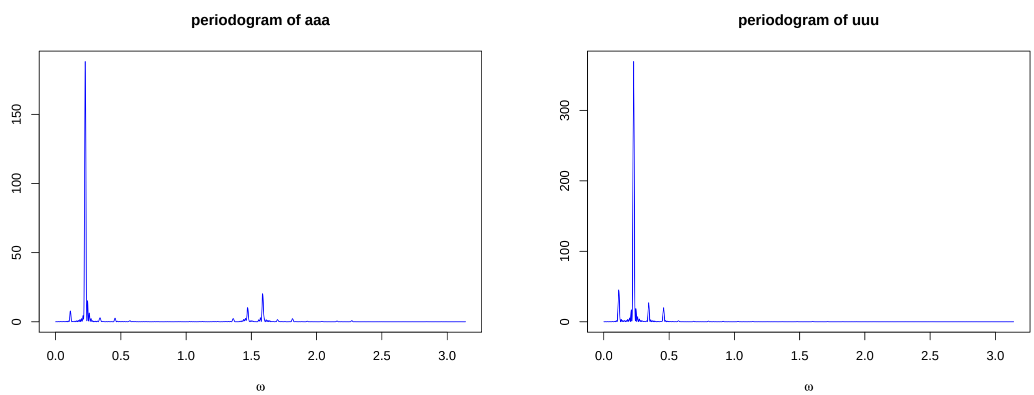

FIGURE 12. The periodogram functions of 'aaa' and 'uuu' data.

example Chapter 5 of Nandi and Kundu [13]. By exploratory analysis, we have obtained $p_1 = 17$ and $p_2 = 6$, the numbers of harmonics in 'aaa' and 'uuu' data, respectively. We first estimated the LSEs of the unknown fundamental frequency and linear parameters and the elements of the error variance matrix $\boldsymbol{\Sigma}$. The error variances and covariance are estimated as $\widetilde{\sigma}_1^2 = 9.93 \times 10^{-2}$, $\widetilde{\sigma}_2^2 = 7.26 \times 10^{-2}$ and $\widetilde{\sigma}_{12} = 3.94 \times 10^{-3}$. The fitted signal using LSEs along with mean corrected data in both the channels are plotted in Figure 13. They match quite well. In order to find the GLSEs of the unknown parameters, the estimates of σ_1^2, σ_2^2 and σ_{12}, obtained above are used. The unknown parameters are estimated by minimizing $S(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \lambda)$ following the methodology described in Section 3. The error variances and the covariance estimated using GLSEs of λ and the linear parameters are as follows:

$$\widehat{\sigma}_1^2 = 9.56 \times 10^{-2}, \quad \widehat{\sigma}_2^2 = 7.68 \times 10^{-2}, \quad \widehat{\sigma}_{12} = 4.26 \times 10^{-3}.$$

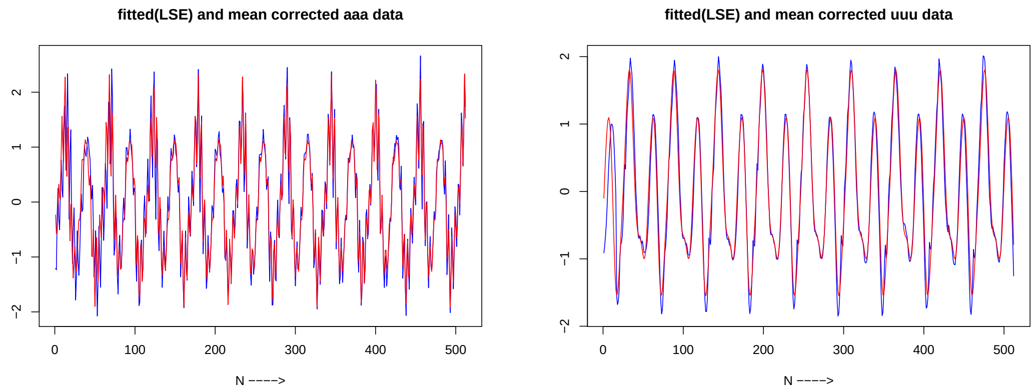

Figure 13. The fitted (red) and the observed mean corrected (blue) 'aaa' and 'uuu' data using LSEs.

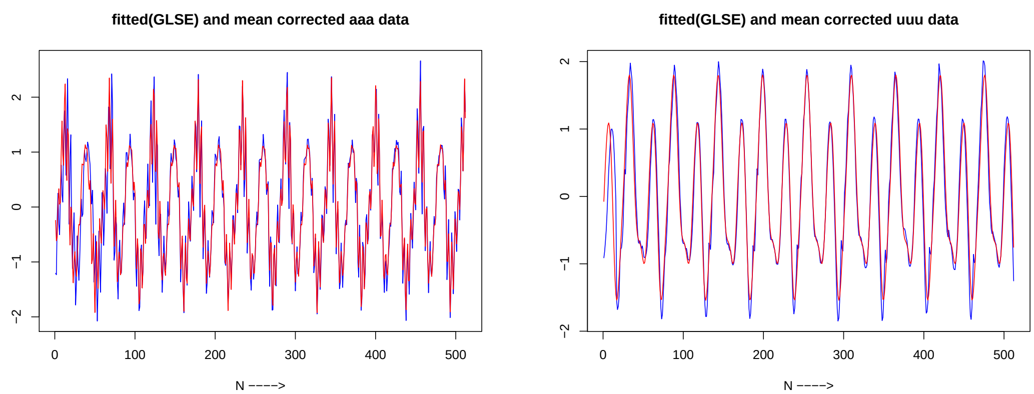

Figure 14. The fitted (red) and the observed mean corrected (blue) 'aaa' and 'uuu' data using GLSEs.

The fitted values using GLSEs and the mean corrected observed data are plotted in Figure 14. By observing Figures 13 and 14, it can be said that using model (1) is a reasonable way of analyzing 'aaa' and 'uuu' data sets simultaneously and both the LSEs and the GLSEs work well in this case.

6. Concluding Remarks

In this article, we propose a multichannel fundamental frequency with harmonics model in its most general form. This model is useful when signals from different channels have the same fundamental frequency, and the effective frequencies are harmonics of this common fundamental frequency. It has been assumed that the noise among the channels may be correlated also, which has not been considered before. We have proposed the GLSEs of the unknown parameters which become the maximum likelihood estimators also when the noise becomes multivariate normal distribution. The theoretical properties of the GLSEs have been established. It is observed that to implement the GLSEs in practice one needs to know the noise dispersion matrix, which may not be available. We have provided an implementation procedure of the GLSEs in practice using LSEs which do not need any information about the variance covariance matrix. Numerical experiments have been conducted based on a two-channel model with the number of components as two. Two short duration voice data have been analyzed using a two-channel model with different number of harmonics. The data analyses reveal that the proposed model works quite well in practice.

In this paper we have assumed that the noise components are independent and identically distributed in each channel, although they may not be independent between the channels. It will be interesting to develop proper methodologies when the error variances in each channel may also not be independent. Although, the dependency structure on the error variance in each channel can be easily incorporated, it is not immediate how it can be generalized to the multichannel case. Moreover, defining the LSEs of the unknown parameters can be done along the same way, but defining the GLSEs needs some careful attention. More work is needed along that direction.

Acknowledgments:

The authors would like to thank the Associate editor and two reviewers for their constructive comments, which have helped to improve the manuscript significantly.

Appendix A

To prove Theorem 2.1, we need the following result and lemmas. In this Appendix, for any vector $\boldsymbol{a}$, $|\boldsymbol{a}|$ means the Euclidean norm of $\boldsymbol{a}$.

Result A.1: If $\omega \in (0, \pi)$, then the following results hold.

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

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

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

$$\lim_{n\to\infty} \frac{1}{n^{\frac{2k+1}{2}}}\sum_{t=1}^{n} t^k \cos(\omega t) = \lim_{n\to\infty} \frac{1}{n^{\frac{2k+1}{2}}}\sum_{t=1}^{n} t^k \sin(\omega t) = 0.$$

Proof of Result A.1: The proofs can be found in Mangulis [11]. ∎

Lemma 1. *Let* $\widehat{\boldsymbol{\xi}}$ *be the GLSE of* $\boldsymbol{\xi}^0$ *that minimizes* $R(\boldsymbol{\xi})$, *and* $\boldsymbol{\xi} \in \Theta = [-M, M]^{2P} \times [0, \pi/p_M]$. *For any* $\epsilon > 0$, *suppose* $S_\epsilon = \{\boldsymbol{\xi} : \boldsymbol{\xi} \in \boldsymbol{\Theta}, |\boldsymbol{\xi} - \boldsymbol{\xi}^0| > (2P+1)\epsilon\}$, *for some fixed* $\boldsymbol{\xi}^0$, *an interior point of* $\boldsymbol{\Theta}$. *If for any* $\epsilon > 0$,

$$\lim_{n\to\infty} \inf_{S_\epsilon} \frac{1}{n}[R(\boldsymbol{\xi}) - R(\boldsymbol{\xi}^0)] > 0, \qquad a.e.,$$

then as $n \to \infty$, $\widehat{\boldsymbol{\xi}}$ *is a strongly consistent estimator of* $\boldsymbol{\xi}^0$.

Proof of Lemma 1: The proof mainly follows using contradiction argument, as that of Lemma 1 of Wu [20].

Lemma 2. *Let* $\{e(t)\}$ *be a sequence of i.i.d. random variables with mean zero and finite variance* $\sigma^2 > 0$, *then as* $n \to \infty$,

$$\sup_\omega \left| \frac{1}{n^{k+1}} \sum_{t=1}^{n} t^k e(t) \cos(\omega t) \right| \xrightarrow{a.e.} 0, \quad and \quad \sup_\omega \left| \frac{1}{n^{k+1}} \sum_{t=1}^{n} t^k e(t) \sin(\omega t) \right| \xrightarrow{a.e.} 0.$$

Proof of Lemma 2: See for example Kundu [9].

Proof of Theorem 2.1: Observe that

$$\frac{1}{n}[R(\boldsymbol{\xi}) - R(\boldsymbol{\xi}^0)] = f_1(\boldsymbol{\xi}) + f_2(\boldsymbol{\xi}),$$

where

$$\begin{aligned}
f_1(\boldsymbol{\xi}) &= \frac{1}{n} \sum_{t=1}^{n} \left(\boldsymbol{\mu}^0(t) - \boldsymbol{\mu}(t)\right)^\top \boldsymbol{\Sigma}^{-1} \left(\boldsymbol{\mu}^0(t) - \boldsymbol{\mu}(t)\right) \\
f_2(\boldsymbol{\xi}) &= \frac{2}{n} \sum_{t=1}^{n} \boldsymbol{e}(t)^\top \boldsymbol{\Sigma}^{-1} \left(\boldsymbol{\mu}^0(t) - \boldsymbol{\mu}(t)\right) = \frac{1}{n} \sum_{t=1}^{n} \boldsymbol{e}(t)^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}^0(t) - \frac{1}{n} \sum_{t=1}^{n} \boldsymbol{e}(t)^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}(t).
\end{aligned}$$

Hence, because of Lemma 2,

$$\lim_{n\to\infty} \sup_{\boldsymbol{\xi} \in S_\epsilon} |f_2(\boldsymbol{\xi})| = \lim_{n\to\infty} \sup_{\boldsymbol{\xi} \in S_\epsilon} \left| \frac{1}{n} \sum_{t=1}^{n} \boldsymbol{e}(t)^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}(t) \right| \leq \lim_{n\to\infty} \sup_{\boldsymbol{\xi} \in \boldsymbol{\Theta}} \left| \frac{1}{n} \sum_{t=1}^{n} \boldsymbol{e}(t)^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}(t) \right| \longrightarrow 0, \ a.e.$$

Now consider the following sets for $m = 1, \ldots, M, \quad j = 1, \ldots, p_m$,

$$S_{\epsilon, A_{mj}} = \{\boldsymbol{\xi} : |A_{mj} - A^0_{mj}| > \epsilon; \boldsymbol{\xi} \in \boldsymbol{\Theta}\}, \quad S_{\epsilon, B_{mj}} = \{\boldsymbol{\xi} : |B_{mj} - B^0_{mj}| > \epsilon; \boldsymbol{\xi} \in \boldsymbol{\Theta}\}$$

and

$$S_{\epsilon, \lambda} = \{\boldsymbol{\xi} : |\lambda - \lambda^0| > \epsilon; \boldsymbol{\xi} \in \boldsymbol{\Theta}\}.$$

Clearly,

$$S_\epsilon \subset \left(\bigcup_{m=1}^{M}\bigcup_{j=1}^{p_m} S_{\epsilon,A_{mj}}\right)\bigcup\left(\bigcup_{m=1}^{M}\bigcup_{j=1}^{p_m} S_{\epsilon,B_{mj}}\right)\bigcup S_{\epsilon,\lambda}.$$

Now let us consider

$$\liminf_{n\to\infty}\inf_{\boldsymbol{\xi}\in S_{\epsilon,A_{11}}}\frac{1}{n}f_1(\boldsymbol{\xi}) = \liminf_{n\to\infty}\inf_{\boldsymbol{\xi}\in S_{\epsilon,A_{11}}}\frac{\sigma^{11}}{n}\sum_{t=1}^{n}(A_{11}^0 - A_{11})^2\cos^2(\lambda^0 t) = \frac{\sigma^{11}}{2}(A_{11}^0 - A_{11})^2 > 0.$$

Similarly, it can be shown that for other sets $S_{\epsilon,A_{mj}}$, $S_{\epsilon,B_{mj}}$, $S_{\epsilon,\lambda}$, the corresponding limits are strictly positive. Hence, using Lemma 1, the result follows. ∎

Appendix B

Proof of Theorem 2.2: In order to obtain the asymptotic distribution of $\widehat{\boldsymbol{\xi}}$, we denote $R'(\boldsymbol{\xi})$ and $R''(\boldsymbol{\xi})$ as the vector of first derivatives and the matrix of second derivatives of $R(\boldsymbol{\xi})$, of the order $2P+1$ and $(2P+1)\times(2P+1)$, respectively. Now we express $R'(\boldsymbol{\xi})$ around the point $\boldsymbol{\xi}^0$, the true parameter vector using multivariate Taylor series expansion as follows

$$R'(\widehat{\boldsymbol{\xi}}) - R'(\boldsymbol{\xi}^0) = R''(\overline{\boldsymbol{\xi}})(\widehat{\boldsymbol{\xi}} - \boldsymbol{\xi}^0), \tag{17}$$

where $\overline{\boldsymbol{\xi}}$ is a point on the line joining $\widehat{\boldsymbol{\xi}}$ and $\boldsymbol{\xi}^0$. Since $\widehat{\boldsymbol{\xi}}$ minimizes $R(\boldsymbol{\xi})$, $R'(\widehat{\boldsymbol{\xi}}) = 0$. Using the $(2P+1)\times(2P+1)$ diagonal matrix $\boldsymbol{D}$, as defined in (11), (17) can be written as

$$\boldsymbol{D}(\widehat{\boldsymbol{\xi}} - \boldsymbol{\xi}^0) = [\boldsymbol{D}^{-1}R''(\overline{\boldsymbol{\xi}})\boldsymbol{D}^{-1}]^{-1}[\boldsymbol{D}^{-1}R'(\boldsymbol{\xi}^0)],$$

provided, $[\boldsymbol{D}^{-1}R''(\overline{\boldsymbol{\xi}})\boldsymbol{D}^{-1}]$ is an invertible matrix. Since, $\widehat{\boldsymbol{\xi}} \longrightarrow \boldsymbol{\xi}^0$, a.s.,

$$\lim_{n\to\infty}[\boldsymbol{D}^{-1}R''(\overline{\boldsymbol{\xi}})\boldsymbol{D}^{-1}] = \lim_{n\to\infty}[\boldsymbol{D}^{-1}R''(\boldsymbol{\xi}^0)\boldsymbol{D}^{-1}].$$

By repeated use of Result A.1, it follows that

$$\lim_{n\to\infty}[\boldsymbol{D}^{-1}R''(\boldsymbol{\xi}^0)\boldsymbol{D}^{-1}] = \boldsymbol{\Pi}, \tag{18}$$

as defined in (12). Note that $\boldsymbol{D}^{-1}R'(\boldsymbol{\xi}^0)$ is a $2P+1$ vector as follows:

$$\begin{aligned}\boldsymbol{D}^{-1}R'(\boldsymbol{\xi}^0) &= \boldsymbol{D}^{-1}\left[\frac{\partial R(\boldsymbol{\xi})}{\partial A_{11}}, \frac{\partial R(\boldsymbol{\xi})}{\partial B_{11}}, \ldots, \frac{\partial R(\boldsymbol{\xi})}{\partial A_{Mp_M}}, \frac{\partial R(\boldsymbol{\xi})}{\partial B_{Mp_M}}, \frac{\partial R(\boldsymbol{\xi})}{\partial \lambda}\right]^\top_{\boldsymbol{\xi}=\boldsymbol{\xi}^0} \\ &= \boldsymbol{D}^{-1}\left[\frac{\partial R(\boldsymbol{\xi}^0)}{\partial A_{11}}, \frac{\partial R(\boldsymbol{\xi}^0)}{\partial B_{11}}, \ldots, \frac{\partial R(\boldsymbol{\xi}^0)}{\partial A_{Mp_M}}, \frac{\partial R(\boldsymbol{\xi}^0)}{\partial B_{Mp_M}}, \frac{\partial R(\boldsymbol{\xi}^0)}{\partial \lambda}\right]^\top. \end{aligned} \tag{19}$$

Now to compute the right hand side of (19), let us use the following notations.

$$\eta_m(t; \boldsymbol{\alpha}_m^0, \lambda^0) = \sum_{j=1}^{p_m} jt[-A_{mj}^0 \sin(j\lambda^0 t) + B_{mj}^0 \cos(j\lambda^0 t)]; \quad m = 1, \ldots, M,$$

and $\boldsymbol{\eta}^0(t) = (\eta_1(t; \boldsymbol{\alpha}_1^0, \lambda^0), \ldots, \eta_M(t; \boldsymbol{\alpha}_M^0, \lambda^0))^\top$, for $t = 1, \ldots, n$. Then, using the M unit vectors $\{\boldsymbol{u}_m; m = 1, \ldots, M\}$, defined in Theorem 2.4, we have

$$\begin{aligned}
\frac{1}{\sqrt{n}} \frac{\partial R(\boldsymbol{\xi}^0)}{\partial A_{mj}} &= -2\boldsymbol{u}_m^\top \boldsymbol{\Sigma}^{-1} \sum_{t=1}^{n} \frac{\cos(j\lambda^0 t)}{\sqrt{n}} \boldsymbol{e}(t) = -2\boldsymbol{\sigma}^m \sum_{t=1}^{n} \frac{\cos(j\lambda^0 t)}{\sqrt{n}} \boldsymbol{e}(t); \\
\frac{1}{\sqrt{n}} \frac{\partial R(\boldsymbol{\xi}^0)}{\partial B_{mj}} &= -2\boldsymbol{u}_m^\top \boldsymbol{\Sigma}^{-1} \sum_{t=1}^{n} \frac{\sin(j\lambda^0 t)}{\sqrt{n}} \boldsymbol{e}(t) = -2\boldsymbol{\sigma}^m \sum_{t=1}^{n} \frac{\sin(j\lambda^0 t)}{\sqrt{n}} \boldsymbol{e}(t); \\
& \qquad\qquad\qquad\qquad\qquad\qquad j = 1, \ldots, p_m, \quad m = 1, \ldots, M, \\
\frac{1}{n^{3/2}} \frac{\partial R(\boldsymbol{\xi}^0)}{\partial \lambda} &= -\frac{2}{n^{3/2}} \sum_{t=1}^{n} (\boldsymbol{\eta}^0(t))^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{e}(t).
\end{aligned}$$

Since, all the elements of $\boldsymbol{D}^{-1} R'(\boldsymbol{\xi}^0)$ satisfy the Lindeberg Feller's conditions (Chung;1974), therefore,

$$\boldsymbol{D}^{-1} R'(\boldsymbol{\xi}^0) \xrightarrow{d} N_{2P+1}(\mathbf{0}, 2\boldsymbol{\Pi}),$$

where $\boldsymbol{\Pi}$ is same as defined in (12).

References

[1] Baldwin, A.J. and Thomson, P.J. (1978). Periodogram analysis of S. Carinae. *R. Astron. Soc., New Zealand* (Variable Star Section), 6, 31–38.

[2] Brown, E.N. Czeisler, C.A. (1992). The statistical analysis of circadian phase and amplitude in constant-routine core-temperature data. *J. Biol. Rhythms*, 7, 177–202.

[3] Brown, E.N. and Liuthardt, H. (1999). Statistical model building and model criticism for human circadian data. *J. Biol. Rhythms*, 14, 609–616.

[4] Chan, F.K.W., So, H.C., Sun, W. (2013), "Accurate estimation of common sinusoidal parameters in multiple channels", *Signal Processing*, vol. 93, 742 – 748.

[5] Christensen, M.G. (2012), "Multi-channel maximum likelihood pitch estimation", *Proceedings IEEE International Conference in ASSP*, pp 409 – 412.

[6] Griffin, A., Hirvonen, T., Tzagkarakis, C., Mouchtaris, A., Tsakalides, P. (2011), "Single-channel and multi-channel sinusoidal audio coding using compresses sensing", *IEEE Transactions on Signal Processing*, vol. 19, 1382 – 1395.

[7] Handel, P. and Host-Madsen, A. (1997), "Estimation of velocity and size particles from two channel laser anemometry measurement", *Measurement*, vol. 21, 113 – 123.

[8] Hannan, E.J. (1973), "The estimation of frequencies", *Journal of Applied Probability*, vol. 10, 510 – 519.

[9] Kundu, D. (1997), "Asymptotic Theory of the Least Squares Estimators of Sinusoidal Signal", *Statistics*, vol. 30, 221–238.

[10] Makino, S., Lee, T.W. and Sawada, H. (2007) Blind Speech Separation, Dordrecht, the Netherlands : Springer.

[11] Mangulis, V. (1965), *Handbook of series for scientists and engineers*, New York, Academic Press.

[12] Nandi, S. and Kundu, D. (2003), "Estimating the Fundamental Frequency of a Periodic Function", *Statistical Methods and Applications*, vol. 12, 341 – 360.

[13] Nandi, S. and Kundu, D. (2020), *Statistical Signal Processing: Frequency Estimation*, 2nd Edition, Springer, Singapore.

[14] Nandi, S. and Kundu, D. (2022), "Estimating parameters in multichannel sinusoidal model", *Circuits, Systems and Signal Processing*, vol. 41, 4604 – 4631.

[15] Quinn, B.G. and Thomson, P.J. (1991), "Estimating the frequency of a periodic function", *Biometrika*, vol. 78, 65 – 74.

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

[17] Sakai, H. (1993), "The determination of the number of terms in a multichannel sinusoidal regression", *Journal of Time Series Analysis*, vol. 14, 621 – 628.

[18] Tzagkarakis, C. and Mouchtaris, A. (2009) "A multichannel sinusoidal model applied to spot microphone signals for immersive audio", *IEEE Trans. Audio, Speech and Language Processing*, vol. 17(8), 1483-1497.

[19] Wang, D. and Brown, G.J. (2006) Computational Auditory Scene Analysis: Principles, Algorithms, and Applications, Hoboken, N.J. : Wiley Interscience.

[20] Wu, C. F. J. (1981), "Asymptotic Theory of Non Linear Squares Estimation", *Annals of Statistics*, vol. 9, 501–513.

[21] Zhou, Z. (2013), "Spectral analysis of sinusoidal signal from multiple channels", Ph.D thesis, City University of Hong Kong.

[22] Zhou, Z., Huang, L., Christensen, M.G. and Zhang, S. (2021), "Robust spectral analysis of multichannel sinusoidal signals in impulsive noise environments", *IEEE Transactions on Signal Processing*, DOI: 10.1109/TSP.2021.3101989.

[23] Zhou, Z., So, H.C. and Christensen, M.G. (2013), "Parameter modelling for damped sinusoids from multiple channel", *IEEE Transactions on Signal Processing*, vol. 61(15), 3895 – 3907.

[1]Theoretical Statistics and Mathematics Unit, Indian Statistical Institute, 7, S.J.S. Sansanwal Marg, New Delhi - 110016, India, nandi@isid.ac.in

[2]Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Pin 208016, India, kundu@iitk.ac.in