Robust Parameter Estimation in Two-Channel Sinusoidal Frequency Model

Debasis Kundu *

Abstract

In this article a robust parameter estimation method for two-channel sinusoidal model has been proposed. In a two-channel sinusoidal frequency model, the frequencies at different channels are same but they may have different amplitudes. The noise random variables in each channel are assumed to follow a stationary linear process, and the errors from the different channels are assumed to be dependent. The generalized least squares (GLS) estimators may be used in this case as they are most efficient estimators and in presence of the bivariate Gaussian errors, the asymptotic variances of the GLSEs attend the Cramer-Rao lower bound. It is well known that the GLS estimators are very sensitive to the outliers. In this article a very efficient weighted GLS (WGLS) estimators has been proposed, which are quite robust to the outliers. The asymptotic properties of the WGLS estimators have been provided. In simulation studies, it is observed that in presence of outliers the WGLS estimators perform better than the GLS estimators. The implementation of the GWLS estimators is quite straight forward. The performances of the WGLS estimators depend on the weight function. The proper choice of the weight function has been discussed and it has been illustrated using a synthetic data set.

Key Words: multichannel model; stationary error; generalized least squares estimators; sinusoidal model; robust estimators; consistency; asymptotic normality.

1 Introduction

In this article the robust estimation of the unknown parameters of the signal obtained from a two-channel sinusoidal model has been addressed. The two-channel sinusoidal model has

*Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Pin 208016, India. E-mail: kundu@iitk.ac.in, Phone no. 91-512-2597141, Fax no. 91-512-2597500.

the following form:

$$\left(\begin{array}{c} y_1(n) \\ y_2(n) \end{array}\right) = \left(\begin{array}{cc} A_1^0 & B_1^0 \\ A_1^2 & B_2^0 \end{array}\right)\left(\begin{array}{c} \cos(\omega^0 n) \\ \sin(\omega^0 n) \end{array}\right) + \left(\begin{array}{c} e_1(n) \\ e_2(n) \end{array}\right); \quad n = 1, \ldots, N. \tag{1}$$

Here $y_1(n)$ and $y_2(n)$ are the observed signals from the two channels, ω^0 is the common frequency of the two channels, (A_1^0, B_1^0) and (A_2^0, B_2^0) are the amplitude parameters, for the respective channel. Further, $e_1(n)$ and $e_2(n)$ are the additive errors associated with channel 1 and channel 2, respectively. We will address the problem of estimation of $(A_1^0, B_1^0, A_2^0, B_2^0, \omega^0)$, based on the observed signals from the two channels namely $\{(y_1(n), y_2(n)); n = 1, \ldots, N\}$.

The two channel model (1) has several applications in various fields, see for example Handel and Host-Madsen [6], Ramos et al. [15], Vucijak and Saranovac [23], Tzagkarakis et al. [22], Griffin et al. [4] and see the references cited there in. The problem was originally studied by Sakai [16], but in recent times it has received a significant amount of attention in the signal processing literature, see for example Clercq et al. [2], Sandgren et al. [17], Papy et al. [14], Handel [5], So and Zhou [18], Chan et al. [1], Zhou et al. [25], Stanović et al. [21], Nandi and Kundu [13] and the references cited therein.

Among different methods the most efficient methods are by So and Zhou [19] and by Nandi and Kundu [13]. So and Zhou [19] proposed the maximum likelihood (ML) estimators based on the assumption that the errors at the two channels are independent Gaussian random variables. Nandi and Kundu [13] proposed generalized least squares (GLS) estimators when the errors are not Gaussian. The GLS estimators become the ML estimators also under the assumption that $(e_1(n), e_2(n))$ follows a bivariate normal distribution. The GLS estimators are consistent and the variances of the estimators attend the Cramer-Rao (CR) bound when the errors are Gaussian. Some of the other proposed methods may be computationally efficient, but their large sample properties have not yet been established theoretically. One major deficiency in all these methods is that they may work well in the absence of any outliers, but the same is not true even in presence of few outliers. No attempt has been made so far to develop an robust estimation procedure in presence of outliers. Moreover, it is not immediate how to implement some of the standard robust estimation procedure like least absolute deviation method or Huber's M-estimator in this multichannel model, particularly when the errors in the two channels are dependent.

The main contribution in this paper is that we have provided a simple robust method of estimation of the unknown parameters of the two channel sinusoidal frequency model. The model can be implemented in practice more conveniently compared to some of the standard robust method of estimation namely LAD or Huber's M-estimators. The properties of the proposed estimators have been obtained under the same set of assumptions as the LSEs. It has been shown that the WLS estimators are strongly consistent, i.e. as the sample size increases the proposed estimators converge to the true parameter values. The same properties may be difficult to establish in case of LAD or Huber's M-estimators under the same set of assumptions. Further, the method can be easily extended for more than two channels also.

The proposed method is based on the weighted least squares (WLS) approach. The WLS method is not very uncommon, see for example Strutz [20], for a comprehensive review on WLS method and its implementation in practice. But here the WLS method has been used as an effective robust estimation procedure in presence of outliers. The basic idea of the proposed WLS method is very intuitive. It puts less weight to the residual squares where outliers are present, as against equal weights as adopted by the LS method, see for example Kundu [8], Kundu et al. [10, 11], Kundu and Grover [9] and the references cited there in. Usually a polynomial weight function can be chosen for this purpose. The implementation involves the same amount of computational complexity as the LS estimators. Where as least absolute deviation (LAD) estimators or Huber's M-estimators are computationally more challenging. Extensive simulations have been performed to show the effectiveness of the proposed method. In practice the choice of the penalty function is an important issues, and it has been addressed here. The proposed method has been illustrated using a synthetic data set. The results are quite satisfactory.

The paper has been organized as follows. In Section 2 the WGLS estimators, it's implementation and their asymptotic properties have been discussed. Extensive simulation results and the choice of the penalty function have been discussed in Section 3 and Section 4, respectively. The analysis of a synthetic data set has been presented in Section 5 for illustrative purposes, and finally the conclusion appears in Section 6. All the proofs have been presented in Appendix A and B, and some more simulation results have been presented in Appendix C.

2 Weighted Generalized Least Squares Estimators

2.1 Model Formulation and GLS Estimators

The model (1) has the following matrix form:

$$\boldsymbol{y}(n) = \boldsymbol{A}^0\boldsymbol{\theta}(\omega^0, n) + \boldsymbol{e}(n) = \boldsymbol{\mu}(n; \boldsymbol{\beta}^0, \omega^0) + \boldsymbol{e}(n), \tag{2}$$

where

$$\boldsymbol{y}(n) = \begin{bmatrix} y_1(n) \\ y_2(n) \end{bmatrix}, \quad \boldsymbol{A}^0 = \begin{bmatrix} A_1^0 & B_1^0 \\ A_1^2 & B_2^0 \end{bmatrix}, \boldsymbol{\theta}(\omega^0, n) = \begin{bmatrix} \cos(\omega^0 n) \\ \sin(\omega^0 n) \end{bmatrix}, \quad \boldsymbol{e}(n) = \begin{bmatrix} e_1(n) \\ e_2(n) \end{bmatrix}, \tag{3}$$

$$\boldsymbol{\beta}^0 = \begin{bmatrix} A_1^0 \\ B_1^0 \\ A_2^0 \\ B_2^0 \end{bmatrix}, \quad \boldsymbol{\mu}(n; \boldsymbol{\beta}^0, \omega^0) = \begin{bmatrix} \beta_1(n; \boldsymbol{\beta}^0, \omega^0) \\ \beta_2(n; \boldsymbol{\beta}^0, \omega^0) \end{bmatrix}. \tag{4}$$

Here

$$y_k(n) = A_k^0 \cos(\omega^0 n) + B_k^0 \sin(\omega^0 n) + e_k(n); \quad n = 1, \ldots, N, \tag{5}$$

$|A_k^0| > 0$, $|B_k^0| > 0$ for $k = 1$ and 2, and $\omega^0 \in (0, \pi)$. The error $\boldsymbol{e}(n)$ satisfies the following assumption.

Assumption 1: The error vector $\boldsymbol{e}(n)$ has the form

$$\boldsymbol{e}(n) = \sum_{m=-\infty}^{\infty} a(m)\boldsymbol{u}(n-m), \tag{6}$$

here the random vectors $\boldsymbol{u}(m)$'s are independently and identically distributed (i.i.d.) random variables with mean vector $\mathbf{0}$ and the variance-covariance matrix $\boldsymbol{\Sigma}$, which is non-singular. The scalar coefficients $a(m)$ satisfy

$$\sum_{m=-\infty}^{\infty} |a(m)| < \infty. \tag{7}$$

If we denote $\boldsymbol{u}(n) = (u_1(n), u_2(n))^\top$, then $e_k(n) = \sum_{m=-\infty}^{\infty} a(m)u_k(n-m)$, for $k = 1$ and 2. Let us denote the $(i,j)-th$ element of the matrix $\boldsymbol{\Sigma}$ by σ_{ij}, for $i,j = 1,2$. Then for $n,k = 1,2,\ldots$

$$\begin{aligned} Corr(e_1(n), e_2(n)) &= \frac{\sigma_{12}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{22}}}, & (8)\\ Corr(e_1(n), e_2(n+k)) &= \frac{\sigma_{12}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{22}}} \times \frac{\sum_{m=-\infty}^{\infty} a(m)a(m+k)}{\sum_{m=-\infty}^{\infty} a^2(m)}. & (9) \end{aligned}$$

It should be noted that the above error assumption is a very general one, and it has not been considered so far in analyzing multichannel sinusoidal model. In most of the cases the errors in two channels are taken to be independently distributed. Nandi and Kundu [13] assumed that $e_1(n)$ and $e_2(n)$ are correlated, but $\boldsymbol{e}(n)$'s are assumed to be i.i.d. The GLS estimators of $\boldsymbol{\xi}^0 = (\boldsymbol{\beta}^0, \omega^0)$ can be obtained by minimizing

$$Q(\boldsymbol{\xi}) = \sum_{n=1}^{N} \left(\boldsymbol{y}(n) - \boldsymbol{\mu}(n;\boldsymbol{\xi})\right)^\top \boldsymbol{\Sigma}^{-1} \left(\boldsymbol{y}(n) - \boldsymbol{\mu}(n;\boldsymbol{\xi})\right), \tag{10}$$

with respect to the unknown parameters. It can be shown that when the errors are Gaussian then the asymptotic variances of the GLS estimator attend the Cramer-Rao bound. Now we propose WGLS estimators, which can be used as robust estimators.

2.2 WGLS Estimators: Methodology

Let a k-degree polynomial $w(t)$ has the following form:

$$w(t) = c_0 + c_1 t + \ldots + c_k t^k; \quad t \geq 0, \tag{11}$$

where $c_1, \ldots, c_k$ satisfy the following conditions

$$\inf_{0 \leq t \leq 1} w(t) \geq \gamma > 0. \tag{12}$$

For known $\boldsymbol{\Sigma}$, the WGLS estimators are the argument minimum of

$$R(\boldsymbol{\xi}) = R(\boldsymbol{\beta}, \omega) = \sum_{n=1}^{N} w\left(\frac{n}{N}\right) (\boldsymbol{y}(n) - \boldsymbol{\mu}(n; \boldsymbol{\beta}, \omega))^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{y}(n) - \boldsymbol{\mu}(n; \boldsymbol{\beta}, \omega)). \tag{13}$$

The WGLS estimator of $\boldsymbol{\beta}$, say $\widehat{\boldsymbol{\beta}}(\omega)$, are the solutions of

$$\frac{\partial}{\partial A_1} R(\boldsymbol{\beta}, \omega) = 0, \quad \frac{\partial}{\partial B_1} R(\boldsymbol{\beta}, \omega) = 0, \quad \frac{\partial}{\partial A_2} R(\boldsymbol{\beta}, \omega) = 0, \quad \frac{\partial}{\partial B_2} R(\boldsymbol{\beta}, \omega) = 0. \tag{14}$$

It can be shown that (14) is equivalent to:

$$(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{M}_N(\omega))\boldsymbol{\beta}^\top(\omega) = (\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{I}_2)\boldsymbol{W}_N(\omega), \tag{15}$$

here '$\otimes$' denotes the Kronecker product, $\boldsymbol{I}_2$ denotes the 2×2 identity matrix,

$$\boldsymbol{M}_N(\omega)) = \frac{2}{N} \begin{bmatrix} \sum_{n=1}^N w\left(\frac{n}{N}\right) \cos^2(\omega n) & \sum_{n=1}^N w\left(\frac{n}{N}\right) \cos(\omega n)\sin(\omega n) \\ \sum_{n=1}^N w\left(\frac{n}{N}\right) \cos(\omega n)\sin(\omega n) & \sum_{n=1}^N w\left(\frac{n}{N}\right) \sin^2(\omega n) \end{bmatrix} \tag{16}$$

and

$$\boldsymbol{W}_N(\omega)) = \frac{2}{N} \begin{bmatrix} \sum_{n=1}^N w\left(\frac{n}{N}\right) y_1(n)\cos(\omega n) \\ \sum_{n=1}^N w\left(\frac{n}{N}\right) y_1(n)\sin(\omega n) \\ \sum_{n=1}^N w\left(\frac{n}{N}\right) y_2(n)\cos(\omega n) \\ \sum_{n=1}^N w\left(\frac{n}{N}\right) y_2(n)\sin(\omega n) \end{bmatrix}. \tag{17}$$

Hence

$$\begin{aligned} \widehat{\boldsymbol{\beta}}^\top(\omega) &= (\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{M}_N(\omega))^{-1}(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{I}_2)\boldsymbol{W}_N(\omega) \\ &= (\boldsymbol{\Sigma} \otimes (\boldsymbol{M}_N(\omega))^{-1})(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{I}_2)\boldsymbol{W}_N(\omega) \\ &= (\boldsymbol{I}_2 \otimes (\boldsymbol{M}_N(\omega))^{-1})\boldsymbol{W}_N(\omega). \end{aligned} \tag{18}$$

Therefore, the WGLS estimator of ω^0, say $\widehat{\omega}$, is the argument minimum of $R(\widehat{\boldsymbol{\beta}}(\omega), \omega)$. The WGLS estimator of $\boldsymbol{\beta}^0$, is $\widehat{\boldsymbol{\beta}} = \widehat{\boldsymbol{\beta}}(\widehat{\omega})$.

Now let us observe that

$$\begin{aligned}
\lim_{N\to\infty} \frac{2}{N} \sum_{n=1}^{N} w\left(\frac{n}{N}\right) \cos^2(\omega n) &= \int_0^1 w(t)dt \\
\lim_{N\to\infty} \frac{2}{N} \sum_{n=1}^{N} w\left(\frac{n}{N}\right) \sin^2(\omega n) &= \int_0^1 w(t)dt \\
\lim_{N\to\infty} \frac{2}{N} \sum_{n=1}^{N} w\left(\frac{n}{N}\right) \sin(\omega n)\cos(\omega n) &= 0. \qquad (19)
\end{aligned}$$

Hence, for large N, $\widehat{\boldsymbol{\beta}}^\top(\omega) = \boldsymbol{W}_N(\omega)$. The approximate WGLS estimator of ω^0, say $\widetilde{\omega}$, is the argument minimum of $R(\widehat{\boldsymbol{W}}(\omega), \omega)$.

It may be mentioned that we have proposed to choose the k-th degree polynomial as the weight function because it can take variety of shapes depending on the degree and coefficients of the polynomials. We need flexibility on shapes of the weight function so that the estimators work effectively in practice. Further, because of the polynomial weight function it is possible to establish the properties of the proposed estimators under the same set of assumptions as the LS estimators.

2.3 WGLS Estimators: Asymptotic Properties

We need the following assumptions.

Assumption 2: Suppose $0 < K < \infty$ and $0 < |A_1^0|, |B_1^0|, |A_2^0|, |B_2^0| < K$, and $0 < \omega^0 < \pi$.

Theorem 1: If the error random vector $\{\boldsymbol{e}(n)\}$ satisfies Assumption 1, and the model parameters satisfy Assumption 2, then $\widehat{\boldsymbol{\xi}} = (\widehat{\boldsymbol{\beta}}^\top, \widehat{\omega})^\top$ is a strongly consistent estimator of $\boldsymbol{\xi}^0$.

Proof: See Appendix A.

Let us use the notations now. For $p = 0,1,2, \ldots,$

$$u_p = \int_0^1 t^p w(t) dt \quad \text{and} \quad v_p = \int_0^1 t^p w^2(t) dt. \tag{20}$$

The 5×5 diagonal matrix $\boldsymbol{D}$ as

$$\boldsymbol{D} = \text{diag}\{N^{1/2}, N^{1/2}, N^{1/2}, N^{1/2}, N^{3/2}\}, \tag{21}$$

the elements of the 5×5 matrix $\boldsymbol{\Gamma} = ((\gamma_{ij}))$ are

$$\begin{aligned}
\gamma_{11} &= \gamma_{22} = \sigma^{11} u_0, \quad \gamma_{33} = \gamma_{44} = \sigma^{22} u_0, \quad \gamma_{13} = \gamma_{24} = \gamma_{31} = \gamma_{42} = \sigma^{21} u_0, \\
\gamma_{12} &= \gamma_{21} = \gamma_{14} = \gamma_{41} = \gamma_{23} = \gamma_{32} = \gamma_{34} = \gamma_{43} = 0 \\
\gamma_{15} &= \gamma_{51} = u_1(B_1^0 \sigma^{11} + B_2^0 \sigma^{12}), \quad \gamma_{25} = \gamma_{52} = -u_1(A_1^0 \sigma^{11} + A_2^0 \sigma^{12}) \\
\gamma_{35} &= \gamma_{53} = u_1(B_2^0 \sigma^{22} + B_1^0 \sigma^{12}), \quad \gamma_{45} = \gamma_{54} = -u_1(A_2^0 \sigma^{22} + A_1^0 \sigma^{12}) \\
\gamma_{55} &= u_2\{\sigma^{11}(A_1^{0^2} + B_1^{0^2}) + \sigma^{22}(A_2^{0^2} + B_2^{0^2}) + 2\sigma^{12}(A_1^0 A_2^0 + B_1^0 B_2^0)\},
\end{aligned} \tag{22}$$

here we denote the $(i,j)-th$ element of the matrix $\boldsymbol{\Sigma}^{-1}$ by σ^{ij}, and the elements of the 5×5 matrix $\boldsymbol{\Psi}$ are same as the elements of the $\boldsymbol{\Gamma}$ matrix, where all the u_i's are replaced by v_i's for $i = 0,1,2$. It may be seen that the matrices $\boldsymbol{\Gamma}$ and $\boldsymbol{\Psi}$ have the following forms:

$$\boldsymbol{\Gamma} = \begin{bmatrix} \boldsymbol{\Gamma}_{11} & \boldsymbol{a} \\ \boldsymbol{a}^\top & b \end{bmatrix} \quad \text{and} \quad \boldsymbol{\Psi} = \begin{bmatrix} \boldsymbol{\Psi}_{11} & \boldsymbol{c} \\ \boldsymbol{c}^\top & d \end{bmatrix}, \tag{23}$$

where $\boldsymbol{\Gamma}_{11} = u_0 \boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{I}_2$, $\boldsymbol{a}^\top = (\gamma_{15}, \gamma_{25}, \gamma_{35}, \gamma_{45})$ and $b = \gamma_{55}$. Similarly, $\boldsymbol{\Psi}_{11}$, $\boldsymbol{c}$ and d are also defined from $\boldsymbol{\Gamma}_{11}$, $\boldsymbol{a}$ and b, by replacing u_i's by v_i's.

Theorem 2: If the matrices $\boldsymbol{\Gamma}$ and $\boldsymbol{\Psi}$ as defined above are non-singular, then

$$\Big((\widehat{A}_1 - A_1^0), (\widehat{B}_1 - B_1^0), (\widehat{A}_2 - A_2^0), (\widehat{B}_2 - B_2^0), (\widehat{\omega} - \omega^0)\Big) \boldsymbol{D} \xrightarrow{d} N_5(\boldsymbol{0}, 2\delta \boldsymbol{\Gamma}^{-1} \boldsymbol{\Psi} \boldsymbol{\Gamma}^{-1}). \tag{24}$$

Here, '$\xrightarrow{d}$' denotes distribution convergence and $N_5(\boldsymbol{0}, \boldsymbol{A})$ is a multivariate (5-variate) normal distribution with mean vector $\boldsymbol{0}$, and the variance-covariance matrix $\boldsymbol{A}$. Further,

$$\delta = \left[\sum_{k=-\infty}^{\infty} a(k) \cos(\omega^0 k)\right]^2 + \left[\sum_{k=-\infty}^{\infty} a(k) \sin(\omega^0 k)\right]^2 = \left|\sum_{k=-\infty}^{\infty} a(k) e^{i\omega^0 k}\right|^2. \tag{25}$$

Proof: See Appendix B.

2.4 WGLS Estimators: Practical Implementation

In the previous sections we have discussed about the WGLS estimators and their asymptotic properties when $\mathbf{\Sigma}$ is known. But in practice often it is not known. Therefore, we need to estimate $\mathbf{\Sigma}$ efficiently, then only WGLS estimators can be used. An efficient estimator of the matrix $\mathbf{\Sigma}$ is discussed here. We need the following results for that purpose. Let

$$S(\boldsymbol{\xi}) = S(\boldsymbol{\beta}, \omega) = \sum_{n=1}^{N} w\left(\frac{n}{N}\right)\left(\boldsymbol{y}(n) - \boldsymbol{\mu}(n; \boldsymbol{\beta}, \omega)\right)^{\top}\left(\boldsymbol{y}(n) - \boldsymbol{\mu}(n; \boldsymbol{\beta}, \omega)\right). \tag{26}$$

The weight function $w(t)$ is same as in (11). If $\widetilde{\boldsymbol{\xi}} = (\widetilde{\boldsymbol{\beta}}, \widetilde{\omega})$ minimizes (26), then it is called weighted least squares (WLS) estimator of $\boldsymbol{\xi}^0$. For fixed ω, the WLS estimator $\boldsymbol{\beta}$, say $\widetilde{\boldsymbol{\beta}}(\omega)$, is same as $\widehat{\boldsymbol{\beta}}(\omega)$,as defined in (18). Therefore, WLS estimator of ω can be obtained by minimizing $I(\omega)$, where

$$I(\omega) = S(\widehat{\boldsymbol{\beta}}(\omega), \omega). \tag{27}$$

The minimum value of $I(\omega)$ at $\dfrac{2\pi j}{N}$, for $j = 1, \ldots, N$, can be used as an effective initial estimate of ω.

The following results can be established similarly as Theorems 1 and 2.

Theorem 3: If Assumptions 1 and 2 are satisfied then $\widetilde{\boldsymbol{\xi}} = (\widetilde{\boldsymbol{\beta}}, \widetilde{\omega})$ is a strongly consistent estimator of $\boldsymbol{\xi}^0$.

Let us define the 5×5 matrices $\boldsymbol{U} = ((u_{ij}))$ and $\boldsymbol{V}$ as follows

$$\boldsymbol{U} = \begin{bmatrix} u_0 & 0 & 0 & 0 & u_{15} \\ 0 & u_0 & 0 & 0 & u_{25} \\ 0 & 0 & u_0 & 0 & u_{35} \\ 0 & 0 & 0 & u_0 & u_{45} \\ u_{51} & u_{52} & u_{53} & u_{54} & u_{55} \end{bmatrix}, \quad \boldsymbol{V} = \begin{bmatrix} v_0\sigma_{11} & 0 & v_0\sigma_{12} & 0 & v_{15} \\ 0 & v_0\sigma_{11} & 0 & v_0\sigma_{12} & v_{25} \\ v_0\sigma_{21} & 0 & v_0\sigma_{22} & 0 & v_{35} \\ 0 & v_0\sigma_{21} & 0 & v_0\sigma_{22} & v_{45} \\ v_{51} & v_{52} & v_{53} & v_{54} & v_{55} \end{bmatrix} \tag{28}$$

here

$$u_{15} = u_{51} = u_1 B_1^0, \quad u_{25} = u_{52} = -u_1 A_1^0, \quad u_{35} = u_{53} = u_1 B_1^0, \quad u_{45} = u_{54} = -u_1 A_2^0,$$

$$u_{55} = u_2(A_1^{0^2} + B_1^{0^2} + A_2^{0^2} + B_2^{0^2}),$$

and

$$v_{15} = v_{51} = v_1(\sigma_{11}B_1^0 + \sigma_{12}B_2^0), \quad v_{25} = v_{52} = -v_1(\sigma_{11}A_1^0 + \sigma_{12}A_2^0,$$

$$v_{35} = v_{53} = v_1(\sigma_{22}B_2^0 + \sigma_{21}B_1^0), \quad v_{45} = v_{54} = -v_1(\sigma_{22}A_2^0 + \sigma_{12}A_1^0),$$

$$v_{55} = v_2(\sigma_{11}(A_1^{0^2} + B_1^{0^2}) + \sigma_{22}(A_2^{0^2} + B_2^{0^2}) + 2\sigma_{12}(A_1^0A_2^0 + B_1^0B_2^0)).$$

Theorem 4: If the matrices $\boldsymbol{U}$ and $\boldsymbol{V}$ are non-singular then

$$\Big((\widetilde{A}_1 - A_1^0), (\widetilde{B}_1 - B_1^0), (\widetilde{A}_2 - A_2^0), (\widetilde{B}_2 - B_2^0), (\widetilde{\omega} - \omega^0)\Big)\,\boldsymbol{D} \xrightarrow{d} N_5(\mathbf{0}, 2\delta\boldsymbol{U}^{-1}\boldsymbol{V}\boldsymbol{U}^{-1}).$$

From Theorems 3 and 4, it follows that $\widetilde{\boldsymbol{\Sigma}}$, where

$$\widetilde{\boldsymbol{\Sigma}} = \frac{1}{N}\sum_{n=1}^{N}\Big(\boldsymbol{y}(n) - \boldsymbol{\mu}(n;\widetilde{\beta},\widetilde{\omega})\Big)\Big(\boldsymbol{y}(n) - \boldsymbol{\mu}(n;\widetilde{\beta},\widetilde{\omega})\Big)^{\top}$$

is a consistent estimator of $\boldsymbol{\Sigma}$. Hence, it can be used to compute WGLS estimators.

3 Simulation Results

In this section we present the results of the simulations to show how the proposed methods behave in presence We have performed some simulation experiments to see the effectiveness of the proposed methods in presence of outliers. We have considered the following parameter values:

$$A_1^0 = B_1^0 = 3.0, \quad A_2^0 = B_2^0 = 2.0 \quad \omega^0 = 0.75\pi. \tag{29}$$

The error random vector $\boldsymbol{e}(n)$'s are assumed to be i.i.d. bivariate random variables with mean vector $\mathbf{0}$ and the dispersion matrix $\boldsymbol{\Sigma}$, where for

$$\textbf{Case 1}: \boldsymbol{\Sigma} = \begin{bmatrix} 1 & \sqrt{3}/2 \\ \sqrt{3}/2 & 1 \end{bmatrix} \quad \text{and} \quad \textbf{Case 2}: \boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{30}$$

In both cases the sample size varies from 100 to 1000, and different proportion of outliers namely 10%, 20%, 30% and 40% are present and they are in the middle. The outliers are also normally distributed with mean zero and variance 100. Two different weight functions have been considered as the following for $0 \leq t \leq 1$,

$$w_1(t) = \frac{1}{12} + 11\left(t - \frac{1}{2}\right)^2 \quad \text{and} \quad w_2(t) = \begin{cases} 2-4t & \text{if} \quad 0 \leq t \leq \frac{1}{4} \\ 1 & \text{if} \quad \frac{1}{4} \leq t \leq \frac{3}{4} \\ 4t-2 & \text{if} \quad \frac{3}{4} \leq t \leq 1. \end{cases}$$

For comparison purposes, we have also considered GLS estimators and LAD and Huber-M estimators. It should be mentioned that the LAD and Huber-M estimators have not been tried for two channel model, and we have adopted the following procedure to compute these estimators. LAD estimators have been obtained by minimizing

$$L(\boldsymbol{\xi}) = \sum_{n=1}^{N} \left\{ |y_1(n) - \mu_1(n;\boldsymbol{\xi})| + |y_2(n) - \mu_2(n;\boldsymbol{\xi})| \right\} \tag{31}$$

and the Huber-M estimators have been obtained by minimizing

$$H(\boldsymbol{\xi}) = \sum_{n=1}^{N} \rho \left\{ (y_1(n) - \mu_1(n;\boldsymbol{\xi}))^2 + (y_2(n) - \mu_2(n;\boldsymbol{\xi}))^2 \right\}, \tag{32}$$

here the ρ is of the following form:

$$\rho(x) = \begin{cases} \frac{1}{2c}x^2 & \text{if} \quad |x| < c \\ |x| - \frac{1}{2}c^2 & \text{if} \quad |x| \geq c. \end{cases} \tag{33}$$

The constant c is known as the tuning parameter, and we have taken $c = 1$ in our simulations. It should be mentioned that the asymptotic properties of the LAD or Huber-M estimators are not known yet for two channel model and moreover the minimization of (32) and (33) involve 5-dimensional optimization problem. The biases and the mean squared errors (MSEs) have been computed for the frequency and the amplitudes based on 10000 replications in both the cases. Since the biases are negligible we have reported only the MSEs. The results for Case 1 have been reported in Figures 1, 2 and 3 and for Case 2, in Figures 4, 5 and 6.

It is observed from the simulation studies that the MSEs decrease as the sample size increases in all cases. The performances of the two WGLS estimators are very similar. It may not be very surprising because in both these cases the weight functions put less weight

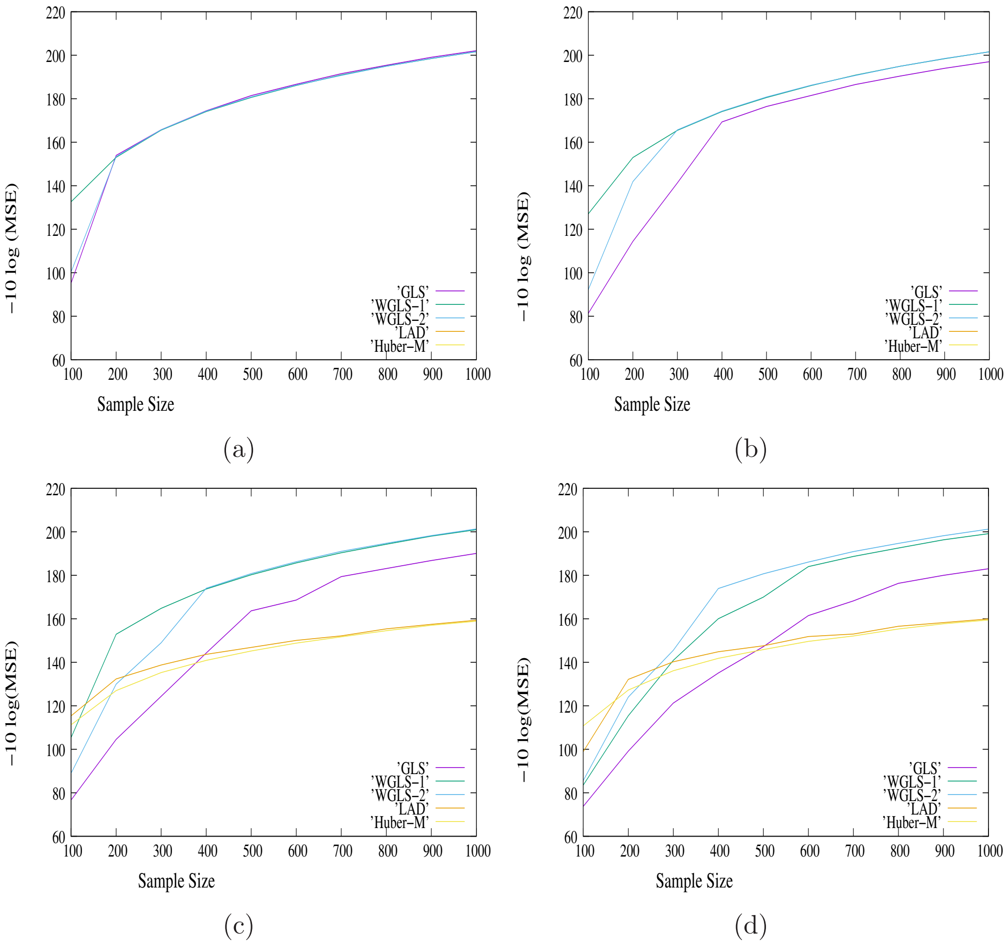
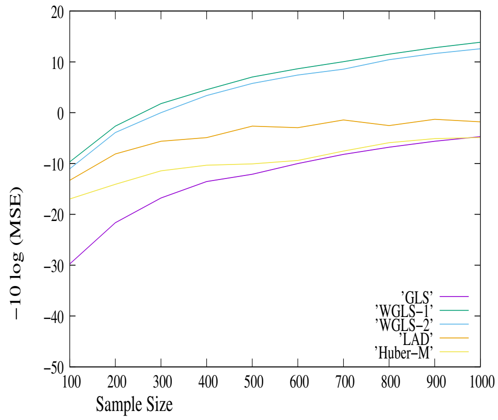
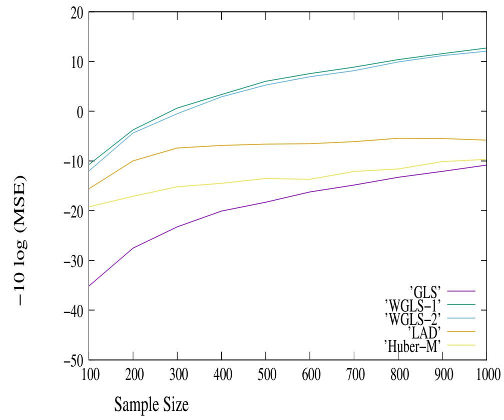


Figure 1: -10log(MSE) of the frequency in presence of outliers: (a) 10% outliers, (b) 20% outliers, (c) 30% outliers, (d) 40% outliers for different estimators.

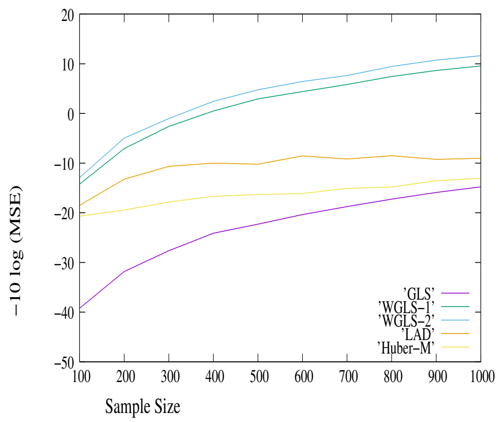
in the middle portion of the data set, and the outliers are present in this portion only. When the proportion of outliers is small, then the GLS and WGLS estimators behave in a very similar manner, although when the proportion of outliers is large WGLS estimators behave much better than the GLS estimators in terms of lower MSEs. Both the robust estimators do not perform that well in case of frequency estimation, although they perform better than the GLS estimators for amplitude estimation. Comparing the two cases it has been observed that the performance of all the estimators are better in terms of lower mean squared errors when the error distribution in two channels are independent compared to when they are



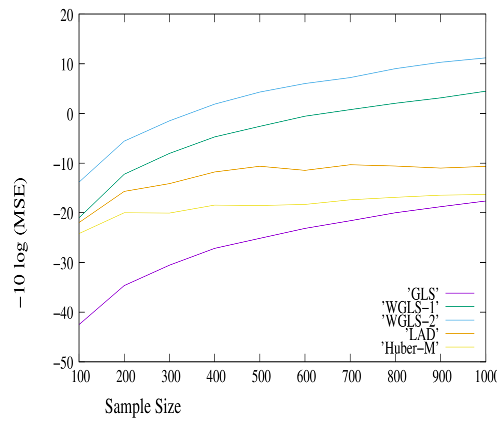
(a)



(b)



(c)



(d)

Figure 2: -10log(MSE) of the first amplitude estimators in presence of outliers: (a) 10% outliers, (b) 20% outliers, (c)30% outliers, (d) 40% outliers for different estimators.

independent, although pattern remains the same.

From these simulation experiments it is clear that if it is known where the outliers are present, then the performances of the WGLS estimators are very good. It may be mentioned that we have performed simulations with other set of parameter values and error structure, but since the performances are very similar in nature they have not been reported due to page restriction. We have reported some more simulation results in Appendix C taking several other weight functions mainly to show how the performances of the estimators depend on the choice of the weight function.

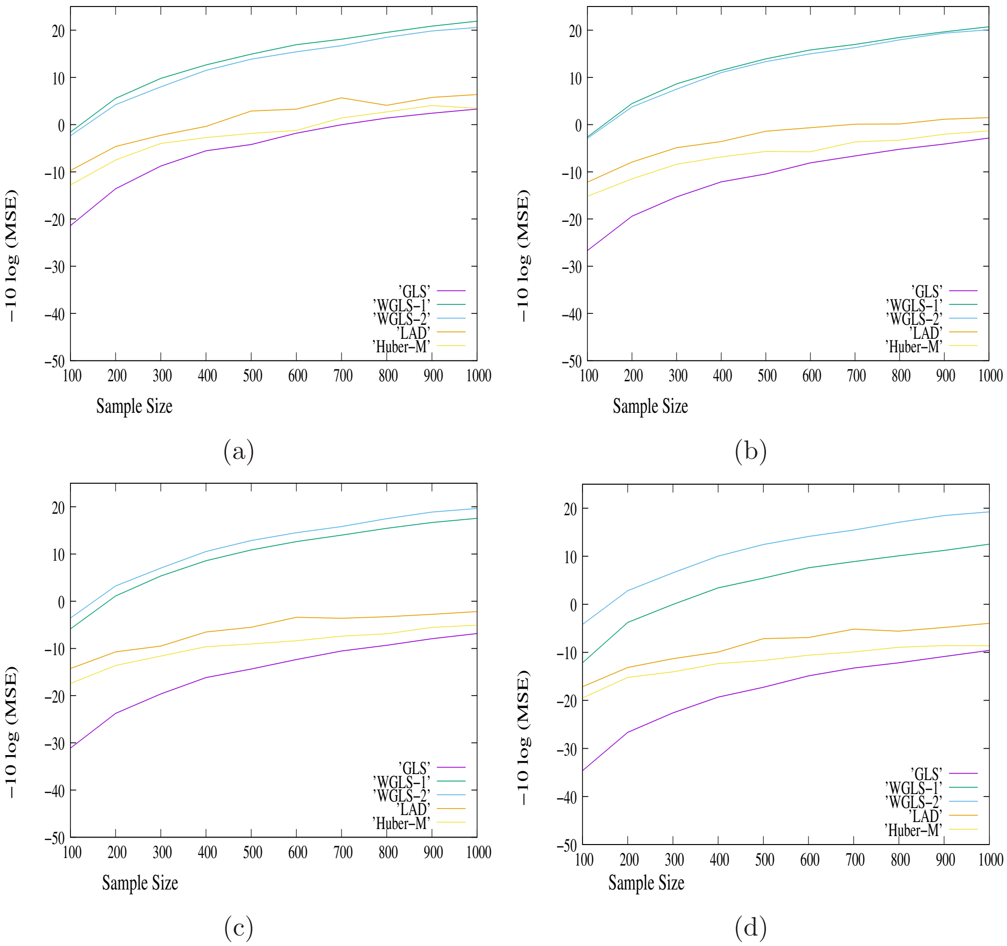


Figure 3: -10log(MSE) of the second amplitude estimators in presence of outliers: (a) 10% outliers, (b) 20% outliers, (c) 30% outliers, (d) 40% outliers for different estimators.

4 Choice of the Weight Function

In practice the weight function is not known and one needs to choose it. It very much depends where the outliers are present in the data set. It is clear that if we have some knowledge before hand that where the outliers are present, then accordingly the weight function can be chosen. For example, if it is known that the outliers are at the beginning of the time scale, then an increasing function or if it is known that they are at the end of the time scale then a decreasing function can be chosen. Similarly, if it is know that the outliers are present in the

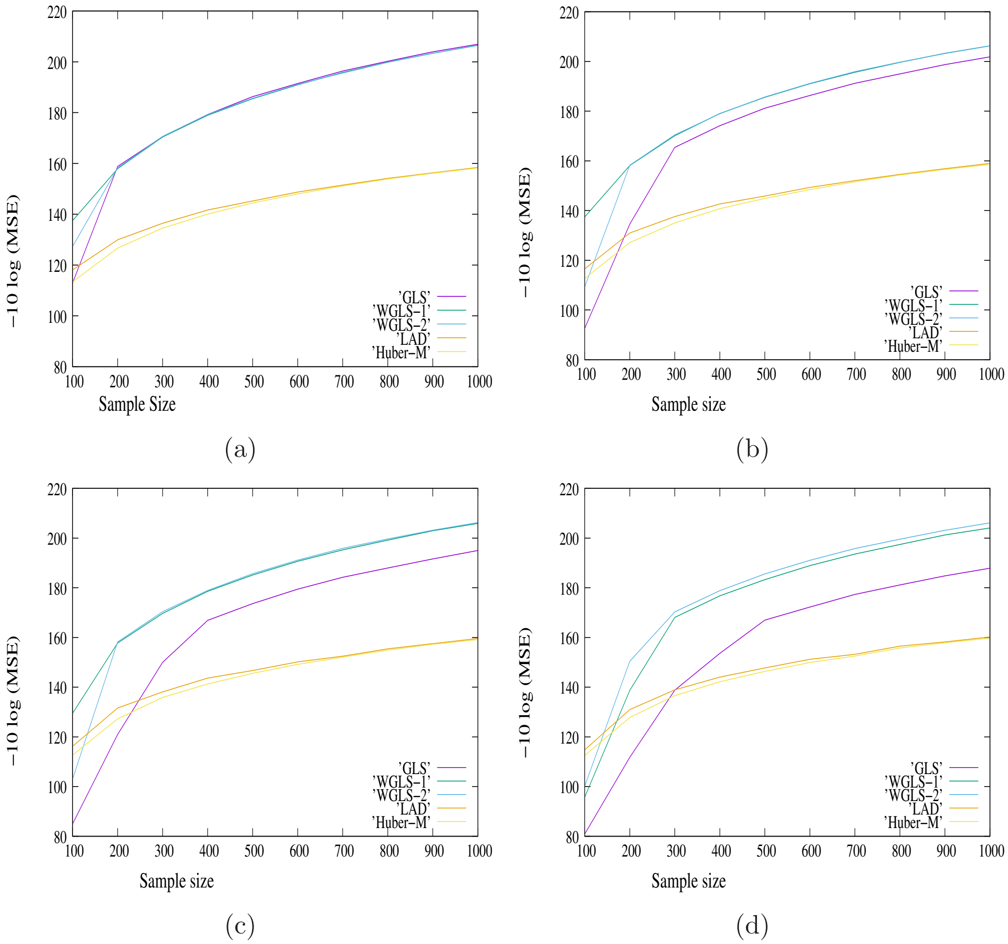


Figure 4: -10log(MSE) of the frequency in presence of outliers: (a) 10% outliers, (b) 20% outliers, (c) 30% outliers, (d) 40% outliers for different estimators.

middle then a V-shaped weight function can be chosen. In some situation it may be known, but often it may be unknown.

Now we will discuss about the choice of the weight function in practice if it is not known where the outliers are present. In this case we propose a class of weight functions which satisfy the condition (12), and they all normalized, i.e. $\int_0^1 w(t)dt = 1$. We choose 5-6 such weight functions of different shapes. Let us recall that the GLS estimators also can be obtained as a special case of WGLS estimators, when $w(t) = 1$. Hence, in this class of weight functions, we keep the weight function $w(t) = 1$, also. We compute the WGLS estimators for

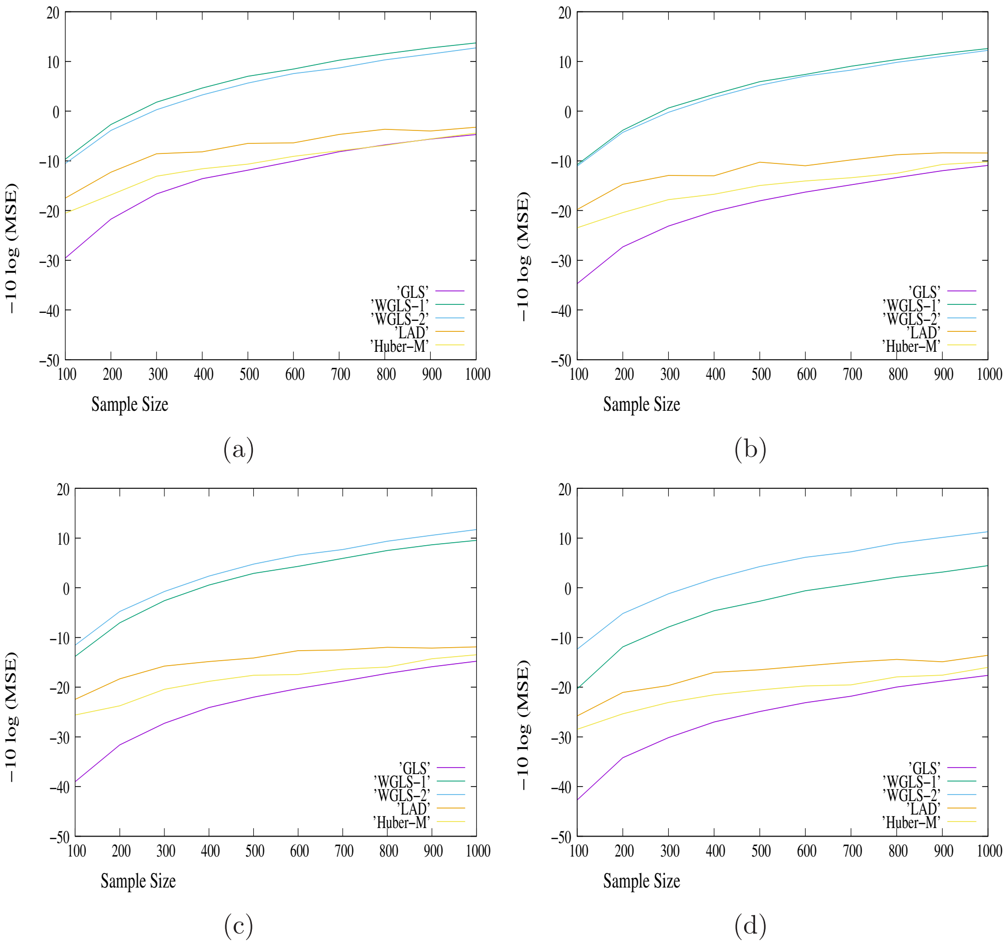


Figure 5: -10log(MSE) of the first amplitude estimators in presence of outliers: (a) 10% outliers, (b) 20% outliers, (c)30% outliers, (d) 40% outliers for different estimators.

all the weight functions of this class, and choose that estimators which produce the smallest weighted residual sums of squares. It is expected that if there are no outliers in the data set, then it should pickup the GLS estimators as the best weight function, otherwise it should pick up that weight function which puts less weight where outliers are present in the data set. The details are explained in the synthetic data analysis section.

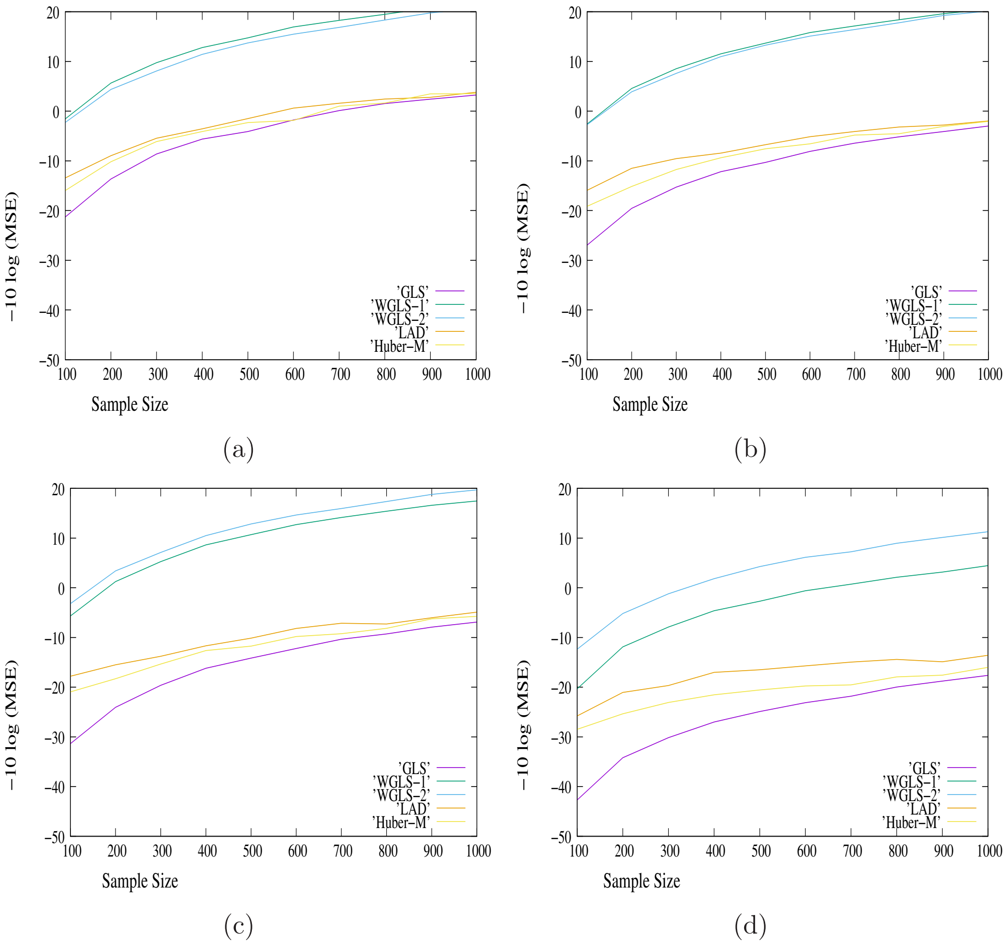


Figure 6: -10log(MSE) of the first amplitude estimators in presence of outliers: (a) 10% outliers, (b) 20% outliers, (c)30% outliers, (d) 40% outliers for different estimators.

5 Analysis of a Synthetic Data Set

A data has been generated using the model parameters:

$$A_1^0 = B_1^0 = 3.0, \quad A_2^0 = B_2^0 = 2.0, \quad \omega^0 = 0.75\pi, \quad N = 300.$$

It is assumed that the noise vector $\boldsymbol{e}(n)$ has the following structure

$$\boldsymbol{e}(n) \stackrel{d}{=} 0.8 N_2(\mathbf{0}, \boldsymbol{\Sigma}) + 0.2 N_2(\mathbf{0}, 100\ \boldsymbol{\Sigma}),$$

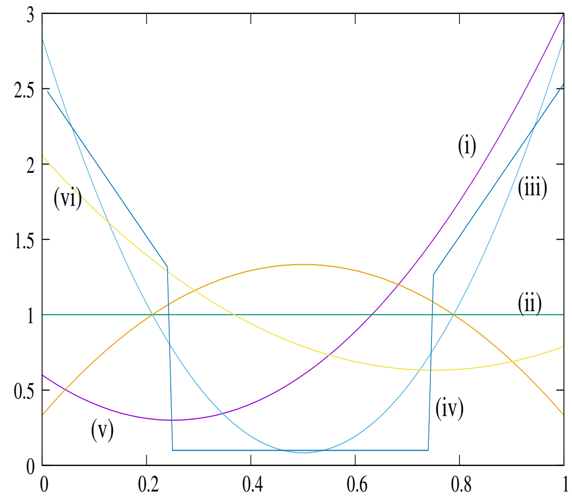


Figure 7: Different weight functions.

here $\boldsymbol{\Sigma}$ has the same form as in (30) and $\boldsymbol{e}(n)$'s are i.i.d. random variables. Hence, there are 20% outliers present in the data set, and the position of the outliers are also not known in advance. Therefore, it is not clear what kind of weight function would work in this case. For this reason we have taken a class of weight functions as follows:

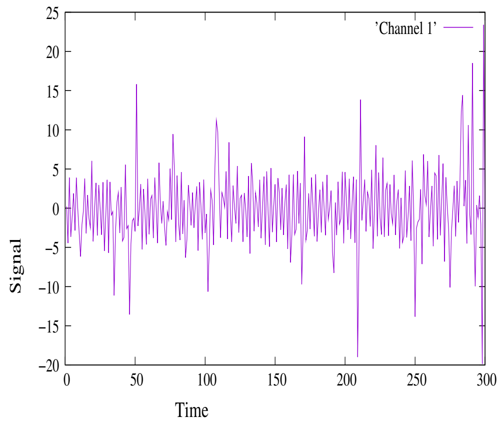
$$w_1(t) = 0.3 + 4.8(t - 0.25)^2, \quad w_2(t) = 1.0, \quad w_3(t) = \frac{1}{12} + 11(t - 0.5)^2$$

$$w_4(t) = \begin{cases} \frac{19}{15}(2 - 4t) & \text{if} \quad 0 \le t \le 0.25 \\ 0.1 & \text{if} \quad 0.25 \le t \le 0.75 \\ \frac{19}{15}(4t - 2) & \text{if} \quad 0.75 \le t \le 1.00 \end{cases}$$

$$w_5(t) = \frac{4}{3} - 4(t - 0.5)^2, \quad w_6(t) = \frac{12}{19} + \frac{48}{19}(t - 0.75)^2.$$

In Figure 7 the weight functions have been plotted. The plot of the signals at two different channels are provided in Figure 8. From Figure 8 it is quite apparent that there are outliers in the data set. The plot of $1/I(\omega)$ at the Fourier frequencies, for $I(\omega)$ as defined in (27), for $w(t) = 1$ is provided in Figure 9. It clearly provides an initial estimate of ω.

Estimates of the amplitudes and the frequency and the associated weighted residual sums of squares (WRSS) based on different weight functions are presented in Table 1. Based on the WRSS, we suggest to use the weight function (v) for this data set.



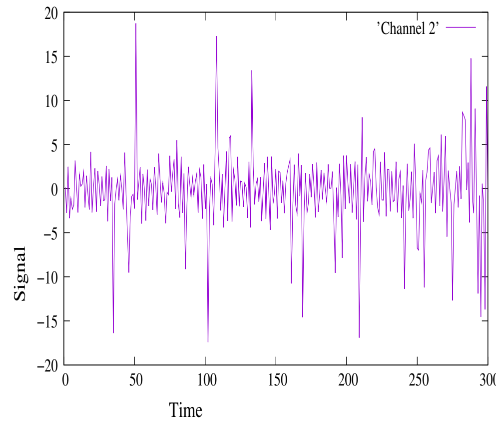


Figure 8: The observed signal with outliers in (a) Channel 1 and (b) Channel 2.

Weight	$\widehat{A_1^2} + \widehat{B_1^2}$	$\widehat{A_2^2} + \widehat{B_2^2}$	$\widehat{\lambda}$	WRSS
(i)	25.2582	10.3506	2.35493	4.6957
(ii)	17.8041	6.7440	2.35499	4.1751
(iii)	20.2789	8.4796	2.35513	5.5052
(iv)	19.6930	8.3425	2.35531	5.2787
(v)	17.0129	6.2107	2.35496	1.3947
(vi)	15.3789	5.8125	2.35510	3.2917

Table 1: GWLS estimators, normalized residual sums of squares and squared error distance of the fitted and the originals based on different weight functions

6 Conclusions

In this article we have considered a two-channel sinusoidal model in presence of outliers when the errors at the different channels are dependent. We have proposed a very efficient weighted generalized least squares estimators which work better than the generalized least squares estimators, LAD estimators and Huber's M-estimators. The large sample properties have been developed under a fairly general set up. It may be mentioned that although we have mainly considered one-component model at both the channel, the results can be established for multi-component sinusoidal model also.

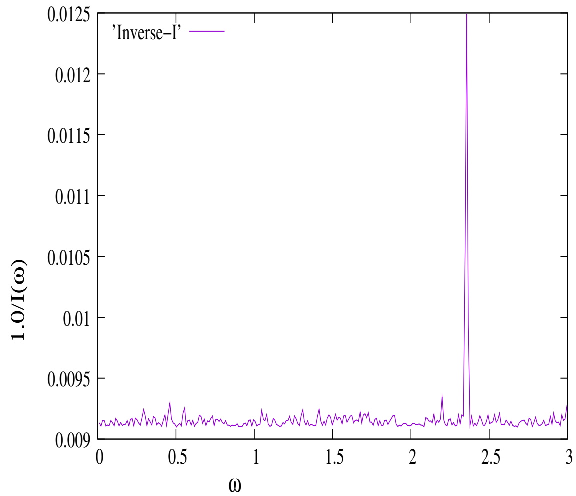


Figure 9: Inverse of the Periodogram Function.

Appendix A

To prove Theorem 1, we need the following lemma.

Lemma 1: Let $\{u_k(n)\}$ be a stationary sequence which satisfies Assumption 1, then for $m = 0, 1, 2, \ldots$.

$$\lim_{N\to\infty} \sup_{\theta} \left| \frac{1}{N^{m+1}} \sum_{n=1}^{N} n^m u_k(n) \cos(n\theta) \right| = 0 \qquad a.s.,$$

for $k = 1$ and 2. The same result holds when $\cos(n\theta)$ is replaced by $\sin(n\theta)$.

Proof For m = 0,1,2, see Lemma 1 and its associated corollaries of Kundu [7]. For general k, the results can be obtained along the same line. ∎

Lemma 2: Let $\{X(t)\}$ be a stationary sequence which satisfies Assumption 1, and $w(t)$ is of the form (11) and it satisfies (12), then

$$\lim_{N\to\infty} \sup_{\theta} \left| \frac{1}{N} \sum_{n=1}^{N} w\left(\frac{n}{N}\right) u_k(n) \cos(t\theta) \right| = 0 \qquad a.s.$$

The same result holds when $\cos(n\theta)$ is replaced by $\sin(n\theta)$.

Proof: Using Lemma 1, it can be obtained. ∎

Lemma 3: Under the same set of assumptions as in Theorem 1, for any given $\delta > 0$

$$\liminf_{N\to\infty} \inf_{\boldsymbol{\xi}\in S_\delta} \frac{1}{N}\{R(\boldsymbol{\xi}) - R(\boldsymbol{\xi}^0)\} > 0, \tag{34}$$

here $R(\boldsymbol{\xi})$ is same as defined in (13) and $S_\delta = \{\boldsymbol{\xi} : \left|\boldsymbol{\xi} - \boldsymbol{\xi}^0\right| > 5\delta\}$

Proof: Let us write

$$\frac{1}{N}\{R(\boldsymbol{\xi}) - R(\boldsymbol{\xi}^0)\} = f(\boldsymbol{\xi}) + g(\boldsymbol{\xi}), where$$

$$\begin{aligned}
f(\boldsymbol{\xi}) &= \frac{1}{N}\sum_{n=1}^{N} w\left(\frac{n}{N}\right)\left(\boldsymbol{\mu}(n;\boldsymbol{\beta}^0,\omega^0) - \boldsymbol{\mu}(n;\boldsymbol{\beta},\omega)\right)^\top \boldsymbol{\Sigma}^{-1}\left(\boldsymbol{\mu}(n;\boldsymbol{\beta}^0,\omega^0) - \boldsymbol{\mu}(n;\boldsymbol{\beta},\omega)\right)\\
g(\boldsymbol{\xi}) &= \frac{2}{N}\sum_{n=1}^{N} w\left(\frac{n}{N}\right)\left(\boldsymbol{\mu}(n;\boldsymbol{\beta}^0,\omega^0) - \boldsymbol{\mu}(n;\boldsymbol{\beta},\omega)\right)^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{e}(n).
\end{aligned}$$

Using Lemma 2, it can be shown that $\lim_{N\to\infty} \sup_{\boldsymbol{\xi}\in S_\delta} |g(\boldsymbol{\xi})| = 0, a.s.$. For $k = 1$ and 2, let us consider the sets

$$S_{\delta,A_k} = \{\boldsymbol{\xi} : |A_k - A_k^0| > \delta\}, S_{\delta,B_k} = \{\boldsymbol{\xi} : |B_k - B_k^0| > \delta\}, S_{\delta,\omega} = \{\boldsymbol{\xi} : |\omega - \omega^0| > \delta\}.$$

We have

$$S_\delta \subset S_{\delta,A_1} \cup S_{\delta,A_2} \cup S_{\delta,B_1} \cup S_{\delta,B_2} \cup S_{\delta,\omega}. \tag{35}$$

If we denote $\boldsymbol{\Sigma}^{-1} = ((\sigma^{ij}))$, then after some calculations it can be shown that

$$\liminf_{N\to\infty} \inf_{\boldsymbol{\xi}\in S_{\delta,A_1}} f(\boldsymbol{\xi}) = \liminf_{N\to\infty} \inf_{|A_1 - A_1^0|>\delta} f(\boldsymbol{\xi}) = \inf_{|A_1 - A_1^0|>\delta} \frac{\sigma^{11}}{2}|A_1 - A_1^0|^2 \int_0^1 w(t)dt > 0.$$

Similar results hold if we replace S_{δ,A_1} by S_{δ,A_2}, S_{δ,B_1}, S_{δ,B_2} and $S_{\delta,\omega}$ also. Hence, using (35), (34) follows. ∎

Proof of Theorem 1: Now using Lemma 1 of Wu [24] the result immediately follows. ∎

Appendix B

Lemma 4: If $\theta \in (0, \pi)$, then for $k = 0, 1, \ldots$

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

Proof: See Mangulis [12]. ∎

Using Lemma 4, we immediately obtain the following Lemma.

Lemma 5: If $w(t) = c_0 + c_1 t + \ldots + c_k t^k$, then

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} w\left(\frac{n}{N}\right) \cos(\theta n) = \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} w\left(\frac{n}{N}\right) \sin(\theta n) = 0,$$
$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} w\left(\frac{n}{N}\right) \cos^2(\theta n) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} w\left(\frac{n}{N}\right) \sin^2(\theta n) = \frac{1}{2} \int_0^1 w(t) dt,$$
$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} w\left(\frac{n}{N}\right) \cos(\theta n) \sin(\theta n) = 0.$$

Proof of Theorem 2:

Let us denote

$$R'(\boldsymbol{\xi}) = \left(\frac{\partial R(\boldsymbol{\xi})}{\partial A_1}, \frac{\partial R(\boldsymbol{\xi})}{\partial B_1}, \frac{\partial R(\boldsymbol{\xi})}{\partial A_2}, \frac{\partial R(\boldsymbol{\xi})}{\partial B_2}, \frac{\partial R(\boldsymbol{\xi})}{\partial \omega} \right)^\top$$

and $R''(\boldsymbol{\xi})$ as the corresponding double derivative 5×5 matrix. Now expanding $R'(\boldsymbol{\xi})$ around $\boldsymbol{\xi}^0$ using multivariate Taylor series expansion gives;

$$R'(\widehat{\boldsymbol{\xi}}) - R'(\boldsymbol{\xi}^0) = R''(\bar{\boldsymbol{\xi}})(\widehat{\boldsymbol{\xi}} - \boldsymbol{\xi}^0), \tag{36}$$

where $\overline{\boldsymbol{\xi}}$ is point on the line joining $\widehat{\boldsymbol{\xi}}$ and $\boldsymbol{\xi}^0$. Since $R'(\widehat{\boldsymbol{\xi}}) = \mathbf{0}$, (36) can be written as

$$\boldsymbol{D}(\widehat{\boldsymbol{\xi}} - \boldsymbol{\xi}^0) = -\left[\boldsymbol{D}^{-1}R''(\overline{\boldsymbol{\xi}})\boldsymbol{D}^{-1}\right]^{-1}\left[\boldsymbol{D}^{-1}R'(\boldsymbol{\xi}^0)\right]. \tag{37}$$

The elements of $R''(\overline{\boldsymbol{\xi}})$ are all continuous functions and using the consistency results of $\widehat{\boldsymbol{\xi}}$, it follows

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

Now by using Lemma 5 repeatedly to the elements of the matrix $\boldsymbol{D}R''(\boldsymbol{\xi}^0)\boldsymbol{D}$, it can be shown that

$$\lim_{N\to\infty}\left[\boldsymbol{D}^{-1}R''(\boldsymbol{\xi}^0)\boldsymbol{D}^{-1}\right] = \boldsymbol{\Gamma}.$$

Now using the Central limit theorem of the Linear processes, see for example Fuller [3] and using Lemma 5, it can be shown that

$$\boldsymbol{D}^{-1}R'(\boldsymbol{\xi}^0) \xrightarrow{d} N_5(\mathbf{0}, 2\delta\boldsymbol{\Psi}),$$

hence, the result follows. ∎

Appendix C

In this appendix we have presented some more simulation results with the same model parameters but five different weight functions mainly to show the effect of the weight functions on the parameter estimation. We have considered the same model parameters as in (29) and error distribution as in Case 1 of Section 3. We have considered the following five different weight functions namely;

$$w_1(t) = 1 + \left(t - \frac{1}{2}\right)^2, \qquad w_2(t) = \frac{1}{12} + 11\left(t - \frac{1}{2}\right)^2 \tag{38}$$

$$w_3(t) = \begin{cases} (2-4t) & \text{if} \quad 0 \le t \le \frac{1}{4} \\ 1 & \text{if} \quad \frac{1}{4} \le t \le \frac{3}{4} \\ (4t-2) & \text{if} \quad \frac{3}{4} \le t \le 1, \end{cases} \tag{39}$$

$$w_4(t) = \frac{4}{3} - 4\left(t - \frac{1}{2}\right)^2, \qquad w_5(t) = 1.0 + 4\left(t - \frac{3}{4}\right)^2. \tag{40}$$

Method ↓ N →	200	400	600	800	1000
GLS	2.3566	2.3561	2.3561	2.3561	2.3562
	(107.8754)	(150.5217)	(168.5121)	(183.1364)	(190.1352)
WGLS-1	2.3567	2.3561	2.3561	2.3561	2.3561
	(107.7767)	(151.4286)	(176.3229)	(184.7373)	(191.6705)
WGLS-2	2.3562	2.3561	2.3561	2.3562	2.3562
	(147.8766)	(173.3804)	(185.5166)	(194.2513)	(201.0490)
WGLS-3	2.3562	2.3561	2.3561	2.3561	2.3562
	(129.4078)	(174.1368)	(186.1527)	(195.0018)	(201.7123)
WGLS-4	2.3570	2.3562	2.3561	2.3561	2.3561
	(102.5301)	(141.3930)	(169.2814)	(177.8650)	(184.8509)
WGLS-5	2.3569	2.3561	2.3561	2.3561	2.3561
	(104.8610)	(146.4052)	(167.8667)	(181.7285)	(188.5213)
LAD	2.3552	2.3551	2.3557	2.3558	2.3559
	(132.3792)	(143.6538)	(150.0876)	(155.3606)	(159.3407)
Huber-M	2.3547	2.3554	2.3556	2.3558	2.3558
	(127.0081)	(140.8901)	(148.8538)	(154.5606)	(158.8891)

Table 2: The average estimates of the frequencies and the associated -10 log (MSE) of the different methods are presented. The minimum MSE for each sample size has been indicated by red.

These weight functions put different weights at different part of the data sequences. We have computed the average estimates of the frequency, and the amplitudes and the corresponding mean squared errors (MSE). We have also computed the same for the GLS estimators, LAD and Huber-M estimators. We have reported the average estimates and the associated -10 log (MSE) (in brackets below) in each case. The results are reported in Tables 2, 3 and 4. Some of the points are quite clear from these simulation experiments. It is clear that the performance of the estimators depend very much of the weight functions as expected. Some of the weight functions perform better than the GLS estimators, LAD and Huber-M estimators. The effect is more on the amplitude estimation than frequency estimation. Hence, the choice of the weight function is an important problem, and it has been discussed in details in Section 4.

Method ↓ N →	200	400	600	800	1000
GLS	18.6191	18.2523	18.2101	18.1873	18.1205
	(-31.9062)	(-24.0497)	(-20.1955)	(-17.0724)	(-14.8960)
WGLS-1	18.4663	18.3157	18.2132	18.1482	18.0931
	(-30.5802)	(-22.7243)	(-18.7609)	(-15.8265)	(-13.5708)
WGLS-2	18.0698	18.0367	18.0231	18.0180	18.0172
	(-6.8317)	(0.3678)	(4.2503)	(7.2169)	(9.6464)
WGLS-3	18.0577	18.0372	18.0235	18.0199	18.0140
	(-5.0970)	(2.4668)	(6.5804)	(9.5966)	(11.4505)
WGLS-4	19.0596	18.5669	18.3521	18.3176	18.1840
	(-37.3611)	(-29.5655)	(-25.5180)	(-22.4119)	(-20.0876)
WGLS-5	18.1584	18.0767	18.0726	18.0576	18.0331
	(-28.1015)	(-20.0969)	(-16.1703)	(-13.4199)	(-10.9940)
LAD	17.9729	17.6732	17.5407	17.4979	17.5293
	(-13.2471)	(-10.0088)	(-8.5431)	(-8.5067)	(-9.0271)
Huber-M	17.1569	17.2612	17.5733	17.5742	17.6247
	(-19.4322)	(-16.6943)	(-16.1348)	(-14.7992)	(-13.0578)

Table 3: The average estimates of the first amplitude and the associated -10 log (MSE) of the different methods are presented. The minimum MSE for each sample size has been indicated by red.

Method ↓ N →	200	400	600	800	1000
GLS	8.6977	8.2960	8.2059	8.1633	8.1317
	(-24.1064)	(-16.2127)	(-12.1213)	(-9.1092)	(-6.9528)
WGLS-1	8.5453	8.2877	8.1951	8.1392	8.0917
	(-22.3859)	(-14.7801)	(-10.6991)	(-7.7348)	(-5.6643)
WGLS-2	8.0655	8.0367	8.0208	8.0193	8.0159
	(1.1455)	(8.5350)	(12.4045)	(15.4045)	(17.7722)
WGLS-3	8.0637	8.0335	8.0239	8.0196	8.0155
	(3.4427)	(10.4362)	(14.6670)	(17.6114)	(19.4527)
WGLS-4	9.1260	8.5647	8.3505	8.3064	8.2008
	(-29.6829)	(-21.8746)	(-17.6258)	(-14.5137)	(-12.2373)
WGLS-5	8.3354	8.1611	8.1199	8.0792	8.0573
	(-19.5132)	(-12.0010)	(-8.3056)	(-5.3805)	(-2.8661)
LAD	7.5416	7.5421	7.6862	7.7560	7.8063
	(-10.7259)	(-6.5125)	(-3.3979)	(-3.2903)	(-2.1861)
Huber-M	7.5156	7.6283	7.8729	7.8839	7.8360
	(-13.6170)	(-9.6339)	(-8.3757)	(-6.8820)	(-5.0903)

Table 4: The average estimates of the second amplitude and the associated -10 log (MSE) of the different methods are presented. The minimum MSE for each sample size has been indicated by red.

Acknowledgements:

The authors would like to thank the Associate Editor and two unknown reviewers for their constructive suggestions which have helped to improve the paper significantly.

Data Availability Statement

The manuscript has no associated data.

Conflict of Interest Statement:

The author does not have any conflict of interest in preparation of this manuscript.

References

[1] 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.

[2] Clercq, W.D., Vanrumste, B., Papy, J.-M., Paesschen, W.V. and Van Huffel, S. (2005), "Modeling common dynamics in multichannel signals with applications to artifact and background removal EEG recordings", IEEE Transactions on Biomedical Engineering, vol. 52, 2006 – 2015.

[3] Fuller, W.A. (1976), *Introduction to Statistical Time Series*, John Wiley and Sons, New York.

[4] 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.

[5] Handel, P. (2008), "Parameter estimation employing a dual-channel sine-wave model under a Gaussian assumption", *IEEE Transactions on Instrumentation and Measurement*, vol. 57, 1661 – 1669.

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

[7] Kundu, D. (1997), "Asymptotic Properties of the Least Squares Estimators of Sinusoidal Signals", *Statistics*, Vol. 30, 221–238.

[8] Kundu, D. (2023), "Weighted least squares: A robust method of estimation for sinusoidal model", *Communications in Statistics - Theory and Methods*, (to appear).

[9] Kundu, D. and Grover, R. (2023), "Robust estimators of two dimensional sinusoidal model parameters", *Signal Processing*, vol. 212, November, 109167.

[10] Kundu, D., Nandi, S. and Grover, R. (2023), "On Weighted Least Squares Estimators of Parameters of a Chirp Model", *Circuits, Systems and Signal Processing*, Vol. 42, No. 1, 493 – 521.

[11] Kundu, D., Nandi, S. and Grover, R. (2023), "On weighted least squares estimators for chirp like model", *Sankhya A*, (to appear).

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

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

[14] Papy, J.-M., Lathauwer, L. De and Van Huffel, S. (2006), "Common pole estimation in multi-channel exponential data modeling", *Signal Processing*, vol. 86, 846 – 858.

[15] Ramos, P.M., da Silva, M.F. and Serra, A.C. (2004), “Low frequency impedance measurement using sine-fitting”, *Measurement*, vol. 35, 89 – 96.

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

[17] Sandgren, N., Stoica, P., Frigo, F.J., Selen, Y. (2005), “Spectral analysis of multichannel MRS data”, *Journal of Magnetic Resonance*, vol. 175, 79 – 91.

[18] So, H.C. and Zhou, Z. (2012), “Linear prediction approach for accurate dual channel sine-wave parameter estimation”, *ETRI Journal*, vol. 34, 641 - 644.

[19] So, H.C., Zhou, Z. (2013), ”Two accurate phase-difference estimators for dual-channel sine-wave model”, *EURASIP Journal of Advanced Signal Processing*, vol. 122, https://doi.org/10.1186/1687-6180-2013-122

[20] Strutz, T. (2016), *Data Fitting and Uncertainty (A practical introduction to weighted least squares and beyond)*, Springer Vieweg. ISBN 978-3-658-11455-8.

[21] Stanković, L., Mandić, D., Daković, M., Brajović, M. (2018), “Time-frequency decomposition of multivariate multicomponent signals”, *Signal Processing*, vol. 142, 468 – 479.

[22] Tzagkarakis, C., Mouchtaris, A. and Tsakalides, P., (2010), “A Multichannel Sinusoidal Model Applied to Spot Microphone Signals for Immersive Audio”, *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 17, no. 8, pp. 1483-1497.

[23] Vucijak, N.M. and Saranovac, L.V. (2010), “A simple algorithm for the estimation of phase difference between two sinusoidal voltages”, *IEEE Transactions on Instrumentation and Measurement*, vol. 59, 3152 – 3158.

[24] Wu, C.F.J. (1981), "Asymptotic theory of the nonlinear least squares estimation", *Annals of Statistics*, vol. 9, 501 – 513.

[25] 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, 3895 – 3907.