On Weighted Least Squares Estimators for Chirp Like Model

Debasis Kundu*& Swagata Nandi†& Rhythm Grover ‡

Abstract

In this paper we have considered the chirp like model which has been recently introduced, and it has a very close resemblance with a chirp model. We consider the weighted least squares estimators of the parameters of a chirp like model in presence of an additive stationary error, and study their properties. It is observed that although the least squares method seems to be a natural choice to estimate the unknown parameters of a chirp like model, the least squares estimators are very sensitive to the outliers. It is observed that the weighted least squares estimators are quite robust in this respect. The weighted least squares estimators are consistent and they have the same rate of convergence as the least squares estimators. We have further extended the results in case of multicomponent chirp like model. Some simulations have been performed to show the effectiveness of the proposed method. In simulation studies, weighted least squares estimators have been compared with the least absolute deviation estimators which, in general, are known to work well in presence of outliers. One EEG data set has been analyzed and the results are quite satisfactory.

KEYWORDS: Chirp like model; chirp signal model; non-linear least squares; weighted least squares; asymptotic distribution; strong consistency; outliers.

AMS SUBJECT CLASSIFICATIONS: 62F10, 62F03, 62H12.

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

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

‡Theoretical Statistics and Mathematics Unit, Indian Statistical Institute, 7, S.J.S. Sansanwal Marg, New Delhi - 110016, India, rhythm.grover91@gmail.com

1 Introduction

The chirp signal has received a considerable amount of attention in the statistical signal processing literature due to its wide applications in various fields. A multicomponent chirp model can be described as follows:

$$y(n) = \sum_{k=1}^{p} \{A_k^0 \cos(\alpha_k^0 n + \beta_k^0 n^2) + B_k^0 \sin(\alpha_k^0 n + \beta_k^0 n^2)\} + X(n). \tag{1}$$

The multicomponent chirp model (1) can be seen as an uniform linear frequency modulated (FM) sinusoidal model. The history of chirp model goes back to 1950, see for example the article by Lancaster [10]. According to Lancaster [10], *Pulse compression (aka 'chirp') radar was invented in the 1950s by Sperry and a couple of other defense contractors.* Klauder et al. [6] used single component chirp model (1) in designing chirp radars that provides a solution for the conflicting requirements of simultaneous long-range and high-resolution performance in radar systems. Since then it has become extremely popular in radar technology. Chirp model has been used quite successfully to analyze sonar data also. The chirp sonar data has been used in estimating the physical and acoustic properties of the seabed quite frequently, see for example Schock [14, 15] and the references cited therein. Chirp model has been used quite extensively in optics also. It has been observed by Guillet et al. [4] that the broadband radio-frequency waveforms can be analyzed quite effectively using chirp models.

Recently, Grover [2], see also Kundu and Grover [7] and Grover, Kundu and Mitra [3], introduced a chirp like model which behaves like a chirp model and it is observed based on extensive data analyses that it is not possible to discriminate a chirp and a chirp like model in many instances. Mathematically, the multicomponent chirp like (MCCL) model in presence of additive noise can be written as follows:

$$y(n) = \sum_{j=1}^{p} \{A_j^0 \cos(\alpha_j^0 n) + B_j^0 \sin(\alpha_j^0 n)\} + \sum_{k=1}^{q} \{C_k^0 \cos(\beta_k^0 n^2) + D_k^0 \sin(\beta_k^0 n^2)\} + X(n). \tag{2}$$

Here $A_j^0, B_j^0, C_k^0, D_k^0$ are linear parameters, α_j^0 and β_k^0 are the frequency and frequency rate, respectively. The sequence of random variables, $\{X(t)\}$ denotes the noise present in the model. It has mean zero and finite variance. The explicit assumptions on $\{X(t)\}$ will be provided later. Based on extensive data analyses by Grover, Kundu and Mitra [3] it has been observed that MCCL model often provides a better fit to a nearly periodic data set than a multicomponent chirp model, and it needs less number of parameters. Moreover, the implementation of the MCCL model is much easier than a multicomponent chirp model. Similar to the chirp model, MCCL also can be used for future prediction also.

Note that the MCCL model (2) is also a non-linear regression model. Hence, the least squares (LS) method seems to be a natural choice in estimating the unknown parameters. It has been shown by Grover [2] that the MCCL does not satisfy the standard sufficient conditions of Jennrich [5] or Wu [17] so that the least squares estimators (LSEs) become consistent. It has been shown by Grover, Kundu and Mitra [3] that under a fairly general set of assumptions the LSEs are consistent and asymptotically normally distributed. Although, the LSEs have these desirable properties, it has been observed that they are quite sensitive to the presence of outliers. Even if only a few outliers are present in the data, it can affect the performance of the LSEs quite significantly. One natural choice in this case is to choose some robust estimators like L_1 norm estimators or M-estimators. But implementing any robust estimators or establishing the properties of these estimators in a general set up are quite challenging. Due to this reason we have explored in this paper the weighted least squares estimators (WLSEs) to estimate the unknown parameters of the MCCL model. It is observed that the WLSEs are quite robust compared to the LSEs and they are easy to implement even in a general set up. Based on an extensive numerical experiments, it has been observed that the performance of the WLSEs are comparable with the robust least absolute deviation estimators (LADEs). But developing theoretical properties of the LADEs and implementing them in practice are quite difficult specially for multicomponent models. Therefore, we propose to use the WLSEs in presence of outliers in this case. We have established the consistency and asymptotic normality properties of the WLSEs and it is observed that the LSEs and WLSEs have the same rate of convergence under a fairly general set of error assumptions. We have further proposed sequential WLSEs, which can be implemented quite conveniently and they have the same asymptotic properties as the WLSEs.

The rest of the paper is organized as follows. In Section 2 and Section 3 we consider the one component chirp like (OCCL) and MCCL models, respectively. Simulation results have been presented in Section 4 and the data analyses have been presented in Section 5. The conclusions have appeared in Section 6. Preliminary results required for the proofs have been presented in Appendix A and all the proofs in Appendices B, C and D.

2 One Component Chirp Like Model

2.1 WLSEs

The OCCL model can be defined as follows:

$$y(n) = A^0 \cos(\alpha^0 n) + B^0 \sin(\alpha^0 n) + C^0 \cos(\beta^0 n^2) + D^0 \sin(\beta^0 n^2) + X(n). \tag{3}$$

Here, A^0, B^0, C^0, D^0 are linear parameters, and α^0, β^0 are frequency and frequency rate, respectively. We make the following assumptions on $\{X(n)\}$.

$$X(n) = \sum_{k=-\infty}^{\infty} a(k)e(n-k); \qquad n = 1, 2, \ldots. \tag{4}$$

Here $e(n)$'s are i.i.d random variables with mean zero, and finite fourth moment. Moreover, $a(k)$'s are such that

$$\sum_{k=-\infty}^{\infty} |a(k)| < \infty. \tag{5}$$

It is assumed that the weight function $w(t)$ satisfies the following assumption:

ASSUMPTION 1: Suppose $w(t)$ is a non-negative continuous function defined on [0,1], such that $\min\limits_{0\le t\le 1} w(t) > \gamma > 0$ and $\max\limits_{0\le t\le 1} w(t) \le K < \infty$.

The WLSEs of the unknown parameter $\boldsymbol{\theta}^0 = (A^0, B^0, C^0, D^0, \alpha^0, \beta^0)^\top$ can be obtained by minimizing

$$Q(\boldsymbol{\theta}) = \sum_{n=1}^{N} w\left(\frac{n}{N}\right) \left(y(n) - A\cos(\alpha n) - B\sin(\alpha n) - C\cos(\beta n^2) - D\sin(\beta n^2)\right)^2 \qquad (6)$$

with respect to the unknown parameter vector $\boldsymbol{\theta} = (A, B, C, D, \alpha, \beta)^\top$. Let us denote the WLSE of $\boldsymbol{\theta}^0$ as $\widehat{\boldsymbol{\theta}} = (\widehat{A}, \widehat{B}, \widehat{C}, \widehat{D}, \widehat{\alpha}, \widehat{\beta})^\top$. The WLSEs cannot be obtained in analytical form as minimization of $Q(\boldsymbol{\theta})$ is a non-linear optimization problem. One needs to use some numerical techniques like Newton-Raphson or Gauss-Newton method to compute the WLSEs. The OCCL model has six unknown parameters, we will show that the WLSEs of the unknown parameters can be obtained by solving a two-dimensional optimization problem. From (6) it is immediate that the LSEs can be obtained as a special case of the WLSEs by assuming $w(t) = 1$.

For a given α and β, the minimization of $Q(\boldsymbol{\theta})$ can be obtained as a simple weighted least squares approach. The WLSEs of A, B, C and D for a given α and β, say $\widehat{A}(\alpha,\beta)$, $\widehat{B}(\alpha,\beta)$,$\widehat{C}(\alpha,\beta)$ and $\widehat{D}(\alpha,\beta)$, can be obtained as

$$\boldsymbol{\delta}(\alpha,\beta) = \begin{pmatrix} \widehat{A}(\alpha,\beta) & \widehat{B}(\alpha,\beta) & \widehat{C}(\alpha,\beta) & \widehat{D}(\alpha,\beta) \end{pmatrix}^\top = \boldsymbol{U}_N^{-1}(\alpha,\beta)\boldsymbol{u}, \qquad (7)$$

where $\boldsymbol{U}_N(\alpha,\beta) = ((u_{ij}))$ is a 4×4 symmetric matrix, as follows:

$$u_{11} = \sum_{n=1}^N w\left(\tfrac{n}{N}\right)\cos^2(\alpha n) \qquad u_{12} = u_{21} = \sum_{n=1}^N w\left(\tfrac{n}{N}\right)\sin(\alpha n)\cos(\alpha n)$$

$$u_{13} = u_{31} = \sum_{n=1}^N w\left(\tfrac{n}{N}\right)\cos(\alpha n)\cos(\beta n^2) \qquad u_{14} = u_{41} = \sum_{n=1}^N w\left(\tfrac{n}{N}\right)\cos(\alpha n)\sin(\beta n^2)$$

$$u_{22} = \sum_{n=1}^N w\left(\tfrac{n}{N}\right)\sin^2(\alpha n) \qquad u_{23} = u_{32} = \sum_{n=1}^N w\left(\tfrac{n}{N}\right)\sin(\alpha n)\cos(\beta n^2)$$

$$u_{24} = u_{42} = \sum_{n=1}^N w\left(\tfrac{n}{N}\right)\sin(\alpha n)\sin(\beta n^2) \qquad u_{33} = \sum_{n=1}^N w\left(\tfrac{n}{N}\right)\cos^2(\beta n^2)$$

$$u_{34} = u_{43} = \sum_{n=1}^N w\left(\tfrac{n}{N}\right)\cos(\beta n^2)\sin(\beta n^2) \qquad u_{44} = \sum_{n=1}^N w\left(\tfrac{n}{N}\right)\sin^2(\beta n^2)$$

and $\boldsymbol{u} = (u_1, u_2, u_3, u_4)^\top$ is a 4×1 vector as given below

$$u_1 = \sum_{n=1}^N w\left(\tfrac{n}{N}\right) y(n)\cos(\alpha n) \qquad u_2 = \sum_{n=1}^N w\left(\tfrac{n}{N}\right) y(n)\sin(\alpha n)$$

$$u_3 = \sum_{n=1}^N w\left(\tfrac{n}{N}\right) y(n)\cos(\beta n^2) \qquad u_4 = \sum_{n=1}^N w\left(\tfrac{n}{N}\right) y(n)\sin(\beta n^2).$$

Therefore, the WLSEs of α and β can be obtained by minimizing $R(\alpha,\beta)$, where

$$R(\alpha,\beta) = Q(\widehat{A}(\alpha,\beta), \widehat{B}(\alpha,\beta), \widehat{C}(\alpha,\beta), \widehat{D}(\alpha,\beta), \alpha, \beta),$$

with respect to α and β. Hence, we have observed that although OCCL model has six parameters, the WLSEs of the nonlinear parameters α and β can be obtained by solving a two dimensional optimization problem. Once the WLSEs of α and β are obtained, then the WLSEs of the other parameters can be obtained in explicit forms. The following result provides the consistency properties of the WLSEs.

Theorem 1: If $\{X(n)\}$ has the structure as given in (4), and $w(t)$ satisfies Assumption 1, then the WLSE $\widehat{\boldsymbol{\theta}}$ is a strongly consistent estimator of $\boldsymbol{\theta}^0$.

Proof: See in Appendix B. ∎

Now we provide the asymptotic distribution of the WLSEs, and for that we need the following notations.

$$\lim_{N\to\infty} \frac{1}{N^{k+1}} \sum_{n=1}^N n^k w\left(\frac{n}{N}\right) \quad = \quad \int_0^1 t^k w(t)dt = c_{k+1} > 0, \tag{8}$$

$$\lim_{N\to\infty} \frac{1}{N^{k+1}} \sum_{n=1}^{N} n^k w^2\left(\frac{n}{N}\right) \quad = \quad \int_0^1 t^k w^2(t)dt = d_{k+1} > 0; \quad k = 0, 1, 2, \ldots. \tag{9}$$

Let us define the following four 3×3 matrices:

$$\boldsymbol{\Sigma}_1 = \begin{bmatrix} d_1 & 0 & B^0 d_2 \\ 0 & d_1 & -A^0 d_2 \\ B^0 d_2 & -A^0 d_2 & (A^{0^2} + B^{0^2}) d_3 \end{bmatrix}, \; \boldsymbol{\Sigma}_2 = \begin{bmatrix} d_1 & 0 & D^0 d_3 \\ 0 & d_1 & -C^0 d_3 \\ D^0 d_3 & -C^0 d_3 & (C^{0^2} + D^{0^2}) d_5 \end{bmatrix}, \tag{10}$$

$$\boldsymbol{G}_1 = 2\begin{bmatrix} c_1 & 0 & B^0 c_2 \\ 0 & c_1 & -A^0 c_2 \\ B^0 c_2 & -A^0 c_2 & (A^{0^2} + B^{0^2}) c_3 \end{bmatrix}, \; \boldsymbol{G}_2 = 2\begin{bmatrix} c_1 & 0 & D^0 c_3 \\ 0 & c_1 & -C^0 c_3 \\ D^0 c_3 & -C^0 c_3 & (C^{0^2} + D^{0^2}) c_5 \end{bmatrix}, \tag{11}$$

and two diagonal matrices of order 3

$$\boldsymbol{D}_1 = \mathrm{diag}\left(N^{1/2}, N^{1/2}, N^{3/2}\right) \quad \text{and} \quad \boldsymbol{D}_2 = \mathrm{diag}\left(N^{1/2}, N^{1/2}, N^{5/2}\right).$$

Theorem 2: Under the same assumptions as in Theorem 1, and if the matrices $\boldsymbol{\Sigma}_i$'s and $\boldsymbol{G}_i$'s as defined in (10) and (11), are of full rank, then

$$\left((\widehat{A} - A^0), (\widehat{B} - B^0), (\widehat{\alpha} - \alpha^0)\right) \boldsymbol{D}_1 \xrightarrow{d} N_3\left(\mathbf{0}, \sigma^2 \zeta \; \boldsymbol{G}_1^{-1} \boldsymbol{\Sigma}_1 \boldsymbol{G}_1^{-1}\right),$$

$$\left((\widehat{C} - C^0), (\widehat{D} - D^0), (\widehat{\beta} - \beta^0)\right) \boldsymbol{D}_2 \xrightarrow{d} N_3\left(\mathbf{0}, \sigma^2 \eta \; \boldsymbol{G}_2^{-1} \boldsymbol{\Sigma}_2 \boldsymbol{G}_2^{-1}\right).$$

Here $\xrightarrow{d}$ means convergence in distribution and using $i = \sqrt{-1}$

$$\zeta = \left[\sum_{k=-\infty}^{\infty} a(k)\cos(\alpha^0 k)\right]^2 + \left[\sum_{k=-\infty}^{\infty} a(k)\sin(\alpha^0 k)\right]^2 = \left|\sum_{k=-\infty}^{\infty} a(k) e^{i\alpha^0 k}\right|^2,$$

$$\eta = \left[\sum_{k=-\infty}^{\infty} a(k)\cos(3\beta^0 k^2)\right]^2 + \left[\sum_{k=-\infty}^{\infty} a(k)\sin(3\beta^0 k^2)\right]^2 = \left|\sum_{k=-\infty}^{\infty} a(k) e^{i3\beta^0 k^2}\right|^2.$$

Also $\left((\widehat{A} - A^0), (\widehat{B} - B^0), (\widehat{\alpha} - \alpha^0)\right) \boldsymbol{D}_1$ and $\left((\widehat{C} - C^0), (\widehat{D} - D^0), (\widehat{\beta} - \beta^0)\right) \boldsymbol{D}_2$ are independently distributed.

Proof: See in Appendix C. ∎

We note that the WLSE of the frequency α is asymptotically independent of the WLSE of the frequency rate β. In fact, the WLSEs of A and B, the linear parameters associated with α are asymptotically independent of the WLSEs of C and D, the linear parameters associated with β. The asymptotic distribution of the WLSEs of the sinusoidal parameters only depends of the true values of the sinusoidal parameters. This has also been observed in case of chirp component. The WLSEs of sinusoidal component parameters are asymptotically independent of the WLSEs of the chirp component parameters, therefore, sequential method works even in one component chirp like model and has been discussed in Section 2.2.

So far we have discussed about the consistency and asymptotic normality properties of the WLSEs. Now we will discuss about the numerical issues related to the WLSEs. We have already mentioned that the WLSEs can be obtained by solving a two-dimensional optimization problem. Let us look at the behavior of $\boldsymbol{U}_N(\alpha, \beta)$ for large N. Observe that using Results A.1 and A.2 (stated in Appendix A), we obtain that

$$\lim_{N \to \infty} \frac{1}{N} u_{ii} = \frac{c_1}{2}; \quad 1 \leq i \leq 4, \tag{12}$$

$$\lim_{N \to \infty} \frac{1}{N} u_{ij} = 0; \quad 1 \leq i \neq j \leq 4. \tag{13}$$

Hence, $\widehat{\boldsymbol{\delta}}(\alpha, \beta)$ for large N can be obtained approximately as

$$\widehat{\boldsymbol{\delta}}^{\top}(\alpha, \beta) = \begin{pmatrix} \widehat{A}(\alpha) & \widehat{B}(\alpha) & \widehat{C}(\beta) & \widehat{D}(\beta) \end{pmatrix} = \begin{pmatrix} \frac{2u_1}{Nc_1} & \frac{2u_2}{Nc_1} & \frac{2u_3}{Nc_1} & \frac{2u_4}{Nc_1} \end{pmatrix},$$

where u_1, u_2, u_3, u_4 are same as defined before. Once, the WLSEs of α and β are estimated, the linear parameters are obtained as $\widehat{\boldsymbol{\delta}}^{\top}(\widehat{\alpha}, \widehat{\beta})$. When $w(t) = 1$, $\widehat{A}(\widehat{\alpha})$ and $\widehat{B}(\widehat{\alpha})$ corresponds to the approximate LSEs of A and B and similarly, $\widehat{C}(\widehat{\beta})$ and $\widehat{D}(\widehat{\beta})$ are those of C and D, respectively.

2.2 Weighted Sequential Estimators

It has been observed that the WLSEs can be obtained by solving a two-dimensional optimization problem, and it needs a $N \times N^2$ order of searching for the initial guesses similar to the LSEs. Finding the initial guesses can take a significant amount of time if N is large. To reduce the amount of computation of searching the initial guesses, Grover, Kundu and Mitra [3] proposed sequential estimators (SE), which have the same asymptotic efficiency as the ordinary LSEs. It reduces the computational burden significantly. First of all it reduces the search of the initial guesses to the order of $N + N^2$, and also instead of solving one two-dimensional optimization problem, one needs to solve two one-dimensional optimization problems. Similar to the SEs, we can provide weighted SEs (WSEs) also. Estimate A, B and α first by minimizing

$$Q_1(A, B, \alpha) = \sum_{n=1}^{N} w\left(\frac{n}{N}\right)\left(y(n) - A\cos(\alpha n) - B\sin(\alpha n)\right)^2 \tag{14}$$

with respect to A, B and α. Let $\widetilde{A}$, $\widetilde{B}$ and $\widetilde{\alpha}$ minimize $Q_1(A, B, \alpha)$. Then consider the modified data after taking out the effect of the sinusoidal component as follows:

$$\widetilde{y}(n) = y(n) - \widetilde{A}\cos(\widetilde{\alpha} n) - \widetilde{B}\sin(\widetilde{\alpha} n); \qquad n = 1, \ldots, N. \tag{15}$$

Estimate C, D and β, by minimizing

$$Q_2(C, D, \beta) = \sum_{n=1}^{N} w\left(\frac{n}{N}\right)\left(\widetilde{y}(n) - C\cos(\beta n^2) - B\sin(\beta n^2)\right)^2. \tag{16}$$

Note that the minimization of $Q_1(A, B, \alpha)$ can be obtained by solving a one-dimensional optimization problem. For a given α, the values of A and B which minimize $Q_1(A, B, \alpha)$ can be obtained as

$$\begin{bmatrix} \widetilde{A}(\alpha) & \widetilde{B}(\alpha) \end{bmatrix}^\top = \boldsymbol{V}_N^{-1}(\alpha)\boldsymbol{v},$$

here $\boldsymbol{V}_N(\alpha) = ((v_{ij}))$ is a 2×2 matrix, $\boldsymbol{v} = (v_1, v_2)^\top$ is a 2×1 vector, as follows:

$$v_{11} = u_{11}, v_{12} = v_{21} = u_{12}, v_{22} = u_{22}, \quad v_1 = u_1, v_2 = u_2.$$

Hence, the WSE of α can be obtained by minimizing $Q_1(\widetilde{A}(\alpha), \widetilde{B}(\alpha), \alpha)$ with respect to α. Similarly, the minimization of $Q_2(C, D, \beta)$ also can be obtained by solving a one-dimensional optimization problem. It may be seen that for a given β, the values of C and D which minimize $Q_2(C, D, \beta)$ can be obtained as

$$\left[\begin{array}{cc} \widetilde{C}(\beta) & \widetilde{D}(\beta) \end{array} \right]^\top = \boldsymbol{W}_N^{-1}(\beta)\boldsymbol{w},$$

here $\boldsymbol{W}_N(\beta) = ((w_{ij}))$ is a 2×2 matrix, as follows:

$$w_{11} = u_{33}, w_{12} = w_{21} = u_{34}, w_{22} = u_{44},$$

and $\boldsymbol{w} = (w_1, w_2)^\top$ is a 2×1 vector, where

$$w_1 = \sum_{n=1}^{N} w\left(\frac{n}{N}\right) \widetilde{y}(n) \cos(\beta n^2) \quad \text{and} \quad w_2 = \sum_{n=1}^{N} w\left(\frac{n}{N}\right) \widetilde{y}(n) \sin(\beta n^2).$$

Therefore, the WSE of β can be obtained by minimizing $Q_2(\widetilde{C}(\beta), \widetilde{D}(\beta), \beta)$ with respect to β. It is immediate that the WSEs can be obtained by solving two one-dimensional optimization problems sequentially. Let $\widetilde{\boldsymbol{\theta}} = (\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D}, \widetilde{\alpha}, \widetilde{\beta})^\top$ denote the WSEs of $\boldsymbol{\theta}^0$. The following results provide the asymptotic properties of $\widetilde{\boldsymbol{\theta}}$.

THEOREM 3: Under the same assumption as in Theorem 1, the WSE $\widetilde{\boldsymbol{\theta}}$ is a strongly consistent estimator of $\boldsymbol{\theta}^0$.

PROOF: The proof follows using the Lemmas required to prove Theorem 1, and based on the similar approaches as in Grover, Kundu and Mitra [3]. ∎

THEOREM 4: Under the same assumption as in Theorem 2, the WSE $\widetilde{\boldsymbol{\theta}}$ has the same asymptotic distribution as the WLSE $\widehat{\boldsymbol{\theta}}$.

PROOF: The proof follows along the same line as the proof of Theorem 2, but in this case we need Lemma C-3, see Appendix C, to establish the result. Please see eqn. (30) and (31) of Prasad et al. [13], and it should be clear why it is needed. ∎

3 Multicomponent Chirp Like Model

3.1 WLSEs

In this section we discuss about the WLSEs of the MCCL model as defined in (2). Let us use the following notations: $\boldsymbol{\theta}_j = (A_j, B_j, \alpha_j)$, $\boldsymbol{\gamma}_i = (C_i, D_i, \beta_i)$,

$$\mu_1(n; \boldsymbol{\theta}_j) = A_j \cos(\alpha_j n) + B_j \sin(\alpha_j n) \quad \text{and} \quad \mu_2(n; \boldsymbol{\gamma}_i) = C_i \cos(\beta_i n^2) + D_i \sin(\beta_i n^2),$$

for $j = 1, \ldots, p$ and $i = 1, \ldots, q$. Therefore, the WLSEs of the MCCL model can be obtained by minimizing the weighted residual sum of squares defined as follows:

$$Q(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_p, \boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_q) = \sum_{n=1}^{N} w\left(\frac{n}{N}\right) \left(y(n) - \sum_{j=1}^{p} \mu_1(n; \boldsymbol{\theta}_j) - \sum_{i=1}^{q} \mu_2(n; \boldsymbol{\gamma}_i) \right)^2. \tag{17}$$

Let us denote $\widehat{\boldsymbol{\theta}}_j = (\widehat{A}_j, \widehat{B}_j, \widehat{\alpha}_j)$, $\widehat{\boldsymbol{\gamma}}_i = (\widehat{C}_i, \widehat{D}_i, \widehat{\beta}_i)$ as the WLSEs of $\boldsymbol{\theta}_j^0 = (A_j^0, B_j^0, \alpha_j^0)$, $\boldsymbol{\gamma}_i^0 = (C_i^0, D_i^0, \beta_i^0)$, respectively, for $j = 1, \ldots, p$ and $i = 1, \ldots, q$. Now we are in a position to state the consistency results of the WLSEs of the MCCL model.

Theorem 5: If $\{X(n)\}$ has the same structure as given in (4), and $w(t)$ satisfies Assumption 1, then the WLSEs $\widehat{\boldsymbol{\theta}}_j = (\widehat{A}_j, \widehat{B}_j, \widehat{\alpha}_j)$, $\widehat{\boldsymbol{\gamma}}_i = (\widehat{C}_i, \widehat{D}_i, \widehat{\beta}_i)$ are consistent estimators of $\boldsymbol{\theta}_j^0 = (A_j^0, B_j^0, \alpha_j^0)$, $\boldsymbol{\gamma}_i^0 = (C_i^0, D_i^0, \beta_i^0)$, respectively, for $j = 1, \ldots, p$ and $i = 1, \ldots, q$.

Proof: The proof can be obtained along the same line as the proof of Theorem 1. ∎

Now we would like to state the asymptotic distribution of the WLSEs. For that purpose, we need to introduce the following notations. The 3×3 matrices $\boldsymbol{\Sigma}_{1j}$ and $\boldsymbol{G}_{1j}$ are obtained from the matrices $\boldsymbol{\Sigma}_1$ and $\boldsymbol{G}_1$, respectively, by replacing A^0, B^0 with A_j^0 and B_j^0, respectively, for $j = 1, \ldots, p$. Similarly, $\boldsymbol{\Sigma}_{2k}$ and $\boldsymbol{G}_{2k}$ are obtained from the matrices $\boldsymbol{\Sigma}_2$ and $\boldsymbol{G}_2$, respectively, by replacing C^0, D^0 with C_k^0 and D_k^0, respectively, for $k = 1, \ldots, q$. The matrices $\boldsymbol{D}_1$ and $\boldsymbol{D}_2$ are same as defined before. Then we have the following result.

Theorem 6: Under the same assumptions as in Theorem 1, and if all the matrices defined above are of full rank, then for $j = 1, \ldots, p$ and $k = 1, \ldots, q$,

$$\left((\widehat{A}_j - A_j^0), (\widehat{B}_j - B_j^0), (\widehat{\alpha}_j - \alpha_j^0)\right) \boldsymbol{D}_1 \xrightarrow{d} N_3\left(\mathbf{0}, \sigma^2 \zeta_j \; \boldsymbol{G}_{1j}^{-1} \boldsymbol{\Sigma}_{1j} \boldsymbol{G}_{1j}^{-1}\right),$$

$$\left((\widehat{C}_k - C_k^0), (\widehat{D}_k - D_k^0), (\widehat{\beta}_k - \beta_k^0)\right) \boldsymbol{D}_2 \xrightarrow{d} N_3\left(\mathbf{0}, \sigma^2 \eta_k \; \boldsymbol{G}_{2k}^{-1} \boldsymbol{\Sigma}_{2k} \boldsymbol{G}_{2k}^{-1}\right).$$

Here

$$\zeta_j = \left[\sum_{l=-\infty}^{\infty} a(l) \cos(\alpha_j^0 l)\right]^2 + \left[\sum_{l=-\infty}^{\infty} a(l) \sin(\alpha_j^0 l)\right]^2 = \left|\sum_{l=-\infty}^{\infty} a(l) e^{i\alpha_j^0 l}\right|^2,$$

$$\eta_k = \left[\sum_{l=-\infty}^{\infty} a(l) \cos(3\beta_k^0 l^2)\right]^2 + \left[\sum_{l=-\infty}^{\infty} a(l) \sin(3\beta_k^0 l^2)\right]^2 = \left|\sum_{l=-\infty}^{\infty} a(l) e^{i3\beta_k^0 l^2}\right|^2,$$

and they are all independently distributed.

Proof: The asymptotic distribution of the WLSE of the parameter correcponding to the j-th sinusoidal component $\left((\widehat{A}_j - A_j^0), (\widehat{B}_j - B_j^0), (\widehat{\alpha}_j - \alpha_j^0)\right) \boldsymbol{D}_1$ and correcponding to the k-th chirp component $\left((\widehat{C}_k - C_k^0), (\widehat{D}_k - D_k^0), (\widehat{\beta}_k - \beta_k^0)\right) \boldsymbol{D}_2$ follow exactly in the same way as the proof of Theorem 2 considering the estimator vector $\widehat{\boldsymbol{\theta}}_F = (\widehat{\boldsymbol{\theta}}_1, \ldots, \widehat{\boldsymbol{\theta}}_p, \widehat{\boldsymbol{\gamma}}_1, \ldots, \widehat{\boldsymbol{\gamma}}_q)$ and $Q(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_p, \boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_q)$ instead of $Q(\boldsymbol{\theta})$ defined in (6). The independence of $\widehat{\boldsymbol{\theta}}_i$ and $\widehat{\boldsymbol{\theta}}_j$, $i \neq j$ can be proved using Lemma C-1, see Appendix C, and that of $\widehat{\boldsymbol{\gamma}}_k$ and $\widehat{\boldsymbol{\gamma}}_l$, $k \neq l$ can be proved using Lemma C-2, see Appendix C. To prove independence between $\widehat{\boldsymbol{\theta}}_j$ and $\widehat{\boldsymbol{\gamma}}_k$ we also need to use Lemma C-2. ∎

3.2 WSE

It has been observed that the WLSEs of the unknown parameters of a MCCL model can be obtained by solving a $(p+q)$ dimensional optimization problem. In most of the practical situations the problem can be a quite numerically challenging problem. Due to this reason we propose the WSEs of the unknown parameters, which can be obtained by solving $(p+q)$

separate one-dimensional optimization problems and they have the same asymptotic properties as the WLSEs. The idea is same as the WSEs of the OCCL model. First obtain $\widetilde{\boldsymbol{\theta}}_1$, the WSE of $\boldsymbol{\theta}_1$, by the argument minimum of (18), where

$$S_1(\boldsymbol{\theta}_1) = \sum_{n=1}^{N} w\left(\frac{n}{N}\right)\left(y(n) - \mu_1(n;\boldsymbol{\theta}_1)\right)^2. \tag{18}$$

Then obtain $\widetilde{\boldsymbol{\gamma}}_1$, the WSE of $\boldsymbol{\gamma}_1$, by the argument minimum of (19), where

$$S_2(\boldsymbol{\gamma}_1) = \sum_{n=1}^{N} w\left(\frac{n}{N}\right)\left(y_1(n) - \mu_2(n;\boldsymbol{\gamma}_1)\right)^2, \tag{19}$$

where $y_1(n) = y(n) - \mu_1(n;\widetilde{\boldsymbol{\theta}}_1)$, for $n = 1,\ldots,N$. By repeating this procedure $(p+q)$ times, we can obtain $\widetilde{\boldsymbol{\theta}}_1,\ldots,\widetilde{\boldsymbol{\theta}}_p$ and $\widetilde{\boldsymbol{\gamma}}_1,\ldots,\widetilde{\boldsymbol{\gamma}}_q$. In each step, whether a sinusoidal component or a chirplet component is estimated, depends on the powers $\widetilde{A}_j^2 + \widetilde{B}_j^2$ and $\widetilde{C}_k^2 + \widetilde{D}_k^2$. By following similar procedure as in Grover, Kundu and Mitra [3], it can be shown that the WSEs have the same asymptotic properties as the WLSEs. Lemma C-3, provided in Appendix C, can be required to derive the asymptotic properties of WSEs.

4 Simulations

In this section, we report simulation results in the form of mean square errors (MSEs) and mean absolute deviations (MADs) for the WLSEs. These results are obtained with 5000 simulation runs for each data set. For the first set of experiments, we consider the following model with one sinusoidal and one chirp component:

$$y(n) = 10\cos(1.5n) + 10\sin(1.5n) + 10\cos(0.1n^2) + 10\sin(0.1n^2) + X(n).$$

Here, $X(n)$s are generated from a white Gaussian noise process with mean 0 and variance σ^2. For different error variances and sample sizes, we generate data from this model. These data sets are then contaminated by adding outliers to middle five percent of observations.

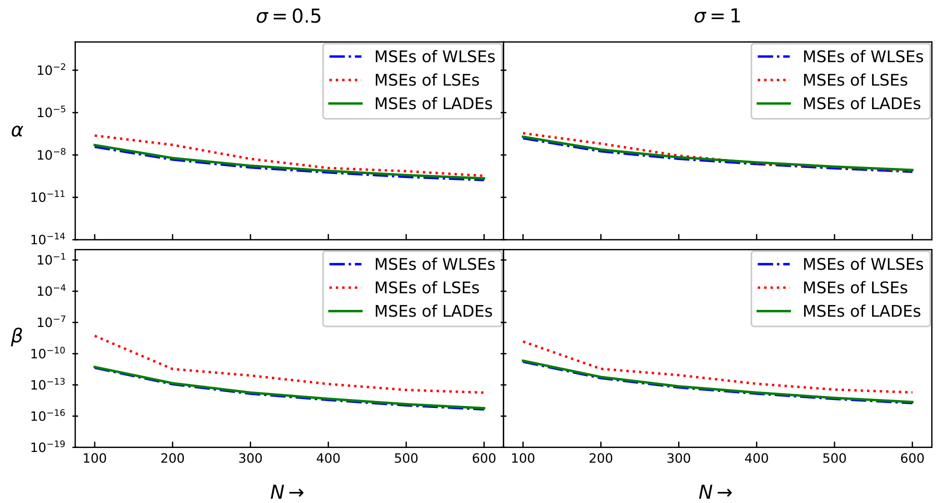

Figure 1: In each sub-plot, the graphed lines represent the MSEs of the WLSEs, the LSEs and the LADEs of the simulated one component model.

We then compute the WLSEs of the frequency and chirp rate parameters using the following weight function:

$$w\left(\frac{n}{N}\right) = \frac{1}{4} - \frac{n}{N} + \frac{n^2}{N^2}.$$

The weight function plays an important role in the computation of the WLSEs and therefore its choice is a crucial step. Since in the simulated data sets, the outliers are the middle observations, we have chosen a convex weight function that gives minimum weightage to these observations. To compute the WLSEs, the initial guesses of both α and β are obtained based on grid search. In case of α, the grid size is $\frac{\pi}{N}$ and for β it is $\frac{\pi}{N^2}$. In both the cases the ranges are $[0, \pi]$. We have used Nelder-Mead algorithm to obtain the final estimates. In the figures to follow, we report the MSEs and MADs of the WLSEs and those of LSEs and LADEs for comparison. It can be seen that the proposed method outperforms the LSEs for all the error variances and sample sizes considered. It has been observed that the WLSEs and LADEs seem to exhibit identical performance as revealed by the curves that lie on top of each other in Figure 1.

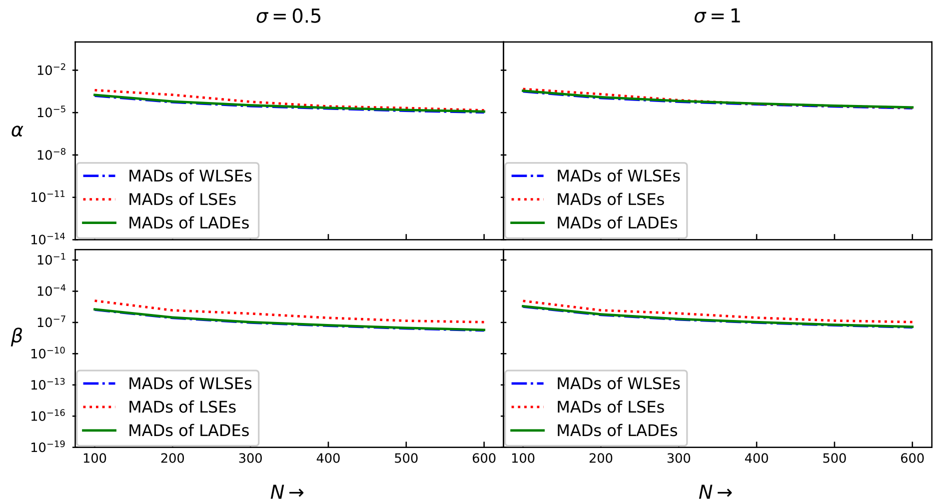

Figure 2: In each sub-plot, the graphed lines represent the MADs of the WLSEs, the LSEs and the LADEs of the simulated one component model.

For a clearer picture of the comparison between the performance of WLSEs and LADEs, we report the ratio of the MSEs of the WLSEs to those of LADEs and the ratio of MADs of the WLSEs to those of LADEs in the subsequent figures. The plots reveal that the MSEs and MADs of WLSEs are slightly smaller than those of LADEs. It should be noted that although the performance of WLSEs is at par with that of the LADEs, there is a significant difference in their computational time. The main difference is in the computational complexity involved in finding the initial values for the two estimators. For the WLSEs, we need to perform initial value computations of the order $O(N^3)$ whereas for the LADEs, it is of the order $O(N^4)$.

In the next set of experiments, we consider a more general model with two sinusoidal components and one chirplet:

$$\begin{aligned} y(n) = 10\cos(1.5n) + 10\sin(1.5n) + 10\cos(0.1n^2) + 10\sin(0.1n^2) \\ +10\cos(2.2n) + 10\sin(2.2n) + X(n). \end{aligned} \tag{20}$$

Different data sets are simulated from the above model with different error variances and sample sizes. To assess the performance of the WLSEs of the parameters of this model, we

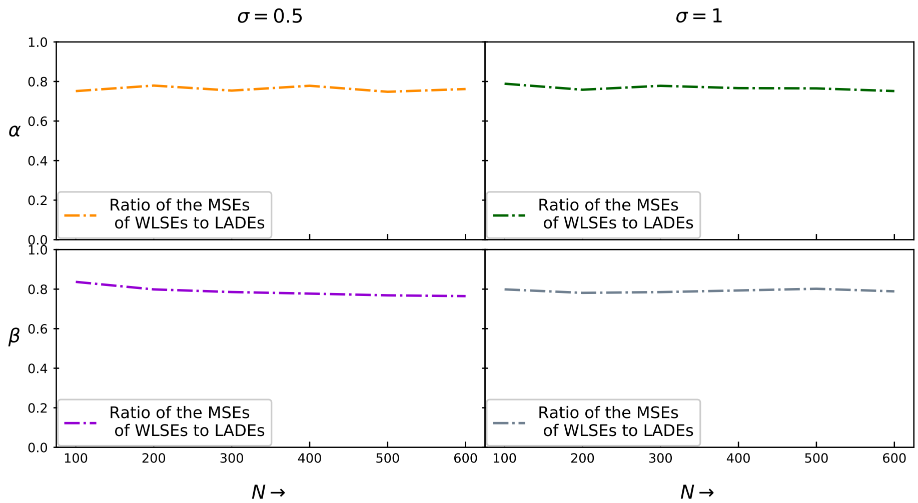

Figure 3: In each sub-plot, the graphed line represents the ratio of MSEs of the WLSEs to those of the LADEs of the simulated one component model.

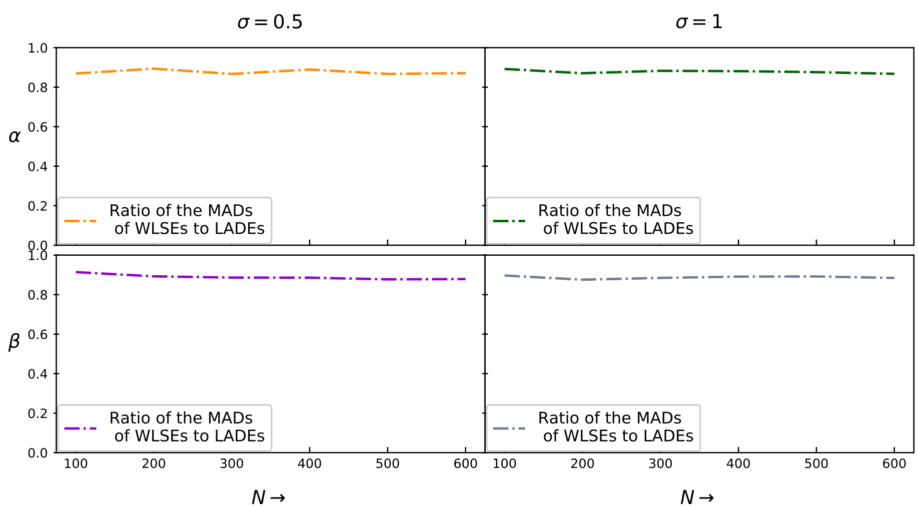

Figure 4: In each sub-plot, the graphed line represents the ratio of MADs of the WLSEs to those of the LADEs of the simulated one component model.

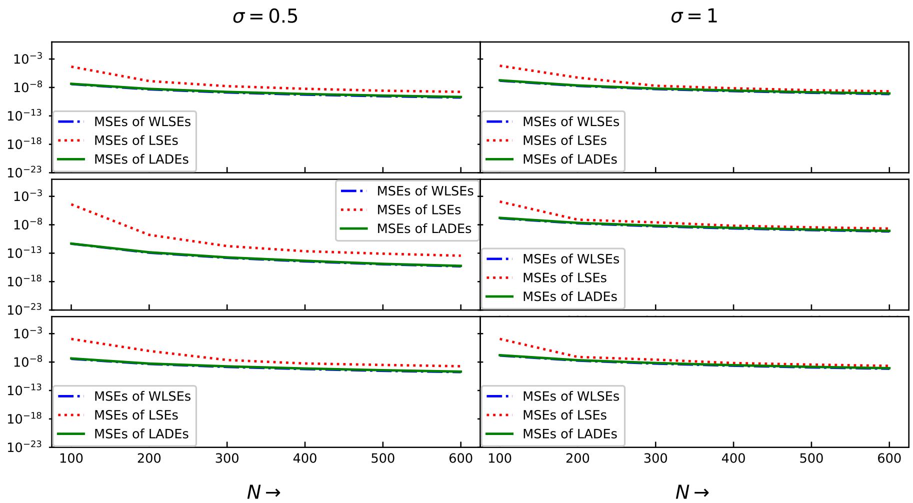

Figure 5: In each sub-plot, the graphed lines represent the MSEs of the WLSEs, the LSEs and the LADEs of the simulated model defined in (20).

again add a few outliers in the middle of each data set. For every value of N and σ^2, 5000 realisations are generated and estimates are obtained. For comparison, we also compute the LSEs and the LADEs of these parameters. The MSEs and the MADs of the three estimators are reported in Figure 5 and 6.

In this case also, we observe the ratio of MSEs and MADs of WLSEs to those of LADEs. The results are shown in Figure 7 and 8. It is observed that for all choices of N and σ^2 considered here, the ratio is less than 1.

Although the difference between the performance of WLSEs and LADEs seem subtle, it is important to note that finding the LADEs even for a model with one sinusoidal component and one chirp component involves solving a six-dimensional optimisation problems. As pointed out before, to find the initial guesses, we need to perform a grid search of order $O(N^3)$ for WLSEs and that of order $O(N^4)$ for the LADEs. However in case of the WLSEs, we can use the proposed sequential algorithm which reduces the multidimensional optimisation problem into a cascade of one-dimensional optimisation problems. In this case, to find the

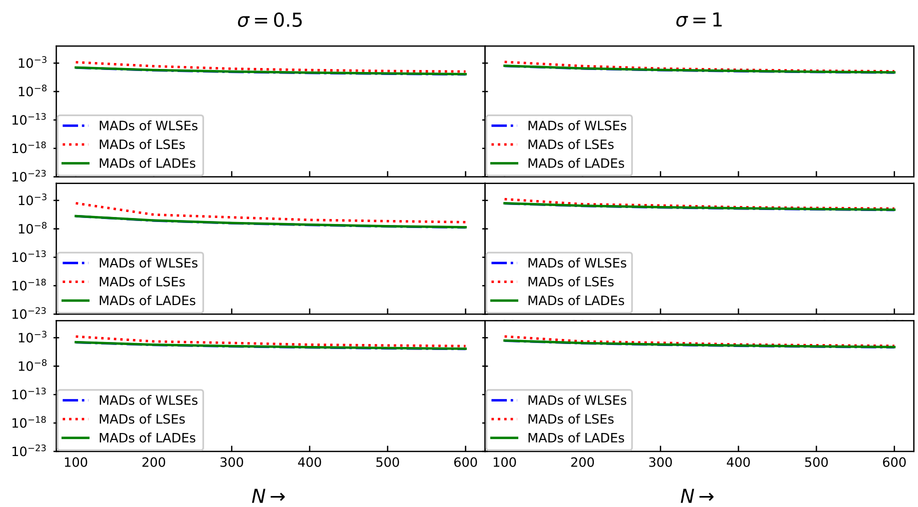

Figure 6: In each sub-plot, the graphed lines represent the MADs of the WLSEs, the LSEs and the LADEs of the simulated model defined in (20).

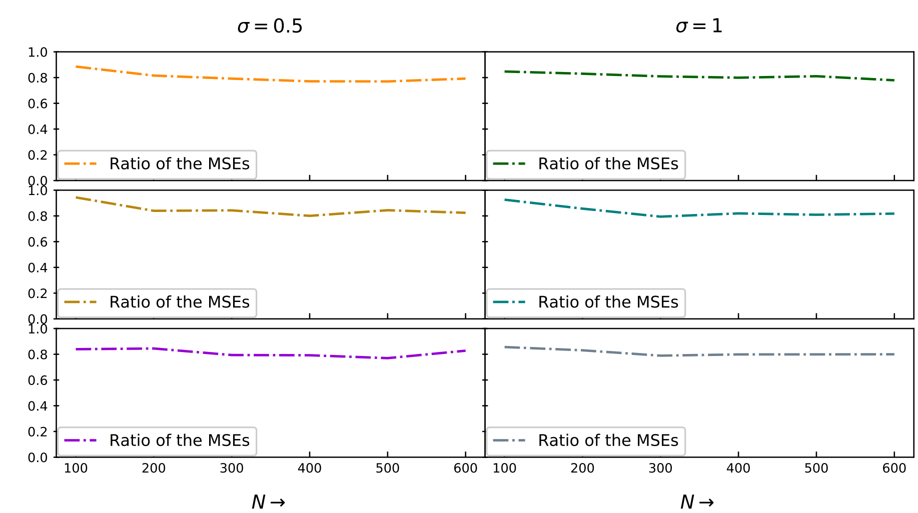

Figure 7: In each sub-plot, the graphed line represents ratio of the MSEs of the WLSEs to those of the LADEs of the simulated model defined in (20).

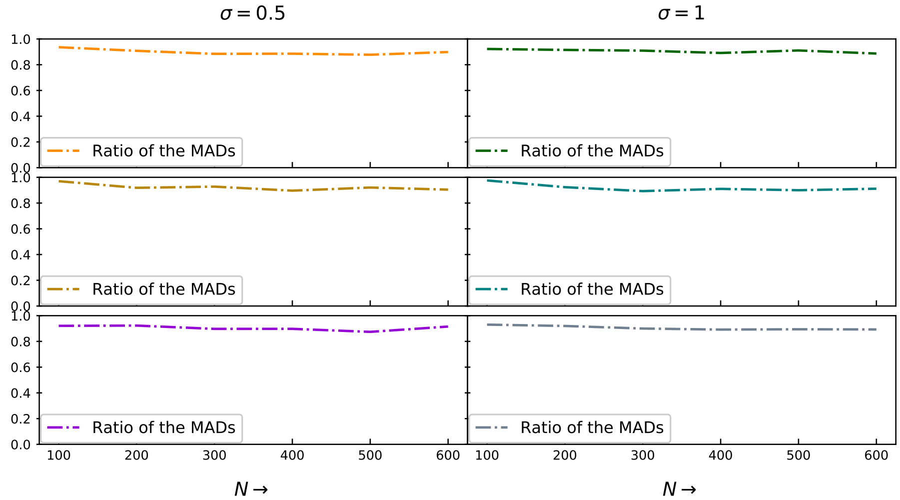

Figure 8: In each sub-plot, the graphed line represents ratio of the MADs of the WLSEs to those of the LADEs of the simulated model defined in (20).

initial values for the unknown parameters, the computational cost reduces from $O(N^3)$ to $O(N + N^2)$ for the sequential WLSEs. This makes the proposed method, as opposed to LADEs, faster as well as practically feasible to implement. It is important to note that for the model with multiple sinusoids and chirplets, say p and q respectively, finding the initial guesses for sequntial WLSEs is of order $O(pN + qN^2)$. On the other hand, finding the initial guesses for LADEs is of the order $O(N^{4(p+q)})$. Finding LADEs is, therefore, highly computationally complex and might be infeasible in practice for large values of p and q, whereas sequential WLSEs overcome this problem in an efficient manner.

5 Data Analyses

In this section, we demonstrate how to implement the proposed sequential method efficiently to describe an observed EEG data using a chirp-like model. This signal originates from a study to examine EEG correlates of genetic predisposition to alcoholism. We consider one segment of the data sampled at 256 Hz for 1 second. We scale the signal values to lie between

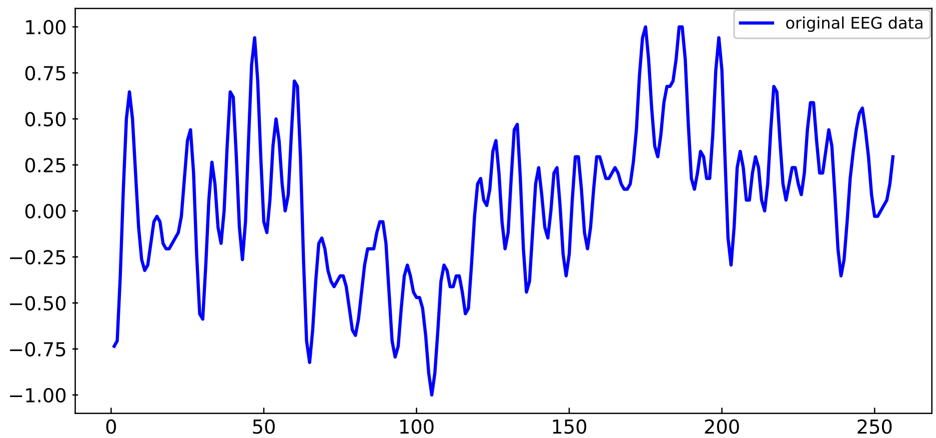

Figure 9: The scaled EEG signal.

-1 and 1 and the resultant signal is plotted in Figure 9.

To assess the applicability and robustness of the sequential estimators, we add white Gaussian noise with mean 0 and variance 100 to the middle most observation. The maximum model order is set to $p + q = 30$. To find the initial guesses for the non-linear parameters, we have used grid search method with the same grid sizes as it has been used for simulation experiments also. The final estimates are then found using the Nelder-Mead optimisation algorithm. Following form of the BIC criterion is used for model selection:

$$\text{BIC}[p, q] = N \log_e(SS_{residuals}[p, q]) + 0.5(5p + 7q + \text{arma}_{a,b} + 1) \log_e(N),$$

where the $SS_{residuals}[p, q]$ is the residuals sum of squares when p number of sinusoidal components and q number of chirp components are fitted to the data. This form of BIC is used to account for the fact that the magnitude of penalty should depend on the type of model parameter. This is inspired from Djuric's asymptotic MAP rule [1], where the frequency parameter of a sinusoidal signal is shown to contribute with a three times larger penalty term than the sinusoidal amplitude. Therefore $p + p + 3p = 5p$ is the penalty term corresponding to the sinusoidal part of the signal and extrapolating that for the chirp parameter,

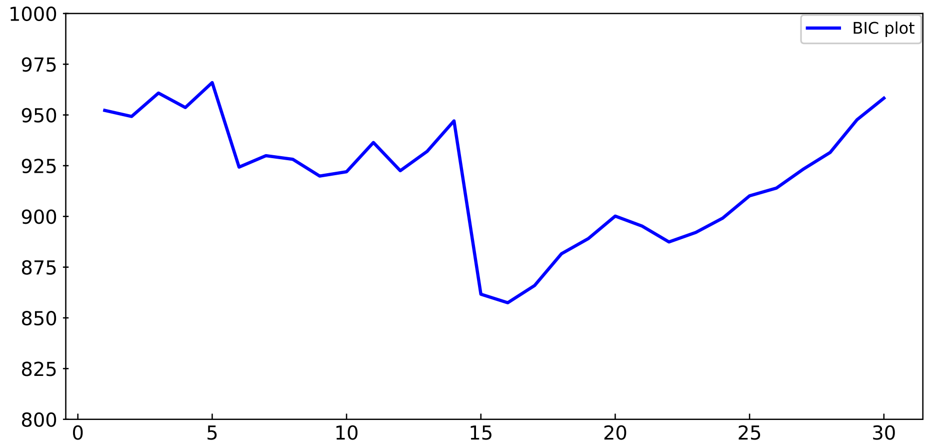

Figure 10: The BIC plot with respect to the number of components.

$q + q + 5q = 7q$ is the penalty term corresponding to the chirp component of the signal. The BIC plot is shown in Figure 10.

The estimated number of components are chosen corresponding to the minimum value of BIC. In Figure 11, the estimated signal is shown with $p = 5$ and $q = 11$. We can see that the algorithm provides a reasonably accurate fit to the data.

In Figure 12, we plot the residuals obtained from the above fit. Using the Augmented Dickey Fuller (ADF) test, we test the stationarity of these residuals. Based on the resultant small p-value, we reject the null hypothesis that the time series has a unit root and thereby conclude that the residuals are stationary. To test for the linear stationary process of the residual, we have used the 'auto.arima' function in 'forecast' package in R to fit an ARMA model to the residuals. It fits ARMA(4,4) with the following coefficients: ar(1) = 0.954, ar(2) = -0.334, ar(3) = -0.278, ar(4) = 0.406, ma(1) = -0.664, ma(2) = 0.169, ma(3) = 0.286 and ma(4) = -0.627. It indicates that the residual is a linear stationary process, which satistifies the model assumptions.

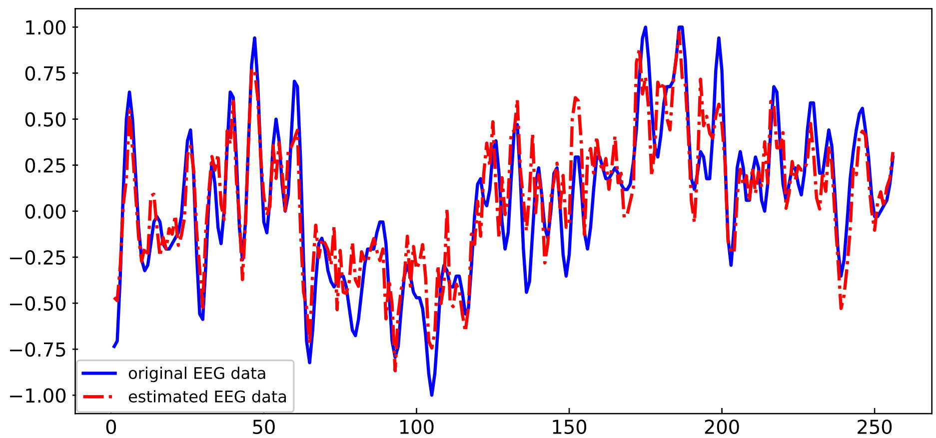

Figure 11: The estimated EEG signal along with the original EEG signal.

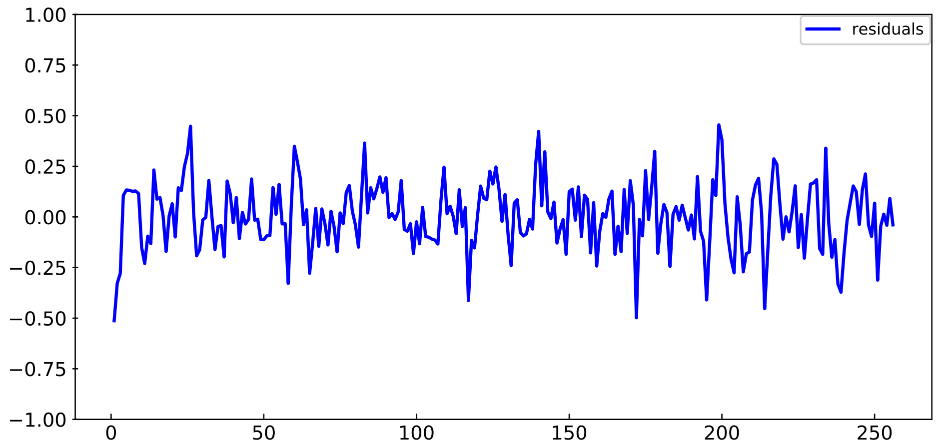

Figure 12: The residuals plot of the EEG signal.

6 Conclusions

In this paper, we have studied the WLSEs to estimate the unknown parameters of a chirp like model. The chirp like model has been proposed as an alternative to the well known chirp model. Similar to the chirp model, this model also can be used for future prediction also. The chirp like model is a combination of sinusoidal components and chirplet components observed in additive noise. We have mainly considered the problem of estimation of parameters in presence of outliers. The least squares method does not work well in presence of outliers. We propose to use the weighted least squares method to estimate the unknown parameters. The LADEs usually work well in presence of outliers, but establishing theoretical properties of LADEs in case of sinusoidal as well as chirplet model is not straight forward specially in case of multicomponent models. We have proved that the WLSEs of the unknown parameters of the chirp like model are strongly consistent and asymptotically normal under the assumption of linear stationary error. Numerical experiments using moderate size samples reveal that WLSEs perform better than LSEs and even LADEs in terms of MSEs and MADs.

Finally it should be mentioned that in this paper we have assumed that the error component is a linear stationary process. All the theoretical developments are based on this assumption. But in practice the error compeonent may not be a linear stationary process. In case of a nonlinear time series residuals all the theoretical results need to be developed. It is definitely a non-trivial extension, more work is needed in that direction.

Aknowledgements

The authors would like to thank two unknown reviewers for their constructive suggestions which have helped to improve the earlier drafts of the manuscript significantly.

Conflict of Interests and Funding Information

The authors do not have any conflict of interets. Part of the work of the first author has been funded by the Science and Engineering Research Board, Government of India.

Appendix A: Preliminaries

To establish the consistency and asymptotic normality of the WLSEs we need some trigonometric and number theoretic results and one famous number theoretic conjecture. We explicitly mention it here for easy reference.

Result A.1: If $\alpha, \beta \in (0, \pi)$, and $\alpha \neq \beta$, then the following results hold.

$$\begin{aligned}
\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \cos(\alpha n) &= \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \sin(\alpha n) = 0, \\
\lim_{N\to\infty} \frac{1}{N^{k+1}}\sum_{n=1}^{N} n^k \cos^2(\alpha n) &= \frac{1}{2(k+1)}, \\
\lim_{N\to\infty} \frac{1}{N^{k+1}}\sum_{n=1}^{N} n^k \sin^2(\alpha n) &= \frac{1}{2(k+1)}, \\
\lim_{N\to\infty} \frac{1}{N^{k+1}}\sum_{n=1}^{N} n^k \cos(\alpha n)\sin(\alpha n) &= 0, \\
\lim_{N\to\infty} \frac{1}{N^{k+1}}\sum_{n=1}^{N} n^k \sin(\alpha n)\sin(\beta n) &= \lim_{N\to\infty} \frac{1}{N^{k+1}}\sum_{n=1}^{N} n^k \cos(\alpha n)\cos(\beta n) = 0,
\end{aligned}$$

where $k = 0, 1, 2, \ldots$.

Proof: The proofs can be found in Mangulis [11]. ∎

Result A.2: If $\alpha, \beta \in (0, \pi)$, and $\alpha \neq \beta$, then except for countable number of points, the following results hold.

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

$$
\begin{aligned}
\lim_{N\to\infty} \frac{1}{N^{k+1}} \sum_{n=1}^{N} n^k \cos^2(\alpha n^2) &= \frac{1}{2(k+1)}, \\
\lim_{N\to\infty} \frac{1}{N^{k+1}} \sum_{n=1}^{N} n^k \sin^2(\alpha n^2) &= \frac{1}{2(k+1)}, \\
\lim_{N\to\infty} \frac{1}{N^{k+1}} \sum_{n=1}^{N} n^k \cos(\alpha n^2) \sin(\beta n^2) &= 0, \\
\lim_{N\to\infty} \frac{1}{N^{k+1}} \sum_{n=1}^{N} n^k \sin(\alpha n^2) \cos(\beta n) &= \lim_{N\to\infty} \frac{1}{N^{k+1}} \sum_{n=1}^{N} n^k \cos(\alpha n^2) \sin(\beta n) = 0, \\
\lim_{N\to\infty} \frac{1}{N^{k+1}} \sum_{n=1}^{N} n^k \sin(\alpha n^2) \sin(\beta n) &= \lim_{N\to\infty} \frac{1}{N^{k+1}} \sum_{n=1}^{N} n^k \cos(\alpha n^2) \cos(\beta n) = 0.
\end{aligned}
$$

In addition if $\alpha \neq \beta$, then for $k = 0, 1, 2, \ldots$,

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

where $k = 0, 1, 2, \ldots$.

Proof: The proofs can be obtained from Vinogradov's [16] results. See Lahiri, Kundu and Mitra [9] for details.

The following well known number theoretic conjecture, see for example Montgomery [12], can not be established formally. But extensive numerical experiments indicate that it holds true.

Conjecture A: If $\alpha, \beta \in (0, \pi)$, then except for countable number of points, for $k = 0, 1, 2, \ldots$,

$$
\begin{aligned}
\lim_{N\to\infty} \frac{1}{\sqrt{N} N^k} \sum_{n=1}^{N} n^k \cos(\alpha n^2) \sin(\beta n^2) &= 0, \\
\lim_{N\to\infty} \frac{1}{\sqrt{N} N^k} \sum_{n=1}^{N} n^k \cos(\alpha n^2) \sin(\beta n) &= 0, \\
\lim_{N\to\infty} \frac{1}{\sqrt{N} N^k} \sum_{n=1}^{N} n^k \sin(\alpha n^2) \cos(\beta n) &= 0,
\end{aligned}
$$

$$\begin{aligned}
\lim_{N\to\infty} \frac{1}{\sqrt{N}N^k} \sum_{n=1}^{N} n^k \cos(\alpha n^2)\cos(\beta n) &= 0, \\
\lim_{N\to\infty} \frac{1}{\sqrt{N}N^k} \sum_{n=1}^{N} n^k \sin(\alpha n^2)\sin(\beta n) &= 0.
\end{aligned}$$

In addition if $\alpha \neq \beta$, then for $k = 0, 1, 2, \ldots$,

$$\begin{aligned}
\lim_{N\to\infty} \frac{1}{\sqrt{N}N^k} \sum_{n=1}^{N} n^k \cos(\alpha n^2)\cos(\beta n^2) &= 0, \\
\lim_{N\to\infty} \frac{1}{\sqrt{N}N^k} \sum_{n=1}^{N} n^k \sin(\alpha n^2)\sin(\beta n^2) &= 0.
\end{aligned}$$

Appendix B: Proof of Theorem 1.

We need the following lemmas to prove Theorem 1.

Lemma B-1: Let $\{e(n)\}$ be a sequence of i.i.d. random variables with mean zero and finite fourth moment, $w(t)$ satisfies Assumption 1, then

$$E\left|\sum_{n=1}^{N} w\left(\frac{n}{N}\right) w^2\left(\frac{n+1}{N}\right) w\left(\frac{n+2}{N}\right) e(n)e^2(n+1)e(n+2)\right| = O(N^{\frac{1}{2}}).$$

Proof:

$$E\left|\sum_{n=1}^{N} w\left(\frac{n}{N}\right) w^2\left(\frac{n+1}{N}\right) w\left(\frac{n+2}{N}\right) e(n)e^2(n+1)e(n+2)\right| \leq$$

$$\left[E\left(\sum_{n=1}^{N} w\left(\frac{n}{N}\right) w^2\left(\frac{n+1}{N}\right) w\left(\frac{n+2}{N}\right) e(n)e^2(n+1)e(n+2)\right)^2\right]^{\frac{1}{2}} = O(N^{\frac{1}{2}}).$$

Similarly, it follows that

$$E\left|\sum_{n=1}^{N} w\left(\frac{n}{N}\right) w\left(\frac{n+1}{N}\right) w\left(\frac{n+k}{N}\right) w\left(\frac{n+k+1}{N}\right) e(n)e(n+1)e(n+k)e(n+k+1)\right| = O(N^{\frac{1}{2}}),$$

for some fixed k, where $k = 2, 3, \ldots$.

LEMMA B-2: Let $\{e(n)\}$ be a sequence of i.i.d. random variables with mean zero and finite fourth moment, $w(t)$ satisfies Assumption 1, then for arbitrary integers $m, k \geq 1$,

$$E \sup_{\theta} \left| \sum_{n=1}^{N} w\left(\frac{n}{N}\right) w\left(\frac{n+k}{N}\right) e(n)e(n+k)e^{im\theta n} \right| = O(N^{\frac{3}{4}}).$$

PROOF:

$$\begin{aligned}
& E \sup_{\theta} \left| \sum_{n=1}^{N} w\left(\frac{n}{N}\right) w\left(\frac{n+k}{N}\right) e(n)e(n+k)e^{im\theta n} \right| \leq \\
& \left[E \sup_{\theta} \left| \sum_{n=1}^{N} w\left(\frac{n}{N}\right) w\left(\frac{n+k}{N}\right) e(n)e(n+k)e^{im\theta n} \right|^2 \right]^{\frac{1}{2}} = \\
& \left[E \sup_{\theta} \left(\sum_{n=1}^{N} w\left(\frac{n}{N}\right) w\left(\frac{n+k}{N}\right) e(n)e(n+k)e^{im\theta n} \right) \right. \\
& \qquad \left. \left(\sum_{n=1}^{N} w\left(\frac{n}{N}\right) w\left(\frac{n+k}{N}\right) e(n)e(n+k)e^{-im\theta n} \right) \right]^{\frac{1}{2}} \\
\leq \; & \left[E \sum_{n=1}^{N} w^2\left(\frac{n}{N}\right) w^2\left(\frac{n+k}{N}\right) e^2(n)e^2(n+k) + \right. \\
& 2E \left| \sum_{n=1}^{N-1} w\left(\frac{n}{N}\right) w\left(\frac{n+k}{N}\right) w\left(\frac{n+1}{N}\right) w\left(\frac{n+k+1}{N}\right) e(n)e(n+k)e(n+1)e(n+k+1) \right| + \\
& \left. \ldots + 2E \left| w\left(\frac{1}{N}\right) w\left(\frac{1+k}{N}\right) w(1)\, w\left(\frac{N+k}{N}\right) e(1)e(1+k)e(N)e(N+k) \right| \right]^{\frac{1}{2}} \\
= \; & O(N + N.N^{\frac{1}{2}})^{\frac{1}{2}} = O(N^{\frac{3}{4}}).
\end{aligned}$$

LEMMA B-3: Let $\{e(n)\}$ be a sequence of i.i.d. random variables with mean zero and finite fourth moment, $w(t)$ satisfies Assumption 1, then

$$E \sup_{\beta} \left| \sum_{n=1}^{N} w\left(\frac{n}{N}\right) e(n)e^{i\beta n^2} \right|^2 = O(N^{\frac{7}{4}}).$$

PROOF:

$$E \sup_{\beta} \left| \sum_{n=1}^{N} w\left(\frac{n}{N}\right) e(n)e^{i\beta n^2} \right|^2 =$$

$$E\sup_{\beta}\left(\sum_{n=1}^{N} w\left(\frac{n}{N}\right)e(n)e^{i\beta n^2}\right)\left(\sum_{n=1}^{N} w\left(\frac{n}{N}\right)e(n)e^{-i\beta n^2}\right) = O(N + N.N^{\frac{3}{4}}) = O(N^{\frac{7}{4}})$$

LEMMA B-4: Let $\{e(n)\}$ be a sequence of i.i.d. random variables with mean zero and finite fourth moment, $w(t)$ satisfies Assumption 1, then

$$E\sup_{\beta}\left|\frac{1}{N}\sum_{n=1}^{N} w\left(\frac{n}{N}\right)e(n)e^{i\beta n^2}\right| \le O(N^{-\frac{1}{8}}).$$

PROOF:

$$E\sup_{\beta}\left|\frac{1}{N}\sum_{n=1}^{N} w\left(\frac{n}{N}\right)e(n)e^{i\beta n^2}\right| \le \left[E\sup_{\beta}\left|\frac{1}{N}\sum_{n=1}^{N} w\left(\frac{n}{N}\right)e(n)e^{i\beta n^2}\right|^2\right]^{\frac{1}{2}} = O(N^{-\frac{1}{8}}).$$

LEMMA B-5: Let $\{e(n)\}$ be a sequence of i.i.d. random variables with mean zero and finite fourth moment, $w(t)$ satisfies Assumption 1, and $\{X(n)\}$ is same as defined in (4), then

$$E\sup_{\beta}\left|\frac{1}{N}\sum_{n=1}^{N} w\left(\frac{n}{N}\right)X(n)e^{i\beta n^2}\right| \le O(N^{-\frac{1}{8}}).$$

PROOF:

$$\begin{aligned}
&E\sup_{\beta}\left|\frac{1}{N}\sum_{n=1}^{N} w\left(\frac{n}{N}\right)X(n)e^{i\beta n^2}\right| = \\
&E\sup_{\beta}\left|\frac{1}{N}\sum_{n=1}^{N}\sum_{k=-\infty}^{\infty} a(k)e(n-k)w\left(\frac{n}{N}\right)e^{i\beta n^2}\right| \le \\
&\sum_{k=-\infty}^{\infty}|a(k)|\left[E\sup_{\beta}\left|\frac{1}{N}\sum_{n=1}^{N} e(n-k)w\left(\frac{n}{N}\right)e^{i\beta n^2}\right|\right] = O(N^{-\frac{1}{8}}).
\end{aligned}$$

Since $E\sup_{\beta}\left|\frac{1}{N}\sum_{n=1}^{N} e(n-k)w\left(\frac{n}{N}\right)e^{i\beta n^2}\right|$ is independent of k, the result follows from Lemma B-4.

LEMMA B-6: Let $\{e(n)\}$ be a sequence of i.i.d. random variables with mean zero and finite fourth moment, $w(t)$ satisfies Assumption 1, and $\{X(n)\}$ is same as defined in (4), then

$$\sup_{\beta}\left|\frac{1}{N}\sum_{n=1}^{N} w\left(\frac{n}{N}\right)X(n)e^{i\beta n^2}\right| \longrightarrow 0, \quad a.s.$$

PROOF: Consider the sequence N^9, then we obtain

$$E \sup_{\beta} \left| \frac{1}{N^9} \sum_{n=1}^{N^9} w\left(\frac{n}{N}\right) X(n) e^{i\beta n^2} \right| \leq O(N^{-\frac{9}{8}}).$$

Therefore, using Borel Cantelli lemma, it follows that

$$\sup_{\beta} \left| \frac{1}{N^9} \sum_{n=1}^{N^9} w\left(\frac{n}{N}\right) X(n) e^{i\beta n^2} \right| \longrightarrow 0, \qquad a.s.$$

Now consider J, such that $N^9 < J \leq (N+1)^9$, then

$$\begin{aligned}
&\sup_{\beta} \sup_{N^9 < J \leq (N+1)^9} \left| \frac{1}{N^9} \sum_{n=1}^{N^9} w\left(\frac{n}{N}\right) X(n) e^{i\beta n^2} - \frac{1}{J} \sum_{n=1}^{J} w\left(\frac{n}{N}\right) X(n) e^{i\beta n^2} \right| = \\
&\sup_{\beta} \sup_{N^9 < J \leq (N+1)^9} \left| \frac{1}{N^9} \sum_{n=1}^{N^9} w\left(\frac{n}{N}\right) X(n) e^{i\beta n^2} - \frac{1}{N^9} \sum_{n=1}^{J} w\left(\frac{n}{N}\right) X(n) e^{i\beta n^2} + \right. \\
&\qquad \left. \frac{1}{N^9} \sum_{n=1}^{J} w\left(\frac{n}{N}\right) X(n) e^{i\beta n^2} - \frac{1}{J} \sum_{n=1}^{J} w\left(\frac{n}{N}\right) X(n) e^{i\beta n^2} \right| \leq \\
&\sup_{\beta} \sup_{N^9 < J \leq (N+1)^9} \left| \frac{1}{N^9} \sum_{n=1}^{N^9} w\left(\frac{n}{N}\right) X(n) e^{i\beta n^2} - \frac{1}{N^9} \sum_{n=1}^{J} w\left(\frac{n}{N}\right) X(n) e^{i\beta n^2} \right| + \\
&\sup_{\beta} \sup_{N^9 < J \leq (N+1)^9} \left| \frac{1}{N^9} \sum_{n=1}^{J} w\left(\frac{n}{N}\right) X(n) e^{i\beta n^2} - \frac{1}{J} \sum_{n=1}^{J} w\left(\frac{n}{N}\right) X(n) e^{i\beta n^2} \right| \leq \\
&\qquad \frac{K}{N^9} \sum_{n=N^9+1}^{(N+1)^9} |X(n)| + K \sum_{n=1}^{(N+1)^9} |X(n)| \left(\frac{1}{N^9} - \frac{1}{(N+1)^9} \right)
\end{aligned}$$

Note that the mean squared error of the first term is of the order $O\left(\frac{1}{N^{18}} \times ((N+1)^9 - N^9)^2\right) = O(N^{-2})$. Similarly, the mean squared error of the second term is of the order $O\left(N^{18} \times \left(\frac{(N+1)^9 - N^9}{N^{18}}\right)^2\right) = O(N^{-2})$. Therefore, both the terms converge to zero almost surely.

Along the same line the following result follows.

LEMMA B-7: Let $\{e(n)\}$ be a sequence of i.i.d. random variables with mean zero and finite fourth moment, $w(t)$ satisfies Assumption 1, and $\{X(n)\}$ is same as defined in (4), then

$$\sup_{\alpha} \left| \frac{1}{N} \sum_{n=1}^{N} w\left(\frac{n}{N}\right) X(n) e^{i\alpha n} \right| \longrightarrow 0, \qquad a.s.$$

Lemma B-8: Let us denote

$$S_c = \{\boldsymbol{\theta} : \boldsymbol{\theta} = (A, B, C, D, \alpha, \beta)^\top, |\boldsymbol{\theta} - \boldsymbol{\theta}^0| \geq 6c\}.$$

If there exists a $c > 0$,

$$\varliminf \inf_{\boldsymbol{\theta} \in S_c} \frac{1}{N}[Q(\boldsymbol{\theta}) - Q(\boldsymbol{\theta}^0)] > 0 \quad a.s. \tag{21}$$

then $\widehat{\boldsymbol{\theta}}$, the WLSE of $\boldsymbol{\theta}^0$, is a strongly consistent estimator of $\boldsymbol{\theta}^0$.

Proof: It follows using simple arguments by contradiction, exactly similar to the lemma by Wu [17]. ∎

Proof of Theorem 1:

Let us denote

$$\mu(n; \boldsymbol{\theta}) = A\cos(\alpha n) + B\sin(\alpha n) + C\cos(\beta n^2) + D\sin(\beta n^2). \tag{22}$$

Consider

$$\begin{aligned}
\frac{1}{N}[Q(\boldsymbol{\theta}) - Q(\boldsymbol{\theta}^0)] &= \frac{1}{N}\left[\sum_{n=1}^{N} w\left(\frac{n}{N}\right)(y(n) - \mu(n;\boldsymbol{\theta}))^2 - \sum_{n=1}^{N} w\left(\frac{n}{N}\right) X^2(n)\right] \\
&= \frac{1}{N}\left[\sum_{n=1}^{N} w\left(\frac{n}{N}\right)(\mu(n;\boldsymbol{\theta}^0) - \mu(n;\boldsymbol{\theta}))^2\right] \\
&\quad + \frac{2}{N}\left[\sum_{n=1}^{N} w\left(\frac{n}{N}\right) X(n)(\mu(n;\boldsymbol{\theta}^0) - \mu(n;\boldsymbol{\theta}))\right] \\
&= f_1(\boldsymbol{\theta}) + f_2(\boldsymbol{\theta}).
\end{aligned}$$

Here

$$\begin{aligned}
f_1(\boldsymbol{\theta}) &= \frac{1}{N}\left[\sum_{n=1}^{N} w\left(\frac{n}{N}\right)(\mu(n;\boldsymbol{\theta}^0) - \mu(n;\boldsymbol{\theta}))^2\right], \\
f_2(\boldsymbol{\theta}) &= \frac{2}{N}\left[\sum_{n=1}^{N} w\left(\frac{n}{N}\right) X(n)(\mu(n;\boldsymbol{\theta}^0) - \mu(n;\boldsymbol{\theta}))\right].
\end{aligned}$$

Consider

$$\begin{aligned}
S_{c,1} &= \{\boldsymbol{\theta} : \boldsymbol{\theta} = (A, B, C, D, \alpha, \beta)^\top, |A - A^0| \geq c\} \\
S_{c,2} &= \{\boldsymbol{\theta} : \boldsymbol{\theta} = (A, B, C, D, \alpha, \beta)^\top, |B - B^0| \geq c\} \\
S_{c,3} &= \{\boldsymbol{\theta} : \boldsymbol{\theta} = (A, B, C, D, \alpha, \beta)^\top, |C - C^0| \geq c\} \\
S_{c,4} &= \{\boldsymbol{\theta} : \boldsymbol{\theta} = (A, B, C, D, \alpha, \beta)^\top, |D - D^0| \geq c\} \\
S_{c,5} &= \{\boldsymbol{\theta} : \boldsymbol{\theta} = (A, B, C, D, \alpha, \beta)^\top, |\alpha - \alpha^0| \geq c\} \\
S_{c,6} &= \{\boldsymbol{\theta} : \boldsymbol{\theta} = (A, B, C, D, \alpha, \beta)^\top, |\beta - \beta^0| \geq c\}.
\end{aligned}$$

Now $S_c \in \cup_{j=1}^{6} S_{c,j} = S$. Therefore,

$$\underline{\lim} \inf_{\theta \in S_c} f_1(\boldsymbol{\theta}) \geq \underline{\lim} \inf_{\boldsymbol{\theta} \in S} f_1(\theta) = \underline{\lim} \inf_{\boldsymbol{\theta} \in \cup_j S_{c,j}} f_1(\theta).$$

Now

$$\begin{aligned}
\underline{\lim} \inf_{\boldsymbol{\theta} \in S_{c,1}} f_1(\theta) &= \underline{\lim} \inf_{|A - A^0| \geq c} (A - A^0)^2 \frac{1}{N} \sum_{n=1}^{N} w\left(\frac{n}{N}\right) \cos^2(\alpha^0 n) \\
&\geq \gamma \ \underline{\lim} \inf_{|A - A^0| \geq c} (A - A^0)^2 \frac{1}{N} \sum_{n=1}^{N} \cos^2(\alpha^0 n) > 0, \qquad \text{(using Result A.1).}
\end{aligned}$$

Similarly using Results A.1 and A.2, it can be shown for $S_{c,2}, \ldots, S_{c,6}$ also. Therefore,

$$\underline{\lim} \inf_{\boldsymbol{\theta} \in S_c} f_1(\boldsymbol{\theta}) > 0.$$

Using Lemma B-7, it follows that

$$\lim \sup_{\boldsymbol{\theta}} |f_2(\boldsymbol{\theta})| = 0,$$

therefore

$$\underline{\lim} \inf_{\boldsymbol{\theta} \in S_c} \frac{1}{N} [Q(\boldsymbol{\theta}) - Q(\boldsymbol{\theta}^0)] > 0 \quad a.s.$$

Using Lemma B-8, the result follows.

∎

Appendix C: Proof of Theorem 2.

We need the following lemmas to prove Theorem 2.

Lemma C-1: If $0 < \alpha, \beta < \pi$, and $w(t)$ satisfies Assumption 1, then for $k = 0, 1, 2, \ldots$,

(a) $$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} w\left(\frac{n}{N}\right)\sin^2(\alpha n) = \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} w\left(\frac{n}{N}\right)\cos^2(\alpha n) = \frac{1}{2}\int_0^1 w(t)dt > \frac{\gamma}{2},$$

(b) $$\lim_{N\to\infty} \frac{1}{N^{k+1}}\sum_{n=1}^{N} n^k w\left(\frac{n}{N}\right)\sin(\alpha n) = \lim_{N\to\infty} \frac{1}{N^{k+1}}\sum_{n=1}^{N} n^k w\left(\frac{n}{N}\right)\cos(\alpha n) = 0,$$

(c) $$\lim_{N\to\infty} \frac{1}{N^{k+1}}\sum_{n=1}^{N} n^k w\left(\frac{n}{N}\right)\sin^2(\alpha n) = \lim_{N\to\infty} \frac{1}{N^{k+1}}\sum_{n=1}^{N} n^k w\left(\frac{n}{N}\right)\cos^2(\alpha n)$$
$$= \frac{1}{2}\int_0^1 t^k w(t)dt = \frac{c_{k+1}}{2} > 0,$$

(d) $$\lim_{N\to\infty} \frac{1}{N^{k+1}}\sum_{n=1}^{N} n^k w\left(\frac{n}{N}\right)\sin(\alpha n)\cos(\alpha n) = \lim_{N\to\infty} \frac{1}{N^{k+1}}\sum_{n=1}^{N} n^k w\left(\frac{n}{N}\right)\sin(\alpha n)\cos(\beta n) = 0,$$

In addition if $\alpha \neq \beta$, then,

(e) $$\lim_{N\to\infty} \frac{1}{N^{k+1}}\sum_{n=1}^{N} n^k w\left(\frac{n}{N}\right)\sin(\alpha n)\sin(\beta n) = \lim_{N\to\infty} \frac{1}{N^{k+1}}\sum_{n=1}^{N} n^k w\left(\frac{n}{N}\right)\cos(\alpha n)\cos(\beta n) = 0.$$

Proof of Lemma C-1: Proof of (a). First we will show

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

For $\epsilon > 0$, there exists a polynomial $p_\epsilon(x)$, such that $|w(x) - p_\epsilon(x)| \leq \epsilon$, for all $x \in [0, 1]$. Hence,

$$\int_0^1 w(x)dx - \epsilon \leq \int_0^1 p_\epsilon(x)dx \leq \int_0^1 w(x)dx + \epsilon.$$

Further

$$\frac{1}{N}\sum_{n=1}^{N} p_\epsilon\left(\frac{n}{N}\right)\cos^2(\alpha n) - \frac{\epsilon}{N}\sum_{n=1}^{N}\cos^2(\alpha t) \leq \frac{1}{N}\sum_{n=1}^{N} w\left(\frac{n}{N}\right)\cos^2(\alpha n) \leq$$
$$\frac{1}{N}\sum_{n=1}^{N} p_\epsilon\left(\frac{n}{N}\right)\cos^2(\alpha n) + \frac{\epsilon}{N}\sum_{n=1}^{N}\cos^2(\alpha n).$$

Suppose

$$p_\epsilon(x) = a_0 + a_1 x + \cdots + a_k x^k \qquad \Rightarrow \qquad \int_0^1 p_\epsilon(x)dx = a_0 + \frac{a_1}{2} + \cdots + \frac{a_k}{k+1}.$$

Now due to Result A.1,

$$\begin{aligned}\frac{1}{N}\sum_{n=1}^{N} p_\epsilon\left(\frac{n}{N}\right)\cos^2(\alpha n) \quad &= \quad \frac{1}{N}\sum_{n=1}^{N}\left\{a_0 + \frac{a_1 n}{N} + \cdots + \frac{a_k n^k}{N^k}\right\}\cos^2(\alpha n) \\ &\longrightarrow \quad \frac{1}{2}\left[a_0 + \frac{a_1}{2} + \cdots + \frac{a_k}{k+1}\right] = \frac{1}{2}\int_0^1 p_\epsilon(x)dx.\end{aligned}$$

Therefore,

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} p_\epsilon\left(\frac{n}{N}\right)\cos^2(\alpha n) - \frac{\epsilon}{2} \le \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} w\left(\frac{n}{N}\right)\cos^2(\alpha n) \le$$
$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} p_\epsilon\left(\frac{n}{N}\right)\cos^2(\alpha n) + \frac{\epsilon}{2}.$$

Hence

$$\frac{1}{2}\int_0^1 w(t)dt - \frac{\epsilon}{2} \le \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} w\left(\frac{n}{N}\right)\cos^2(\alpha n) \le \frac{1}{2}\int_0^1 w(t)dt + \frac{\epsilon}{2}.$$

Since ϵ is arbitrary, the result follows.

The result involving $\sin^2(\alpha n)$ will go through exactly in the same way. Note that using (a), Result A.2 and by properly choosing $w(\cdot)$, (b), (c), (d) and (e) follow. ∎

Lemma C-2: If $0 < \alpha, \beta < \pi$, and $w(t)$ satisfies Assumption 1, then except for countable number of points, for $k = 0, 1, 2, \ldots$,

(a) $$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} w\left(\frac{n}{N}\right)\sin^2(\beta n^2) = \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} w\left(\frac{n}{N}\right)\cos^2(\beta n^2) = \frac{1}{2}\int_0^1 w(t)dt > \frac{\gamma}{2},$$

(b) $$\begin{aligned}\lim_{N\to\infty}\frac{1}{N^{k+1}}\sum_{n=1}^{N} n^k w\left(\frac{n}{N}\right)\sin^2(\beta n^2) &= \lim_{N\to\infty}\frac{1}{N^{k+1}}\sum_{n=1}^{N} n^k w\left(\frac{n}{N}\right)\cos^2(\beta n^2) \\ &= \frac{1}{2}\int_0^1 t^k w(t)dt = \frac{c_{k+1}}{2} > 0,\end{aligned}$$

(c) $$\lim_{N\to\infty}\frac{1}{N^{k+1}}\sum_{n=1}^{N} n^k w\left(\frac{n}{N}\right)\sin(\alpha n^2) = \lim_{N\to\infty}\frac{1}{N^{k+1}}\sum_{n=1}^{N} n^k w\left(\frac{n}{N}\right)\cos(\alpha n^2) = 0,$$

(d) $\displaystyle \lim_{N\to\infty} \frac{1}{N^{k+1}} \sum_{n=1}^{N} n^k w\left(\frac{n}{N}\right) \sin(\alpha n^2)\cos(\alpha n^2) = \lim_{N\to\infty} \frac{1}{N^{k+1}} \sum_{n=1}^{N} w\left(\frac{n}{N}\right) \sin(\alpha n^2)\cos(\beta n^2) = 0,$

(f) $\displaystyle \lim_{N\to\infty} \frac{1}{N^{k+1}} \sum_{n=1}^{N} n^k w\left(\frac{n}{N}\right) \sin(\alpha n^2)\sin(\beta n) = \lim_{N\to\infty} \frac{1}{N^{k+1}} \sum_{n=1}^{N} w\left(\frac{n}{N}\right) \cos(\alpha n^2)\cos(\beta n) = 0,$

(g) $\displaystyle \lim_{N\to\infty} \frac{1}{N^{k+1}} \sum_{n=1}^{N} n^k w\left(\frac{n}{N}\right) \sin(\alpha n^2)\cos(\beta n) = \lim_{N\to\infty} \frac{1}{N^{k+1}} \sum_{n=1}^{N} w\left(\frac{n}{N}\right) \cos(\alpha n^2)\sin(\beta n) = 0,$

In addition if $\alpha \neq \beta$, then,

(e) $\displaystyle \lim_{N\to\infty} \frac{1}{N^{k+1}} \sum_{n=1}^{N} n^k w\left(\frac{n}{N}\right) \sin(\alpha n^2)\sin(\beta n^2) = \lim_{N\to\infty} \frac{1}{N^{k+1}} \sum_{n=1}^{N} w\left(\frac{n}{N}\right) \cos(\alpha n^2)\cos(\beta n^2) = 0.$

Proof of Lemma C-2: The proof follows along the same line as the proof of Lemma C-1.

■

Lemma C-3: If $\alpha, \beta \in (0, \pi)$, and if Conjecture A is true, then except for countable number of points, for $k = 0, 1, 2, \ldots,$

$$\begin{aligned}
\lim_{N\to\infty} \frac{1}{\sqrt{N}N^k} \sum_{n=1}^{N} n^k w\left(\frac{n}{N}\right) \cos(\alpha n^2)\sin(\beta n^2) &= 0,\\
\lim_{N\to\infty} \frac{1}{\sqrt{N}N^k} \sum_{n=1}^{N} n^k w\left(\frac{n}{N}\right) \cos(\alpha n^2)\cos(\beta n) &= 0,\\
\lim_{N\to\infty} \frac{1}{\sqrt{N}N^k} \sum_{n=1}^{N} n^k w\left(\frac{n}{N}\right) \sin(\alpha n^2)\sin(\beta n) &= 0,\\
\lim_{N\to\infty} \frac{1}{\sqrt{N}N^k} \sum_{n=1}^{N} n^k w\left(\frac{n}{N}\right) \sin(\alpha n^2)\cos(\beta n) &= 0,\\
\lim_{N\to\infty} \frac{1}{\sqrt{N}N^k} \sum_{n=1}^{N} n^k w\left(\frac{n}{N}\right) \cos(\alpha n^2)\sin(\beta n) &= 0.
\end{aligned}$$

In addition if $\alpha \neq \beta$, then

$$\begin{aligned}
\lim_{N\to\infty} \frac{1}{\sqrt{N}N^k} \sum_{n=1}^{N} n^k w\left(\frac{n}{N}\right) \cos(\alpha n^2)\cos(\beta n^2) &= 0,\\
\lim_{N\to\infty} \frac{1}{\sqrt{N}N^k} \sum_{n=1}^{N} n^k w\left(\frac{n}{N}\right) \sin(\alpha n^2)\sin(\beta n^2) &= 0.
\end{aligned}$$

Proof of Lemma C-3: If Conjecture A is true, then proof follows along the same line as the proof of Lemma C-1. ∎

Proof of Theorem 2:

This proof can be obtained by expanding $Q(\boldsymbol{\theta})$ around the point $\boldsymbol{\theta}^0$. We will use the structure of the weight function $w(t)$, Lemmas C-1 and C-2 and the Central Limit Theorem of the stochastic process to obtain the asymptotic distribution of $\widehat{\boldsymbol{\theta}}$.

The criterion function is

$$Q(\boldsymbol{\theta}) = \sum_{n=1}^{N} w\left(\frac{n}{N}\right)\left(y(n) - A\cos(\alpha n) - B\sin(\alpha n) - C\cos(\beta n^2) - D\sin(\beta n^2)\right)^2,$$

therefore, the vector of first order derivatives is

$$Q'(\boldsymbol{\theta}^0) = \begin{bmatrix} \frac{\partial Q(\boldsymbol{\theta})}{\partial A} \\ \frac{\partial Q(\boldsymbol{\theta})}{\partial B} \\ \frac{\partial Q(\boldsymbol{\theta})}{\partial \alpha} \\ \frac{\partial Q(\boldsymbol{\theta})}{\partial C} \\ \frac{\partial Q(\boldsymbol{\theta})}{\partial D} \\ \frac{\partial Q(\boldsymbol{\theta})}{\partial \beta} \end{bmatrix}_{\boldsymbol{\theta}=\boldsymbol{\theta}^0} = -2 \begin{bmatrix} \sum_{n=1}^{N} w\left(\frac{n}{N}\right) X(n)\cos(\alpha^0 n) \\ \sum_{n=1}^{N} w\left(\frac{n}{N}\right) X(n)\sin(\alpha^0 n) \\ \sum_{n=1}^{N} nw\left(\frac{n}{N}\right) X(n)(B^0\cos(\alpha^0 n) - A^0\sin(\alpha^0 n)) \\ \sum_{n=1}^{N} w\left(\frac{n}{N}\right) X(n)\cos(\beta^0 n^2) \\ \sum_{n=1}^{N} w\left(\frac{n}{N}\right) X(n)\sin(\beta^0 n^2) \\ \sum_{n=1}^{N} n^2 w\left(\frac{n}{N}\right) X(n)(D^0\cos(\beta^0 n^2) - C^0\sin(\beta^0 n^2)) \end{bmatrix}$$

and the matrix of second order derivatives is

$$Q''(\boldsymbol{\theta}^0) = \begin{bmatrix} \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial A^2} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial A\partial B} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial A\partial \alpha} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial A\partial C} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial A\partial D} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial A\partial \beta} \\ \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial B\partial A} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial B^2} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial B\partial \alpha} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial B\partial C} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial B\partial D} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial B\partial \beta} \\ \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial \alpha\partial A} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial \alpha\partial B} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial \alpha^2} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial \alpha\partial C} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial \alpha\partial D} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial \alpha\partial \beta} \\ \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial C\partial A} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial C\partial B} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial C\partial \alpha} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial C^2} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial C\partial D} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial C\partial \beta} \\ \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial D\partial A} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial D\partial B} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial D\partial \alpha} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial D\partial C} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial D^2} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial D\partial \beta} \\ \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial \beta\partial A} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial \beta\partial B} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial \beta\partial \alpha} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial \beta\partial C} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial \beta\partial D} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial \beta^2} \end{bmatrix}_{\boldsymbol{\theta}=\boldsymbol{\theta}^0}.$$

The elements of $Q''(\boldsymbol{\theta}^0)$ are given in Appendix D. Consider the following diagonal matrix

$$\boldsymbol{D} = \text{diag}(N^{-1/2}, N^{-1/2}, N^{-3/2}, N^{-1/2}, N^{-1/2}, N^{-5/2}) = \begin{pmatrix} \boldsymbol{D}_1^{-1} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{D}_2^{-1} \end{pmatrix}, \tag{23}$$

where $\boldsymbol{D}_1$ and $\boldsymbol{D}_2$ are same as used in the statement of Theorem 2. Then using Lemmas C-1 and C-2, it follows that

$$\boldsymbol{D}Q'(\boldsymbol{\theta}^0) \xrightarrow{d} N_6(\mathbf{0}, \sigma^2\ \boldsymbol{\Sigma}), \tag{24}$$

where

$$\boldsymbol{\Sigma} = \begin{pmatrix} \zeta\boldsymbol{\Sigma}_1 & \mathbf{0} \\ \mathbf{0} & \eta\boldsymbol{\Sigma}_2 \end{pmatrix}.$$

Here $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ are same as defined in equation (10) and

$$\zeta = \left| \sum_{k=-\infty}^{\infty} a(k) e^{i\alpha^0 k} \right|^2, \quad \eta = \left| \sum_{k=-\infty}^{\infty} a(k) e^{i3\beta^0 k^2.} \right|^2.$$

Now expanding $Q'(\widehat{\boldsymbol{\theta}})$ around $\boldsymbol{\theta}^0$ using multivariate Taylor series expansion, we obtain

$$Q'(\widehat{\boldsymbol{\theta}}) = Q'(\boldsymbol{\theta}^0) + Q''(\bar{\boldsymbol{\theta}})(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0),$$

where $\bar{\boldsymbol{\theta}}$ lies on the line joining $\widehat{\boldsymbol{\theta}}$ and $\boldsymbol{\theta}^0$. Since $\widehat{\boldsymbol{\theta}}$ minimizes $Q(\boldsymbol{\theta})$, we have $Q'(\widehat{\boldsymbol{\theta}}) = \mathbf{0}$, therefore

$$\boldsymbol{D}Q'(\boldsymbol{\theta}^0) = -\boldsymbol{D}Q''(\bar{\boldsymbol{\theta}})\boldsymbol{D}\boldsymbol{D}^{-1}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0). \tag{25}$$

Theorem 1 implies that $\widehat{\boldsymbol{\theta}} \xrightarrow{a.s.} \boldsymbol{\theta}^0$. Therefore, using the continuous mapping theorem and repeated use of Lemmas C-1 and C-2, we observe that

$$\lim_{N\to\infty} \boldsymbol{D}Q^{''}(\bar{\theta})\boldsymbol{D} = \lim_{N\to\infty} \boldsymbol{D}Q^{''}(\theta^0)\boldsymbol{D} = \boldsymbol{G}, \tag{26}$$

where

$$\boldsymbol{G} = \begin{pmatrix} \boldsymbol{G}_1 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{G}_2 \end{pmatrix}.$$

with $\boldsymbol{G}_1$ and $\boldsymbol{G}_2$ same as defined in (11). Now using (24) and (26) in (25), we obtain

$$\boldsymbol{D}^{-1}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0) \xrightarrow{d} N_6(\mathbf{0}, \sigma^2\ \boldsymbol{\Sigma}^{-1}\boldsymbol{G}\boldsymbol{\Sigma}^{-1}).$$

Hence, the result follows. ∎

Appendix D

The second order derivatives of $Q(\boldsymbol{\theta})$ with respect to elements of $\boldsymbol{\theta}$ at $\boldsymbol{\theta}^0$ are provided in this Appendix.

$$\frac{\partial^2 Q(\boldsymbol{\theta}^0)}{\partial A^2} = 2\sum_{n=1}^{N} w\left(\frac{n}{N}\right)\cos^2(\alpha^0 n),$$

$$\frac{\partial^2 Q(\boldsymbol{\theta}^0)}{\partial A \partial B} = 2\sum_{n=1}^{N} w\left(\frac{n}{N}\right)\cos(\alpha^0 n)\sin(\alpha^0 n),$$

$$\frac{\partial^2 Q(\boldsymbol{\theta}^0)}{\partial A \partial \alpha} = 2\sum_{n=1}^{N} n\ w\left(\frac{n}{N}\right)(B^0\cos(\alpha^0 n) - A^0\sin(\alpha^0 n))\cos(\alpha^0 n) + 2\sum_{n=1}^{N} n\ w\left(\frac{n}{N}\right)X(n)\sin(\alpha^0 n),$$

$$\frac{\partial^2 Q(\boldsymbol{\theta}^0)}{\partial A \partial C} = 2\sum_{n=1}^{N} w\left(\frac{n}{N}\right)\cos(\alpha^0 n)\cos(\beta^0 n^2),$$

$$\frac{\partial^2 Q(\boldsymbol{\theta}^0)}{\partial A \partial D} = 2\sum_{n=1}^{N} w\left(\frac{n}{N}\right)\cos(\alpha^0 n)\sin(\beta^0 n^2),$$

$$\frac{\partial^2 Q(\boldsymbol{\theta}^0)}{\partial A \partial \beta} = 2\sum_{n=1}^{N} n^2 w\left(\frac{n}{N}\right)\cos(\alpha^0 n)(D^0\cos(\beta^0 n^2) - C^0\sin(\beta^0 n^2))$$

$$\frac{\partial^2 Q(\boldsymbol{\theta}^0)}{\partial B^2} = 2\sum_{n=1}^{N} w\left(\frac{n}{N}\right)\sin^2(\alpha^0 n),$$

$$\frac{\partial^2 Q(\boldsymbol{\theta}^0)}{\partial B \partial \alpha} = 2\sum_{n=1}^{N} n\ w\left(\frac{n}{N}\right)(B^0\cos(\alpha^0 n) - A^0\sin(\alpha^0 n))\sin(\alpha^0 n) - 2\sum_{n=1}^{N} n\ w\left(\frac{n}{N}\right)X(n)\cos(\alpha^0 n),$$

$$\frac{\partial^2 Q(\boldsymbol{\theta}^0)}{\partial B \partial C} = 2\sum_{n=1}^{N} w\left(\frac{n}{N}\right)\sin(\alpha^0 n)\cos(\beta^0 n^2),$$

$$\frac{\partial^2 Q(\boldsymbol{\theta}^0)}{\partial B \partial D} = 2\sum_{n=1}^{N} w\left(\frac{n}{N}\right)\sin(\alpha^0 n)\sin(\beta^0 n^2),$$

$$\frac{\partial^2 Q(\boldsymbol{\theta}^0)}{\partial B \partial \beta} = 2\sum_{n=1}^{N} n^2 w\left(\frac{n}{N}\right)\sin(\alpha^0 n)(D^0\cos(\beta^0 n^2) - C^0\sin(\beta^0 n^2))$$

$$\begin{aligned}\frac{\partial^2 Q(\boldsymbol{\theta}^0)}{\partial \alpha^2} &= 2\sum_{n=1}^{N} n^2 w\left(\frac{n}{N}\right)\left(B^0\cos(\alpha^0 n) - A^0\sin(\alpha^0 n)\right)^2 \\ &\quad +2\sum_{n=1}^{N} n^2 w\left(\frac{n}{N}\right)X(n)\left\{A^0\cos(\alpha^0 n) + B^0\sin(\alpha^0 n)\right\},\end{aligned}$$

$$
\begin{aligned}
\frac{\partial^2 Q(\boldsymbol{\theta}^0)}{\partial \alpha \partial C} &= 2\sum_{n=1}^{N} n\ w\left(\frac{n}{N}\right)\cos(\beta^0 n^2)(B^0\cos(\alpha^0 n) - A^0\sin(\alpha^0 n)),\\
\frac{\partial^2 Q(\boldsymbol{\theta}^0)}{\partial \alpha \partial D} &= 2\sum_{n=1}^{N} n\ w\left(\frac{n}{N}\right)\sin(\beta^0 n^2)(B^0\cos(\alpha^0 n) - A^0\sin(\alpha^0 n)),\\
\frac{\partial^2 Q(\boldsymbol{\theta}^0)}{\partial \alpha \partial \beta} &= 2\sum_{n=1}^{N} n^3 w\left(\frac{n}{N}\right)(B^0\cos(\alpha^0 n) - A^0\sin(\alpha^0 n))(D^0\cos(\beta^0 n^2) - C^0\sin(\beta^0 n^2))\\
\frac{\partial^2 Q(\boldsymbol{\theta}^0)}{\partial C^2} &= 2\sum_{n=1}^{N} w\left(\frac{n}{N}\right)\cos^2(\beta^0 n^2),\\
\frac{\partial^2 Q(\boldsymbol{\theta}^0)}{\partial C \partial D} &= 2\sum_{n=1}^{N} w\left(\frac{n}{N}\right)\cos(\beta^0 n^2)\sin(\beta^0 n^2),\\
\frac{\partial^2 Q(\boldsymbol{\theta}^0)}{\partial C \partial \beta} &= 2\sum_{n=1}^{N} n^2 w\left(\frac{n}{N}\right)(D^0\cos(\beta^0 n^2) - C^0\sin(\beta^0 n^2))\cos(\beta^0 n^2) + 2\sum_{n=1}^{N} nw\left(\frac{n}{N}\right)X(n)\sin(\beta^0 n^2),\\
\frac{\partial^2 Q(\boldsymbol{\theta}^0)}{\partial D^2} &= 2\sum_{n=1}^{N} w\left(\frac{n}{N}\right)\sin^2(\beta^0 n^2),\\
\frac{\partial^2 Q(\boldsymbol{\theta}^0)}{\partial D \partial \beta} &= 2\sum_{n=1}^{N} n^2 w\left(\frac{n}{N}\right)(D^0\cos(\beta^0 n^0) - C^0\sin(\beta^0 n^2))\sin(\beta^0 n^2) - 2\sum_{n=1}^{N} n^2 w\left(\frac{n}{N}\right)X(n)\cos(\beta^0 n^2),\\
\frac{\partial^2 Q(\boldsymbol{\theta}^0)}{\partial \beta^2} &= 2\sum_{n=1}^{N} n^4 w\left(\frac{n}{N}\right)\left(D^0\cos(\beta^0 n^2) - C^0\sin(\beta^0 n^2)\right)^2\\
&\quad +2\sum_{n=1}^{N} n^4 w\left(\frac{n}{N}\right)X(n)\left\{C^0\cos(\beta^0 n^2) + D^0\sin(\beta^0 n^2)\right\}.
\end{aligned}
$$

Now consider the $(1,1)^{th}$ element of $\mathbf{D}Q(\boldsymbol{\theta}^0)\mathbf{D}$ for large N.

$$
\begin{aligned}
\lim_{N\to\infty}\frac{1}{N}\frac{\partial^2 Q(\boldsymbol{\theta}^0)}{\partial A^2} &= \lim_{N\to\infty}\frac{2}{N}\sum_{n=1}^{N} w\left(\frac{n}{N}\right)\cos^2(\alpha^0 n)\\
&= \lim_{N\to\infty}\frac{2}{N}\sum_{n=1}^{N}\left(1 + a_1\frac{n}{N} + a_2\frac{n^2}{N^2} + \cdot + a_m\frac{n^m}{N^m}\right)\cos^2(\alpha^0 n)\\
&= (1 + \frac{a_1}{2} + \frac{a_2}{3} + \cdots + \frac{a_m}{m+1}) = c_1,
\end{aligned}
$$

using Result A.1. Here $c_{k+1}, k = 0, 1, \ldots$ are same as defined in (8). Similarly the other elements of $\lim\limits_{N\to\infty}\mathbf{D}Q(\boldsymbol{\theta}^0)\mathbf{D}$ can be obtained using Lemmas C-1 and C-2.

References

[1] Djuric, P.M., 1996. *A model selection rule for sinusoids in white Gaussian noise. IEEE Transactions on Signal Processing*, vol. 44(7), 1744–1751.

[2] Grover, R. 2020, *Frequency and frequency rate estimation of some non-stationary signal processing models*, Ph.D. thesis, Indian Institute of Technology, Kanpur.

[3] Grover, R., Kundu, D. and Mitra, A. (2021), "Asymptotic properties of least squares estimators and sequential least squares estimators of a chirp-like signal model parameters", *Circuits, Systems and Signal Processing*, vol. 40, no. 11, 5421 - 5467.

[4] Guillet de Chatellus, H., Romero Cortes, L., Schnebelin, C., Burla, M. and Azana, J. (2018), "Reconfigurable photonic generation of broadband chirped waveforms using a single CW laser and low-frequency electronics", . *Nature Communications*, 9, Article No. 2438. https://doi.org/10.1038/s41467-018-04822-4.

[5] Jennrich, R.I. (1969), "Asymptotic properties of the nonlinear least squares estimators", *Annals of Mathematical Statistics*, vol. 40, 633 – 643.

[6] Klauder, J.R., Price, A.C., Darlington, S., Albersheim, W.J. (1960), "The theory and design of chirp radars", *The Bell System Technical Journals*, vol. XXXIX, 745 – 808.

[7] Kundu, D. and Grover, R. (2021), "On a chirp-like model and its parameter estimation using periodogram-type estimators", *Journal of Statistical Theory and Practice*, vol.15, no. 2, Paper no. 37.

[8] Lahiri, A., Kundu, D. and Mitra, A. (2014), "On least absolute deviation estimator of one dimensional chirp model", *Statistics*, vol. 48, no. 2, 405 - 420.

[9] Lahiri, A., Kundu, D. and Mitra, A. (2015), "Estimating the parameters of multiple chirp signals", *Journal of Multivariate Analysis*, **139**, 189 - 205.

[10] Lancaster, D. (1965), "Chirp, a new radar technique", *Electronics World*, vol. 73, 42 – 53.

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

[12] Montgomery H.L. (1990), *Ten Lectures on the Interface Between Analytic Number Theory and Harmonic Analysis*, American Mathematical Society, 196.

[13] Prasad, A., Kundu, D and Mitra, A. (2008), "Sequential estimation of the sum of sinusoidal model parameters", *Journal of Statistical Planning and Inference*, vol. 138, no. 5, 1297 – 1313.

[14] Schock, S.G. (2004), "A method for estimating the physical and acoustic properties of the seabed using chirp sonar data", *IEEE Transactions on Oceanic Engineering.*, vol.29, 1200–1217.

[15] Schock, S.G. (2004), "Remote estimates of physical and acoustic sediment properties in the South China Sea using chirp sonar data and the Biot Model", *IEEE Transactions on Oceanic Engineering.*, vol.29, 1218–1230.

[16] Vinogradov, I.M. (1954), "The method of trigonometrical sums in the theory of numbers", *Interscience*, Translated from Russian. Revised and annotated by K.F. Roth and Anne Davenport. Reprint of the 1954 translation. Dover Publications,Inc., Mineola, NY, 2004.

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