On Weighted Least Squares Estimation of Elementary Chirp Model

Anjali Mittal, Debasis Kundu, and Amit Mitra

***Abstract*—In this paper, we consider the estimation of the parameters of multiple component elementary chirp model based on weighted least squares estimators (WLSEs). Recently, Mittal et al. [17] proposed least squares estimators (LSEs) to estimate the parameters of the elementary chirp model. But it has been observed that LSEs are quite sensitive to outliers. The least absolute deviation estimators and Huber's M-estimators are the robust estimators for estimation in the presence of outliers. However, computing these robust estimators is numerically challenging, and to derive their statistical properties is not straightforward under the same error assumption as the LSEs. We prove the strong consistency and asymptotic normality of the WLSEs under the same error assumptions as those of the LSEs. We have also considered sequential WLSEs to estimate the parameters, which are less computationally involved and have the same asymptotic properties as the WLSEs. Based on the extensive numerical studies, it has been observed that the behaviour of the proposed estimators is better than LSEs in the presence of the outliers. We also discuss the choice of weight function, as the performance of the proposed estimators depends on the weight function. For illustration, we have presented a simulation study, where the outliers lie randomly in the data, not only in a specific portion. We have also analyzed a synthetic and a real data set. The performance is quite satisfactory.**

***Index Terms*—Frequency rate, elementary chirp, weighted least squares, strong consistency, asymptotic normality, sequential weighted least squares.**

I. Introduction

In this paper, we consider the multiple component elementary chirp model with an additive error as follows:

$$y(t) = \sum_{k=1}^{p} A_k^0 e^{i\beta_k^0 t^2} + \epsilon(t);\ t = 1, \ldots, N, \qquad (1)$$

where, $i = \sqrt{-1}$, A_k^0's are the amplitude parameters, β_k^0's are the frequency rate parameters and $\epsilon(t)$'s are independent and identically distributed (i.i.d.) the complex-valued error random variables, which satisfy assumption 2. Here, p is the number of chirp components in the model which is assumed to be known. The problem here is to estimate the unknown parameters A_k^0's and β_k^0's for the observed sample of size N.

A. Mittal, D. Kundu and A. Mitra are with Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Kanpur - 208016, India.

The elementary chirp model has a wide range of applications in signal processing. The estimation of frequency rate, which is widely known as chirp rate estimation in signal processing literature, is quite important here. For instance, their applications can be seen in synchrosqueezing [10], in mode reconstruction [15], in radar [3], [23], in sonar [2], in speech modeling [21] and in several other areas. In the literature, chirp rate estimation based on least squares estimators (LSEs) and approximate least squares estimators (ALSEs) [17], efficient methodology [18], convolution neural network [24], discrete polynomial transform and singular value decomposition [22], discrete polynomial transform and weighted combination [7], discrete chirp-Fourier transform [29], modified discrete chirp-Fourier transform [26], Fourier transform [16], fractional Fourier transform [4], [5], time frequency rate distribution [25] and machine learning [30] have been discussed. Estimators based on cubic phase function [19], cubic phase transform [8], product cubic phase function [28], adaptive algorithm [11] and integrated cubic phase function [27] are also being mentioned in the literature for estimating the instantaneous frequency rate, i.e., twice the frequency rate.

There are some situations where outliers can be found in the observed signal data. To name a few, outliers may occur in radar and sonar systems as a result of electromagnetic [1] and acoustic interference [9], in magnetic resonance imaging [6], in signal measurement [35] and in various other areas; please see [36] and references cited therein.

Recently, LSEs [17] and efficient estimators (EFFEs) based on a finite step iterative procedure [18] have been proposed to estimate the elementary chirp model parameters (1). It has been established that LSEs and EFFEs are consistent and asymptotically normally distributed under a fairly general set of error assumptions, see, for example, Assumption 2 in Section II. The LSEs are used to estimate the parameters of elementary chirp model, which have desired statistical properties, but they are sensitive to outliers present in the data. In this situation, when outliers are present in the data, robust estimators like least absolute deviation estimators (LADEs) and Huber's M-estimators (HMEs) are preferred for estimating the model parameters. To prove the statistical properties of these robust estimators, we need stronger assumptions on error than those required for the LSEs. For illustration, please see the work by Lahiri et al. [13] on LADEs for one component chirp model. Moreover, to establish the properties of LADEs for multiple component model is even more challenging and it has not been investigated yet. The theoretical properties of HMEs are not available even for one component model. Further, to compute LADEs and HMEs numerically, we have to solve a $3p$-dimensional optimization. This optimization becomes numerically challenging for large value of p.

Our aim is to propose estimators which are easy to compute, robust to the outliers, and have a rate of convergence identical to the LSEs. Here we propose weighted least squares estimators (WLSEs) of the multiple component elementary chirp model (1) parameters and derive their asymptotic properties under the same set of assumptions as that are needed for the LSEs. It is observed that WLSEs are robust to the outliers than the LSEs and their performance is similar to LADEs and HMEs. The WLSEs are strongly consistent and have a rate of convergence identical to the LSEs. To find them numerically, we have to do a p-D optimization. To make this problem less computationally expensive, we propose sequential WLSEs for elementary chirp model with p-components (1), where we solve p, 1-D optimizations. We also prove that the sequential WLSEs are strongly consistent and have an asymptotic distribution identical to the WLSEs.

Next, we do extensive simulations to study the performance of the proposed estimators for several combinations of parameters. We compare their performance with LSEs, EFFEs, LADEs, HMEs and cubic phase function based estimators (CPFEs) for one component model (2). It has been reported that the performance of the WLSEs is better than LSEs, EFFEs, and CPFEs and is at par with LADEs and HMEs. The performance of sequential WLSEs has been compared with sequential LSEs, sequential EFFEs, LADEs, HMEs, and product cubic phase function based estimators (PCPFEs) for the multiple component model (1). It has been observed that the performance of the sequential WLSEs is better than sequential LSEs, sequential EFFEs, and PCPFEs and is at par with LADEs and HMEs. For the multiple component model, we have done simulations for the case when frequency rates are close to each other. In the above set of simulations, we contaminated the data from a specific portion to emphasize the fact that how the performance proposed estimators depends on the choice of weight function. Further, we provide a method to choose an appropriate weight function along with a simulation experiment. In this simulation study, we have contaminated the data with outliers from the entire portion, not only from a specific portion, by generating them randomly using complex-valued mixture normal error random variables, as done in [34]. We have also done a synthetic data analysis and a real data analysis based on the method provided for the choice of weight function. We have analyzed the real data using the proposed method by fitting an elementary chirp model to it. In real data analysis presented in Section VII, we need $p = 27$, and in that case, for LADEs or HMEs, we have to solve a 81-D optimization problem. Also, to solve a 81-D optimization problem, we need the same number of the initial values, which makes the problem even more difficult to solve in practice. Hence, it has not been attempted here. Whereas in the proposed sequential WLSEs, one needs to solve 27, 1-D optimization problem. It is easy to compute the initial value for the 1-D optimization problem. Therefore, it is apparent that the proposed method provides a practically feasible solution to the real world problems.

The key contributions of this paper are mentioned as follows: In this paper, we put forward weighted least squares estimation method to estimate the parameters of the elementary chirp model (1). We have proved strong consistency and asymptotic normality of these proposed estimators, namely WLSEs and sequential WLSEs, under the same set of assumptions as those required for LSEs. The proposed estimators are robust to the outliers than LSEs. The proposed estimators are less computationally expensive, less time consuming and easier to implement in practice compared to the other robust estimators like HMEs and LADEs. The performance of the proposed estimators is at par with HMEs and LADEs. Moreover, the proposed sequential WLSEs are convenient to use in practice. Also, the asymptotic properties of the proposed estimators are easily proved. Therefore, it appears that we have offered a thorough and comprehensive solution to this particular problem.

The organization of rest of the paper is given as follows. In Section II, we present WLSEs for one component elementary chirp model (2) and study their asymptotic properties. In Section III, we present WLSEs and sequential WLSEs for the multiple component model (1) and study their statistical properties. Numerical results are presented in Section IV. Discussion on the choice of weight function along with a simulation experiment is presented in Section V. In Section VI, we present a synthetic data analysis. In Section VII, we present analysis on a real data. In Section VIII, we conclude the paper. In the appendices, we provide proofs of all the results.

II. One Component Elementary Chirp Model

In this section, we consider the following one component elementary chirp model:

$$y(t) = A^0 e^{i\beta^0 t^2} + \epsilon(t); \quad t = 1, \ldots, N, \qquad (2)$$

where, A^0 is the complex-valued non-zero amplitude parameter. Here, β^0 is the frequency rate with $\beta^0 \in (0, 2\pi)$. Also, $\epsilon(t)$'s are the error random variables inherent in the observed signal $y(t)$. The problem is to estimate the parameters A^0 and β^0.

The following are the notations that we use in this paper: $A = A_R + iA_I$, where, A_R and A_I denote the real and the imaginary parts of A, respectively. Similarly, $\epsilon(t) = \epsilon_R(t) + i\epsilon_I(t)$, where $\epsilon_R(t)$ and $\epsilon_I(t)$ denote the real and the imaginary parts of $\epsilon(t)$, respectively. The parameter vector $\boldsymbol{\theta}$ is denoted by (A_R, A_I, β). Also, $\boldsymbol{\theta}^0$ denotes the true parameter vector and $\breve{\boldsymbol{\theta}}$ is the WLSEs of $\boldsymbol{\theta}^0$.

In the following subsection, we provide WLSEs for one component elementary chirp model (2). The explicit use of the weight function $w(s)$ will be shown in Subsection II-A. We also provide statistical properties of WLSEs under the following assumptions on the weight function $w(s)$ and the error random variables.

Assumption 1: The weight function $w(s)$ is a non-negative continuous function defined on $[0, 1]$, such that $\min\limits_{0 \le s \le 1} w(s) > \gamma > 0$.

Assumption 2: The error random variables $\epsilon(t)$'s are complex-valued i.i.d. random variables with mean 0 and variance $\frac{\sigma^2}{2}$ for both real and imaginary parts. The real and imaginary parts of $\epsilon(t)$ are independently distributed. The fourth order moment of $\epsilon(t)$ exists. Please observe that if real and imaginary parts of $\epsilon(t)$'s are real-valued normal random variables then $\epsilon(t)$'s would be circularly symmetric complex-valued normal random variables, i.e., $\epsilon(t) \sim \mathcal{CN}(0, \sigma^2)$ $\forall t = 1, \ldots, N$. Mathematically the expression $X \sim \mathcal{CN}(0, \sigma^2)$ denotes that X follows circularly symmetric complex-valued normal distribution with mean 0 and variance σ^2. and the probability density function (PDF) of a circularly symmetric complex-valued normal random variable is given as follows:

$$f_X(x) = \frac{1}{\pi\sigma^2} \exp\left(-\frac{|x|^2}{\sigma^2}\right), x \in \mathbb{C}. \tag{3}$$

Here, $\mathbb{C}$ denotes complex plane. Please see the book by Ladipoth [14] for more details on complex-valued random variables.

A. Weighted Least Squares Estimators

Now we discuss WLSEs for the elementary chirp model (2), which are useful in the estimation in the presence of outliers. The WLSEs of the parameters of the model (2) are determined by minimizing the following weighted residual sum of squares, say:

$$Q(\boldsymbol{\theta}) = \sum_{t=1}^{N} w\left(\frac{t}{N}\right) \left| y(t) - Ae^{i\beta t^2} \right|^2, \tag{4}$$

with respect to A and β simultaneously. For fixed β, WLSEs of A_R^0 and A_I^0 can be obtained as follows:

$$\breve{A}_R(\beta) = \frac{a_1}{b_1} \quad \text{and} \quad \breve{A}_I(\beta) = \frac{a_2}{b_1}, \quad \text{where,} \tag{5}$$

$$a_1 = \sum_{t=1}^{N} w\left(\frac{t}{N}\right)\left(y_R(t)\cos(\beta t^2) + y_I(t)\sin(\beta t^2)\right),$$

$$a_2 = \sum_{t=1}^{N} w\left(\frac{t}{N}\right)\left(y_I(t)\cos(\beta t^2) - y_R(t)\sin(\beta t^2)\right),$$

$$b_1 = \sum_{t=1}^{N} w\left(\frac{t}{N}\right)\left(\cos^2(\beta t^2) + \sin^2(\beta t^2)\right).$$

Once $\breve{A}_R(\beta)$ and $\breve{A}_I(\beta)$ are obtained, WLSE of β^0 can be determined by minimizing the following function:

$$Q\left(\breve{A}_R(\beta), \breve{A}_I(\beta), \beta\right) = \sum_{t=1}^{N} w\left(\frac{t}{N}\right) \left| y(t) - \breve{A}e^{i\beta t^2} \right|^2,$$

with respect to β. Then the WLSE of $\breve{A}$ is obtained by replacing β by its WLSE $\breve{\beta}$ in (5). Here, note that for $w(s) = 1$, (4), provides LSEs of $\boldsymbol{\theta}^0$ as a special case of WLSEs. Next, we provide the results of the strong consistency and the asymptotic normality of the WLSEs in the following theorems:

Theorem 1: If assumptions 1 and 2 hold true, then $\breve{\boldsymbol{\theta}}$ is a strongly consistent estimator of $\boldsymbol{\theta}^0$, i.e.,

$$\breve{\boldsymbol{\theta}} \xrightarrow{a.s.} \boldsymbol{\theta}^0 \text{ as } N \to \infty.$$

Here $\xrightarrow{a.s.}$ stands for almost sure convergence.

Proof: See Appendix B. ∎

Theorem 2: If assumptions 1 and 2 are satisfied, then $(\breve{\boldsymbol{\theta}} - \boldsymbol{\theta}^0)\boldsymbol{D}^{-1} \xrightarrow{d} \mathcal{N}_3\left(\mathbf{0}, \sigma^2\boldsymbol{G}^{-1}\boldsymbol{\Sigma}\boldsymbol{G}^{-1}\right)$ as $N \to \infty$, where $\boldsymbol{D} = diag\left(\frac{1}{\sqrt{N}}, \frac{1}{\sqrt{N}}, \frac{1}{N^2\sqrt{N}}\right)$, $\mathcal{N}_p$ denotes p-dimensional normal distribution and

$$\boldsymbol{G} = 2\begin{bmatrix} c_1 & 0 & -A_I^0 c_3 \\ 0 & c_1 & A_R^0 c_3 \\ -A_I^0 c_3 & A_R^0 c_3 & |A^0|^2 c_5 \end{bmatrix},$$
$$\boldsymbol{\Sigma} = 2\begin{bmatrix} d_1 & 0 & -A_I^0 d_3 \\ 0 & d_1 & A_R^0 d_3 \\ -A_I^0 d_3 & A_R^0 d_3 & |A^0|^2 d_5 \end{bmatrix}, \text{ with} \tag{6}$$

$$\lim_{N\to\infty} \frac{1}{N^{k+1}} \sum_{t=1}^{N} t^k w\left(\frac{t}{N}\right) = \int_0^1 s^k w(s) ds = c_{k+1} > 0, \tag{7}$$

$$\lim_{N\to\infty} \frac{1}{N^{k+1}} \sum_{t=1}^{N} t^k w^2\left(\frac{t}{N}\right) = \int_0^1 s^k w^2(s) ds = d_{k+1} > 0. \tag{8}$$

Here $\xrightarrow{d}$ stands for convergence in distribution.

Proof: See Appendix B. ∎

Note that for $w(s) = 1$, $s \in [0, 1]$, the asymptotic distribution of WLSEs of $\boldsymbol{\theta}^0$ is same as the LSEs of

$\boldsymbol{\theta}^0$ [17]. Also note that for large N, (5) becomes

$$\breve{A}_R(\beta) = \frac{a_1}{Nc_1} \quad \text{and} \quad \breve{A}_I(\beta) = \frac{a_2}{Nc_1}, \tag{9}$$

using Lemma C-1 of [12]. The WLSEs of amplitude parameters are $\breve{A}_R\left(\breve{\beta}\right)$ and $\breve{A}_I\left(\breve{\beta}\right)$, once $\breve{\beta}$ is obtained. These are equivalent to the ALSEs of the amplitude parameters of A_R^0 and A_I^0 for $w(s)=1$.

III. Multiple Component Elementary Chirp Model

In the following subsection, we present the natural extension of WLSEs for the multiple component elementary chirp model (1). Without loss of generality, we also assume the following relationship among A_k^0's

$$2M^2 > \left|A_1^0\right|^2 > \left|A_2^0\right|^2 > \cdots > \left|A_p^0\right|^2 > 0.$$

A. Weighted Least Squares Estimators

The WLSEs of the unknown parameters of the multiple component elementary chirp model (1) can be obtained by minimizing the following weighted residual sum of squares:

$$Q(\boldsymbol{u}) = \sum_{t=1}^{N} w\left(\frac{t}{N}\right)\left|y(t) - \sum_{k=1}^{p} A_k e^{i\beta_k t^2}\right|^2, \tag{10}$$

with respect to $\boldsymbol{u}$ where, $\boldsymbol{u} = (A_{R1}, A_{I1}, \beta_1, \ldots, A_{Rp}, A_{Ip}, \beta_p)$. Here, $\boldsymbol{u}^0$ is the true parameter vector and $\breve{\boldsymbol{u}}$ is the WLSEs of $\boldsymbol{u}^0$. Now the WLSEs, $\breve{\boldsymbol{u}}$ of $\boldsymbol{u}^0$, can be obtained similarly as obtained for the one component elementary chirp model in Subsection II-A. Thus, to obtain WLSEs will be a p-D optimization problem here.

The following assumption on the parameters is required to prove the asymptotic properties, namely, strong consistency and asymptotic distribution, of the WLSEs.

Assumption 3: Let $\boldsymbol{u}^0$ be an interior point of the parameter space $\boldsymbol{\mathcal{U}} = \boldsymbol{\Theta}^{(p)}$; $\boldsymbol{\Theta} = [-L, L] \times [-L, L] \times [0, 2\pi]$ and $\beta_k^0 \neq \beta_j^0\ \forall\ k \neq j$; for $k,\ j = 1, \cdots, p$.

Theorem 3: If assumptions 1, 2 and 3 are satisfied, then

$$\breve{\boldsymbol{u}} \xrightarrow{a.s.} \boldsymbol{u}^0 \text{ as } N \to \infty.$$

Proof: See Appendix C-A. ■

Theorem 4: If assumptions 1, 2 and 3 are satisfied, then

$$(\breve{\boldsymbol{u}} - \boldsymbol{u}^0)\mathscr{D}^{-1} \xrightarrow{d} \mathcal{N}_{3p}\left(0, \sigma^2\boldsymbol{\mathcal{H}}\right) \text{ as } N \to \infty.$$

Here, $\mathscr{D} = diag(\underbrace{\boldsymbol{D}, \ldots, \boldsymbol{D}}_{p\ times})$, where $\boldsymbol{D}$ is same as defined in Theorem 2 and $\boldsymbol{\mathcal{H}}$ is a $3p \times 3p$ block diagonal matrix whose k^{th} block is $\boldsymbol{G}_k^{-1}\boldsymbol{\Sigma}_k\boldsymbol{G}_k^{-1}$. Here, $\boldsymbol{G}_k$ and $\boldsymbol{\Sigma}_k$ can be obtained by replacing A_R^0, A_I^0, β^0 with $A_{Rk}^0, A_{Ik}^0, \beta_k^0$ in $\boldsymbol{G}$ and $\boldsymbol{\Sigma}$, respectively, as defined in (6).

Proof: See Appendix C-A. ■

Here we have seen that WLSEs are strongly consistent and follow normal distribution asymptotically, but to obtain WLSEs we have to solve a p-D optimization problem and the inverse of a $2p \times 2p$ matrix, which would be computationally challenging for large p. This problem can be reduced to p, 1-D optimization problems and solving the inverse of a 2×2 matrix p times using the sequential method. In the next subsection, we propose a sequential technique to estimate the model (1) parameters, which calculates the estimators with lower computational complexity while having the same asymptotic properties as the WLSEs.

B. Sequential Weighted Least Squares Estimators

Let $\tilde{\boldsymbol{u}}$ is the sequential WLSEs of $\boldsymbol{u}^0$. In this procedure, we estimate the components of the model (1) sequentially. The procedure for the sequential technique is provided as follows:

Step 1: Estimate the parameters of the first component, i.e., $\tilde{\boldsymbol{\theta}}_1 = \left(\tilde{A}_{R1}, \tilde{A}_{I1}, \tilde{\beta}_1\right)$ of the model (1), using the method described in Subsection II-A.

Step 2: Obtain the new data vector by eliminating the estimated component:

$$y_2(t) = y_1(t) - \tilde{A}_1 e^{i\tilde{\beta}_1 t^2};\ y_1(t) = y(t).$$

Step 3: To estimate $\tilde{\boldsymbol{\theta}}_2 = \left(\tilde{A}_{R2}, \tilde{A}_{I2}, \tilde{\beta}_2\right)$, minimize the following weighted residual sum of squares:

$$Q_2(\boldsymbol{\theta}) = \sum_{t=1}^{N} w\left(\frac{t}{N}\right)\left|y_2(t) - Ae^{i\beta t^2}\right|^2.$$

Continue doing so until all p-components are estimated.

The results of the strong consistency of the proposed sequential WLSEs are provided as follows:

Theorem 5: If assumptions 1, 2 and 3 hold true, then

$$\tilde{\boldsymbol{\theta}}_1 \xrightarrow{a.s.} \boldsymbol{\theta}_1^0 \text{ as } N \to \infty.$$

Proof: See Appendix C-B. ■

Theorem 6: If assumptions 1, 2 and 3 hold true, then

$$\tilde{\boldsymbol{\theta}}_2 \xrightarrow{a.s.} \boldsymbol{\theta}_2^0 \text{ as } N \to \infty.$$

Proof: See Appendix C-B. ■

This result can be further extended for $k \leq p$.

Theorem 7: If assumptions 1, 2 and 3 hold true, then

$$\tilde{\boldsymbol{\theta}}_k \xrightarrow{a.s.} \boldsymbol{\theta}_k^0 \text{ as } N \to \infty, \forall k \leq p.$$

Proof: It follows similarly as Theorem 6. ■

Theorem 8: If assumptions 1, 2 and 3 hold true, then

$$\tilde{A}_{R(p+1)} \xrightarrow{a.s.} 0,\ \tilde{A}_{I(p+1)} \xrightarrow{a.s.} 0 \text{ as } N \to \infty.$$

Proof: See Appendix C-B. ■

The following theorem provides the asymptotic distribution of the proposed sequential WLSEs:

Theorem 9: If assumptions 1, 2 and 3 hold true, then

$$(\tilde{\boldsymbol{\theta}}_k - \boldsymbol{\theta}_k^0)\boldsymbol{D}^{-1} \xrightarrow{d} \mathcal{N}_3\left(0, \sigma^2\boldsymbol{G}_k^{-1}\boldsymbol{\Sigma}_k\boldsymbol{G}_k^{-1}\right) \text{ as } N \to \infty, \forall k \leq p,$$

where $\boldsymbol{D}$ is same as defined in Theorem 2 and $\boldsymbol{G}_k$ and $\boldsymbol{\Sigma}_k$ are same as defined in Theorem 4.

Proof: See Appendix C-B. ■

From these results, it is evident that the sequential WLSEs are strongly consistent and asymptotically equivalent to WLSEs due to their identical asymptotic distribution. Also, to obtain sequential WLSEs, we have to do 1-D optimization at every step, which is less computationally involved than calculating WLSEs, specially when we have large number of components in the model. Therefore, we can use sequential WLSEs for the implementation purposes.

IV. Numerical Results

In this section, we illustrate simulation results for a one component and a multiple component elementary chirp model. We do these simulations for various choices of weight function, sample size N, error variance σ^2, percentage of outliers, and error variance of outliers σ_{out}^2 to analyze the performance of the proposed weighted estimators in presence of the outliers in the data. The form of the error structure is explicitly mentioned in (11). Also, the performance of the proposed estimators is compared with the prominent existing methods in the literature. Through the numerical simulations presented in this section, our aim is to emphasize the importance of the appropriate weight function for the proposed methodology. Thus, we contaminate the data from a specific portion of the data and perform the experiments. Whereas, in the next section, we do a simulation experiment, where the data is contaminated randomly from the entire portion.

A. One Component Elementary Chirp Model

We consider model (2) with parameters $A^0 = 6$, $\beta^0 = 0.1$, $N = 101$ to 1001 and $\sigma^2 = 0.1, 0.25, 0.5, 0.75, 1, 2$. We have generated the data with outliers, contaminating $10, 20, 30$ and 40 percent of the middle part of the data. The indices of the middle part of the data are obtained as:

$$\frac{N}{2}\left(1 - p_{out}\right) + 1 : \frac{N}{2}\left(1 + p_{out}\right),$$

where, p_{out} denotes the percentage of outliers in the data. Here, $\epsilon\left(t\right)$ has the following form:

Structure :

$$\epsilon\left(t\right) \sim \begin{cases} \mathcal{CN}\left(0, 50\right) & \text{if} \quad t_1 \leq t \leq t_2 \\ \mathcal{CN}\left(0, \sigma^2\right) & \text{Otherwise,} \end{cases} \tag{11}$$

where, $t_1 = \frac{N}{2}\left(1 - p_{out}\right) + 1$ and $t_2 = \frac{N}{2}\left(1 + p_{out}\right)$.

We repeat the experiment 1000 times for each combination of above parameters. We estimate the parameters using five WLSEs discussed in subsection II-A, LADE, Huber's M-estimator (HME), LSE [17], EFFE [18] and CPFE [19] method. The five WLSEs

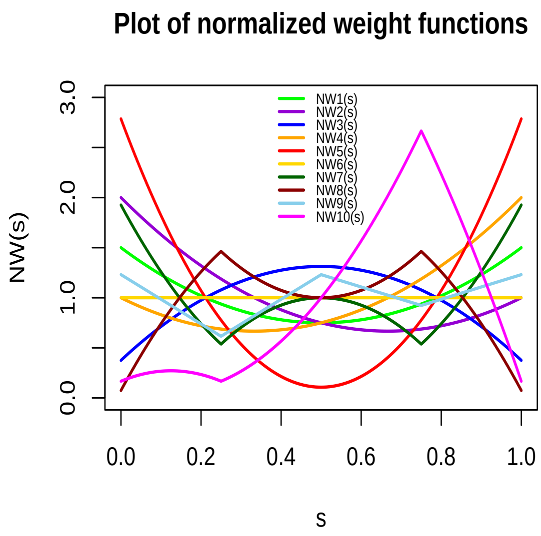


Fig. 1: Plot of normalized weight functions.

are computed based on the following weight functions: $w_1(s) = \frac{(1-s)^2}{2} + \frac{s^2}{2}$, $w_2(s) = \frac{2(1-s)^2}{3} + \frac{s^2}{3}$, $w_3(s) = -s^2 + s + \frac{1}{10}$, $w_4(s) = \frac{2s^2}{3} + \frac{(1-s)^2}{3}$ and $w_5(s) = s^2 - s + 0.26, s \in [0, 1]$. Fig. 1 shows the plot of these weight functions in their normalized form. The normalized weight (NW) function is obtained as follows:

$$NW(s) = \frac{w\left(s\right)}{\int_0^1 w\left(s\right)}; \; s \in \left[0, 1\right]. \tag{12}$$

We take a class of weight functions to see that how crucial it is to choose the appropriate weight function. In this class, w_1 and w_5 give less weight in the middle of the data set, w_2 gives less weight on the right side of the data set, w_3 gives less weight on the edges of the data set and w_4 gives less weight on the left side of the data set than the other regions. The weight function which provides less weight to regions of the data set where outliers are present, should perform better than the other weight functions. To see this practically, we take variously shaped weight functions. Please note that in this section, we have contaminated the data from the middle portion of the data to show the importance of the weight function in the performance of the proposed estimators. The choice of weight function is discussed in more detail in the next Section V along with a simulation experiment where we have contaminated the data with outliers from the entire portion, not only from a specific portion. We compute mean squared errors (MSEs) of the frequency rate for all these methods. We report $-10\log_{10}(MSE)$ for the best and worst perfmoring WLSEs, which are the W5LSE and the W3LSE, respectively, and LADE, HME, LSE, EFFE and CPFE. Here, W5LSE performs best in terms of MSE among five WLSEs as w_5 gives the least weight to the outliers whereas w_3 gives the greatest weight to the outliers therefore, W3LSE performs worst. We compute all these estimates of the frequency rate using the Nelder-Mead Simplex algorithm. We take true parameter values as the initial values in this Simplex algorithm. Observe that we have to do a 1-dimensional optimization to compute WLSE, LSE, and CPFE whereas we have to do a 3-dimensional optimization to obtain LADE and HME. Fig. 2 depicts $-10\log_{10}(MSE)$ of estimates of frequency rate versus sample size for various percentage of outliers in the data.

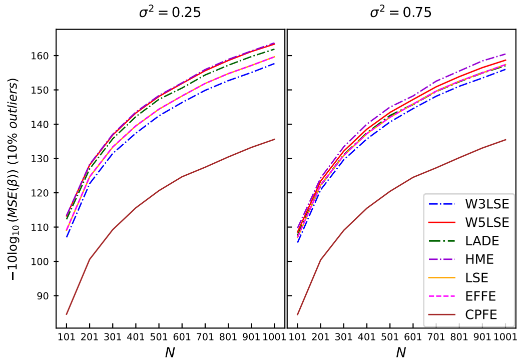
The following are the observations from this figure:

1) As N increases, $-10\log_{10}(MSE)$ increases, this validates the consistency of these estimators.
2) As the error variance, i.e., σ^2 increases $-10\log_{10}(MSE)$ of all the estimators decreases, i.e., as the signal-to-noise ratio (SNR) increases $-10\log_{10}(MSE)$ of all the estimators increases.
3) The $-10\log_{10}(MSE)$ of W5LSE, LADE and HME are greater than W3LSE, LSE, EFFE and CPFE.
4) The $-10\log_{10}(MSE)$ of W5LSE, LADE and HME are close to each other.
5) Therefore, if one chooses an appropriate weight function then WLSEs provide estimators with less computational complexity than the other discussed estimators.

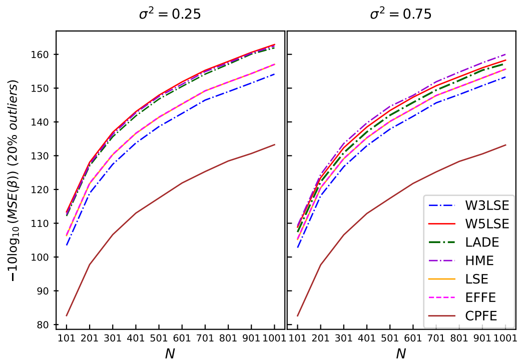
B. Multiple Component Elementary Chirp Model

We consider model (1) with $p = 2$ and $A_1^0 = 8, \beta_1^0 = 0.8, A_2^0 = 6, \beta_2^0 = 0.3$. We do simulation study for this model for different N, σ^2, percentage of outliers, and σ^2_{out}, same as those taken for simulation study in previous subsection IV-A. We estimate the model parameters using the WLSEs discussed in subsection III-A, sequential WLSEs discussed in subsection III-B, LADE, HME, sequential LSEs [17], sequential EFFEs [18], and PCPFE [28]. We computed these estimates using the Nelder-Mead Simplex algorithm. It is important to note that we have to do two 1-dimensional optimization to obtain sequential WLSEs, sequential LSEs, and PCPFE and a 2-dimensional optimization to obtain WLSEs whereas we have to do a 6-dimensional optimization to obtain LADE and HME. This tells about the lower computational complexity of the sequential WLSEs than the other discussed methods. We also compute the MSE of the frequency rates for these methods based on 1000 repetition. We report $-10\log_{10}(MSE)$ of frequency rate estimates in Figs. 3 and 4.

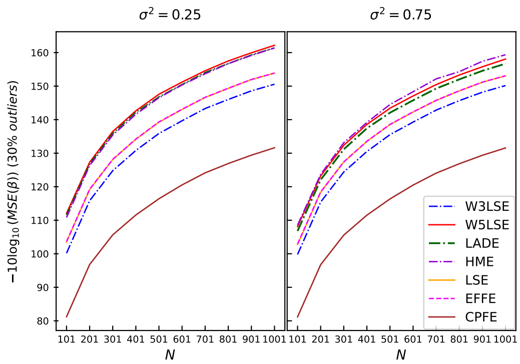
Following are the observations from these figures:



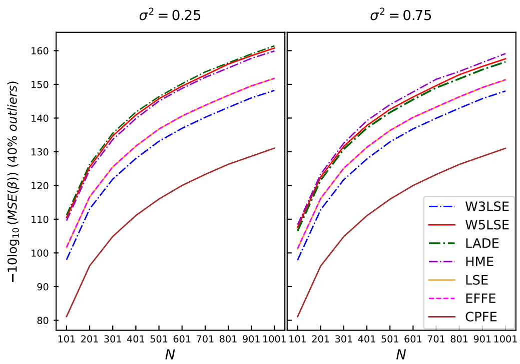
(a)



(b)

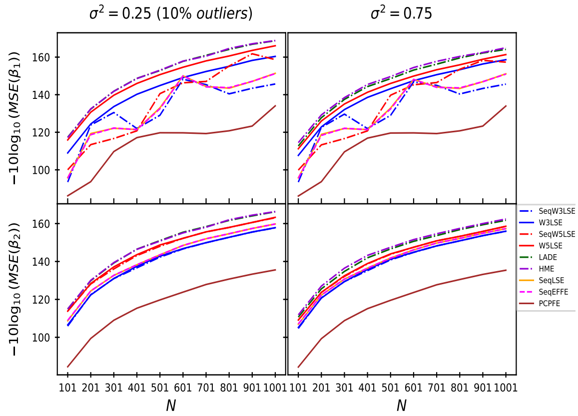


(c)

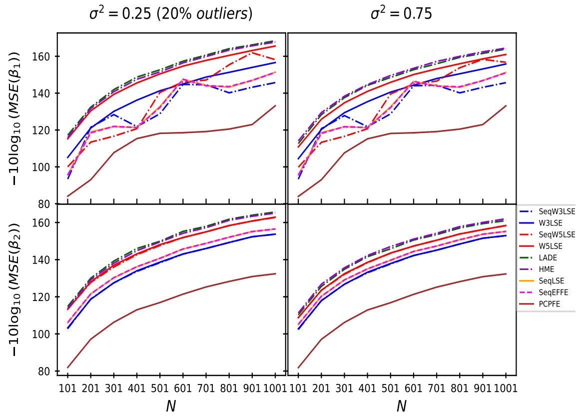


(d)

Fig. 2: $-10\log_{10}(MSE)$ of frequency rate estimates of the model (2) with $A^0 = 6$, $\beta^0 = 0.1$ and (a) 10% (b) 20% (c) 30% (d) 40% outliers with $\sigma^2_{out} = 50$ for various error variances versus sample size.

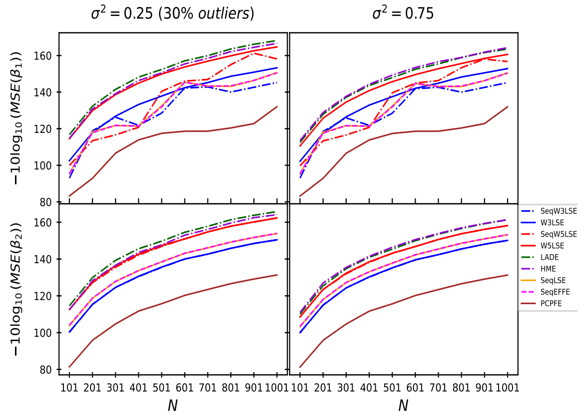


(a)

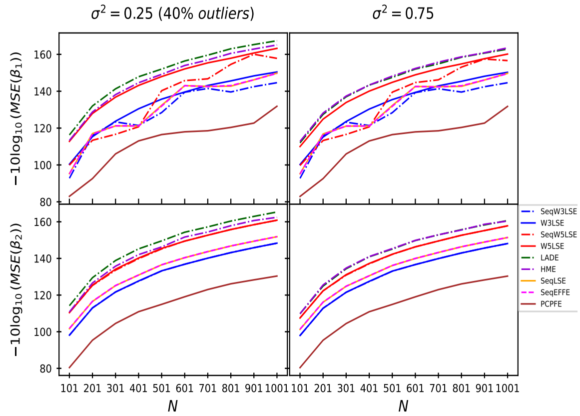


(b)

Fig. 3: $-10\log_{10}(MSE)$ of frequency rates estimates of the model (1) with $A_1^0 = 8$, $\beta_1^0 = 0.8$, $A_2^0 = 6$, $\beta_2^0 = 0.3$ and (a) 10% (b) 20% outliers with $\sigma_{out}^2 = 50$ for different error variances versus sample size.



(a)



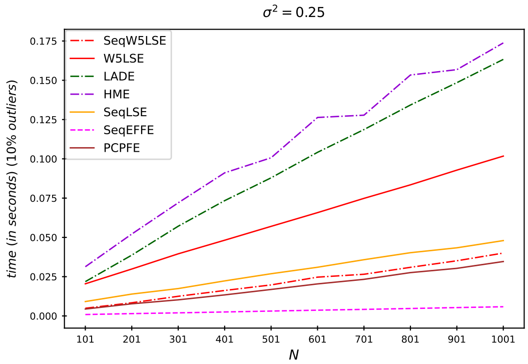
(b)

Fig. 4: $-10\log_{10}(MSE)$ of frequency rates estimates of the model (1) with $A_1^0 = 8$, $\beta_1^0 = 0.8$, $A_2^0 = 6$, $\beta_2^0 = 0.3$ and (a) 30% (b) 40% outliers with $\sigma_{out}^2 = 50$ for different error variances versus sample size.

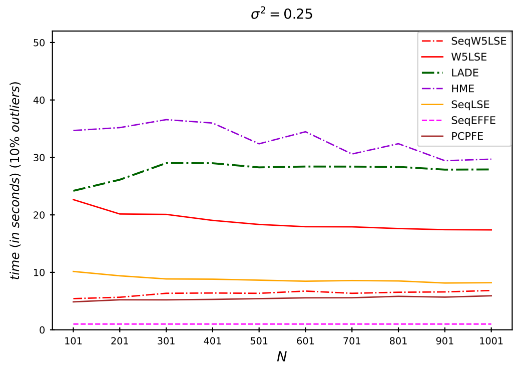
1) As N increases, $-10\log_{10}(MSE)$ increases in most of the cases.
2) As σ^2 increases $-10\log_{10}(MSE)$ of all the estimaors decreases.
3) The $-10\log_{10}(MSE)$ of sequential W5LSE, LADE and HME are close to each other and greater than W5LSE, sequential W3LSE, sequential LSE, sequential EFFE and PCPFE for the higher number of outliers.
4) Hence, by choosing an appropriate weight function sequential WLSEs provide estimators with less computational complexity than the other discussed estimators for multiple component model.

C. Time Comparison

In this subsection we provide time comparison in computing the estimates of the frequency rates and amplitude parameters based on the discussed estimation methods in this paper. We consider model (1) with $p = 3$, and parameters $A_1^0 = 12$, $\beta_1^0 = 3$, $A_2^0 = 10$, $\beta_2^0 = 2.5$, $A_3^0 = 8$, $\beta_3^0 = 2$, $\sigma^2 = 0.25$ and 10% outliers with $\sigma_{out}^2 = 50$ and $N = 101$ to 1001. Structure 1 has been considered for error random variables. We report computation time in estimating the associated frequency rates and amplitude parameters versus sample size in Fig. 5 (a) using various discussed methods. In Fig. 5 (b), we represent the time ratio with respect to the computation time involved in sequential EEFE method. From here, it is evident that the HME takes the greatest computing time, whereas sequential EFFE takes the least time as compared with the other discussed techniques. Also it is interesting to note that the computing time taken by sequential WLSE is lesser than the HME, LADE and WLSE. From here we can observe that the sequential WLSE with the appropriate weight function performs well in terms of



(a)



(b)

Fig. 5: (a) Computational time and (b) Computational time ratio in estimating the frequency rates and amplitude parameters using various estimation techniques for the model (1) with $A_1^0 = 12$, $\beta_1^0 = 3$, $A_2^0 = 10$, $\beta_2^0 = 2.5$, $A_3^0 = 8$, $\beta_3^0 = 2$ and 10% outliers with $\sigma_{out}^2 = 50$ versus sample size.

MSE and also take lesser time in computation with lower complexities than HME, LADE and WLSE and also has desired asymptotic properties.

Remark: Note that in above subsections we have taken true parameter values as the initial values. But we have also done these simulations while randomly choosing the initial values among the neighbourhood of the true parameter values. We have found that in this case performance of the HME and LADE deteriorates than the sequential WLSE with the appropriate weight function. Therefore, in real life data analysis, we suggest to use the proposed sequential WLSE method. In next subsection, we show the performance of all the discussed estimators based on the initial values randomly choosen from the neighbourhood of the true parameter values for a special case.

D. *Case of Close Frequency Rates*

In this subsection, we assess the performance of the proposed estimators and other discussed estimators

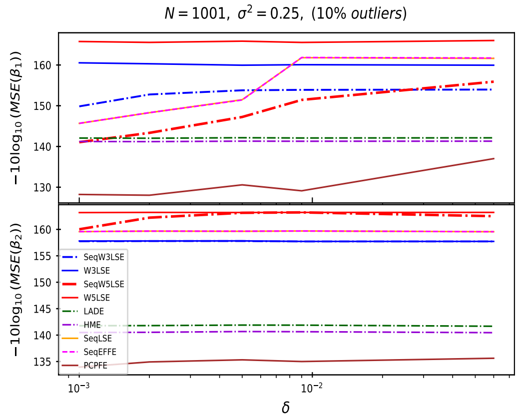


Fig. 6: $-10\log_{10}(MSE)$ of frequency rates estimates of the model (1) with $A_1^0 = 8$, $A_2^0 = 6$, $\beta_2^0 = 0.3$ and $\beta_1^0 = \beta_2^0 + \delta$ with $\delta = 10^{-3}$, 2×10^{-3}, 5×10^{-3}, 9×10^{-3} , 6×10^{-2} and 10% outliers with $\sigma_{out}^2 = 50$ versus δ.

in this paper for the case of close frequency rates. We consider a two-component model with $A_1^0 = 8$, $A_2^0 = 6$, $\beta_2^0 = 0.3$ and $\beta_1^0 = \beta_2^0 + \delta$, where, $\delta = 10^{-3}$, 2×10^{-3}, 5×10^{-3}, 9×10^{-3} , and 6×10^{-2}, and $N = 1001$. The error random variables follow Structure 1 with 10% outliers lying in the middle portion of the data, $\sigma^2 = 0.25$ and $\sigma_{out}^2 = 50$. Note that here we fixed other parameters while varying β_1^0 such that it becomes closer to $\beta_2^0 = 0.3$. We calculate WLSEs and sequential WLSEs based on five weight functions with different shapes as taken in subsection IV-A. We repeat the experiment 1000 times for each set of parameters. We report $-10\log_{10}(MSE)$ of frequency rate estimates versus gap between frquency rates in Fig. 6. From Fig. 6 it is clear that WLSEs and sequential WLSEs perform better than the other discussed estimators in terms of MSE when frequency rates are close to each other to a certain degree. From this discussion, it is recommended to use the proposed estimation method with appropriate weight function.

V. Choice of Weight Function

In the above section, we see that the estimators associated with the weight function, which gives less weight to the region of the data where outliers lie than the other region, perform better than the other estimators in terms of MSEs. This shows how important it is to choose the appropriate weight function. In this section, we discuss how to choose a weight function such that it gives less weight to the part of the data where outliers are present compared to the other parts of the data. There are two scenarios in the choice of weight function. One is that if we know in advance where the outliers are present in the data, then we should choose the weight function in such a manner that the weight function gives less weight in the region where the outliers are present compared to the other regions. The second scenario is that if we do not know that the data contains outliers or does not contain outliers, then to choose the weight function, we consider a class of variously shaped weight functions, i.e., some weight functions give less weight on the left side, some on the right side, some in the middle, some on the edges, or a mixture of these. Next, we see for which function the normalized weighted residual sum of squares (NWRSS) is minimum, then choose that corresponding weight function. Here, NWRSS is obtained by using the normalized weight function, $nw\left(s\right)$, (12) in the weighted residual sum of squares (WRSS) instead of the weight function $w\left(s\right)$. Note that we must include $w(s) = 1,\ s \in [0,1]$, in these functions. Because if no outliers are present in the data, then it should choose LSEs as the estimates of the parameters based on the minimum NWRSS. To demonstrate this method practically, we do the following simulation and synthetic and real data analysis in the succeeding sections. In the following simulation and synthetic data analysis, we have contaminated the data with outliers from the entire portion, not only in the middle, by generating them randomly using complex-valued mixture normal error random variables, which is also known as a Gaussian mixture model as done in [34].

For the simulation experiment, we consider model (1) with $p = 10$ such that

$$y\left(t\right) = 20e^{i0.5t^2} + 18e^{i1.4t^2} + 16e^{i1.1t^2} + 14e^{i0.51t^2} + 12e^{i1.7t^2} + 10e^{i0.8t^2} + 8e^{i2.3t^2} + 6e^{i2.6t^2} + 4e^{i2t^2} + 2e^{i2.9t^2} + \epsilon\left(t\right);\ t = 1, \ldots, N. \tag{13}$$

Here, $N = 101$ to 1001 and $\sigma^2 = 0.75$. Here, $\epsilon\left(t\right)$'s satisfy assumption 2 and it's real and imaginary parts are mixture normal random variables then $\epsilon\left(t\right)$'s are i.i.d. complex-valued mixture normal random variables and has the following forms:

$$\epsilon\left(t\right) \sim (1 - p_{out})\mathcal{CN}\left(0, \sigma^2\right) + p_{out}\mathcal{CN}\left(0, \sigma^2_{out}\right); \tag{14}$$

i.e., the data is contaminated with $p_{out}\%$ outliers with variance σ^2_{out}. Here, $p_{out} = 5\%, 10\%$ and $\sigma^2_{out} = 50$. We repeat the experiment 1000 times for each combination of above parameters. Here our main aim is to choose the best weighted estimator among the pool of different weighted estimators based on the above proposed criterion. Therefore, we compute the estimate of the parameters using sequential WLSEs for a class of weight functions and compare it with sequential LSEs in terms of the NWRSS. The sequential WLSEs are computed based on the following weight functions: $w_1(s) = \frac{(1-s)^2}{2} + \frac{s^2}{2}$, $w_2(s) = \frac{2(1-s)^2}{3} + \frac{s^2}{3}$, $w_3(s) = -s^2 + s + \frac{1}{10}$, $w_4(s) = \frac{2s^2}{3} + \frac{(1-s)^2}{3}$, $w_5(s) = s^2 - s + 0.26$, $w_6(s) = 1, s \in [0,1]$,

$$w_7(s) = \begin{cases} s^2 - s + 0.26 & \text{if } 0 \le s \le 0.25 \\ -s^2 + s - 0.115 & \text{if } 0.25 < s \le 0.75 \\ s^2 - s + 0.26 & \text{if } 0.75 < s \le 1, \end{cases}$$

$$w_8(s) = \begin{cases} -s^2 + s + 0.01 & \text{if } 0 \le s \le 0.25 \\ s^2 - s + 0.385 & \text{if } 0.25 < s \le 0.75 \\ -s^2 + s + 0.01 & \text{if } 0.75 < s \le 1, \end{cases}$$

$$w_9(s) = \begin{cases} 2 - 4s & \text{if } 0 \le s \le 0.25 \\ 4s & \text{if } 0.25 < s \le 0.5 \\ 3 - 2s & \text{if } 0.5 < s \le 0.75 \\ 2s & \text{if } 0.75 < s \le 1, \end{cases} \quad \text{and}$$

$$w_{10}(s) = \begin{cases} -4s^2 + s + 0.1 & \text{if } 0 \le s \le 0.25 \\ 4s^2 - s + 0.1 & \text{if } 0.25 < s \le 0.75 \\ -4s^2 + s + 3.1 & \text{if } 0.75 < s \le 1. \end{cases}$$

In this class, w_1 and w_5 put less weight in the middle of the data, w_2 puts less weight on the right side of the data, w_3 puts less weight on the edges of the data, w_4 puts less weight on the left side of the data than the other regions, w_6 puts uniform weight to data and w_7, w_8, w_9 and w_{10} put different weights at different regions of the data. See the shape of these functions, in Fig. 1. Note that if we use WLSEs to fit the data, then we have to solve the inverse of a 20×20 matrix and a 10-dimensional optimization. And if we use HMEs and LADEs, then we have to solve a 30-dimensional optimization. Therefore, it becomes computationally challenging if we use WLSEs, LADEs and HMEs in this analysis. Hence, sequential WLSEs is convenient to use in practice. We compute the NWRSS of the fitted data using sequential WLSEs in each run, which provides the best fit of the fitted data. Then, we obtain the minimum NWRSS and take the mean of these minimum NWRSSs (MNWRSS). We report $-10\log_{10}(MNWRSS)$ in Fig. 7 and compare it with NWRSS of the fitted data obtained using sequential LSE. From Fig. 7 it is visible that $-10\log_{10}(MNWRSS)$ of sequential WLSE is lower or at par with that of the sequential LSE. It indicates that the sequential WLSE fits the data well in comparison with the sequential LSE in the presence of the outliers.

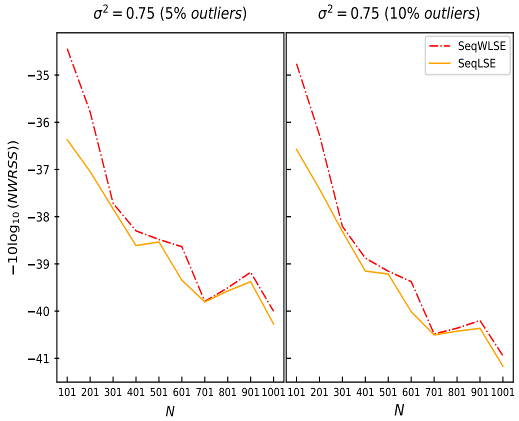
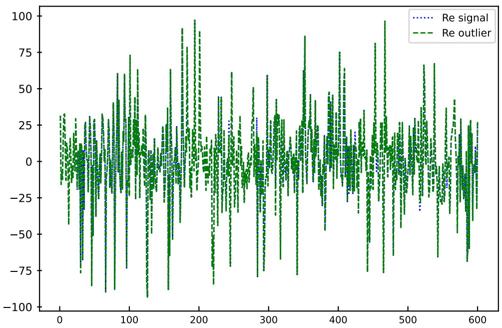


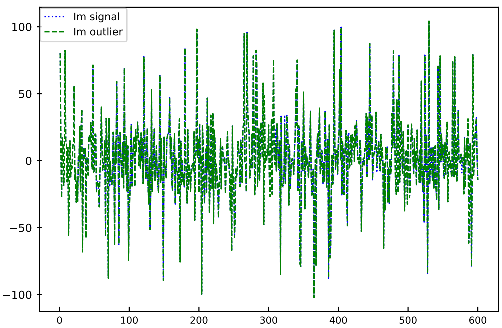
Fig. 7: $-10\log_{10}(MNWRSS)$ of the fitted data of the model (13) with 5% and 10% outliers with $\sigma^2_{out} = 50$ for various error variances versus sample size. (when outliers are generated randomly)

VI. Illustrative Example

In this section, we perform a simulated data analysis to illustrate the performance of the proposed sequential WLSEs based on choice of weight function. We generate the data from the ten-component elementary chirp model (13) with $N = 600$. Here, $\epsilon(t)$'s satisfy assumption 2 and it's real and imaginary parts are assumed to be real-valued normal random variables, i.e., $\epsilon(t) \sim \mathcal{CN}\left(0, \sigma^2\right)$ $\forall t = 1, \ldots, N$. Here σ^2 is 0.5. Further, we contaminate the original data by taking the error structure (14) with $p_{out} = 5\%$, $\sigma^2 = 0.5$ and $\sigma^2_{out} = 96.73624$. Fig. 8 shows original data and data with outliers. In this Fig. "Re" and "Im" denote the real and imaginary parts of the data, respectively. Note that by looking at Fig. 8 it is difficult to distinguish the outlier data from the original data. We analyze this contaminated data using nine sequential WLSEs and sequential LSEs by fitting elementary chirp model to it. We consider a class of variously shaped weight functions, as the positions of the outliers are random. The nine sequential WLSEs are computed based on the weight functions $w_1(s), w_2(s), w_3(s), w_4(s), w_5(s), w_7(s), w_8(s), w_9(s)$ and $w_{10}(s)$ whose mathematical form is given in the previous Section V. The $w_6(s) = 1$ is used to obtain the sequential LSEs. We have used initial values obatined by minimizing $Q_k\left(\breve{A}(\beta), \beta\right)$ over the grid, $\beta_{grid} = \dfrac{2\pi l}{N^2}$, $l = 1, \ldots, N^2 - 1$. We fit the data at 10^{th} component. We report NWRSS at 10^{th} component, i.e., NWRSS(10) which we computed based on above weight functions, in Table I. From Table I, it can be observed that the sequential WLSEs based on $w_{10}(s)$, gives the minimum NWRSS. Thus, we use sequential WLSEs based on $w_{10}(s)$ to fit an elementary chirp model to the data. In Fig. 9, fitted data using sequential WLSEs has been shown. This figure shows that the original data is enveloped by the fitted data using sequential WLSEs.



(a) Real part



(b) Imaginary part

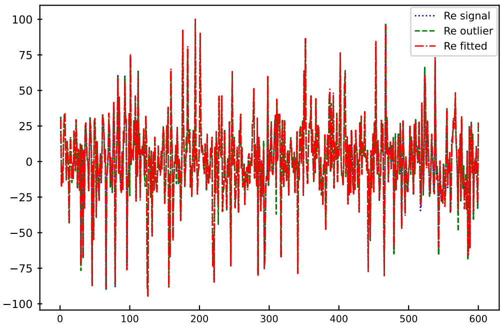
Fig. 8: The original data and data with outliers

Weight	NWRSS(10)	Weight	NWRSS(10)
w_1	11326.88	w_6	10848.71
w_2	10726.98	w_7	12888.10
w_3	10599.61	w_8	9420.25
w_4	12069.91	w_9	11535
w_5	13332.76	w_{10}	9038.41

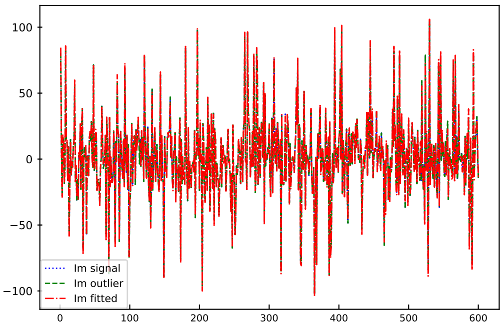
TABLE I: Reporting NWRSS(10) based on different weight functions for the fitted data.

VII. Real Data Analysis

In this section, we consider radar simulator data of 110 point scatterers from chapter 4 of [20], which are used to generate the outline of a hypothetical airplane. Then this data is used to obtain the backscattered electric field for 64 different frequencies and 64 different aspects. Further, by using the obatined backscattered electric field, the conventional small-bandwidth small-angle ISAR imaging algorithm is exploited to produce



(a) Real part



(b) Imaginary part

Fig. 9: Fitting of the outlier data with sequential WLSEs using 10CEC.

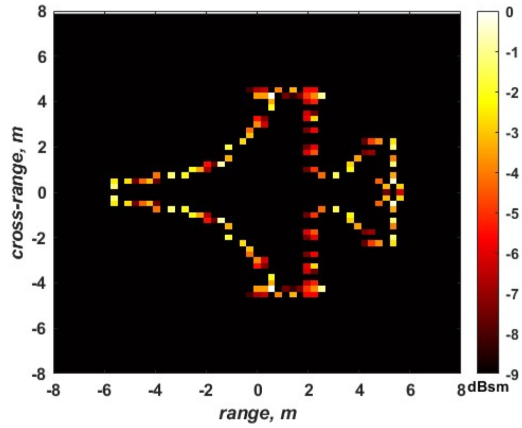
the ISAR data and the image of the airplane, as shown in Fig. 10(a). Now, we fit elementary chirp model with 27 components to the ISAR data observed in each range bin using sequential WLSEs with different weight functions, the same as those taken in Section IV. We report NWRSS at the 27^{th} component, i.e., NWRSS(27), which we computed based on different weight functions, in Table II. From Table II, it can be observed that the sequential WLSEs based on $w_3(s)$ give the minimum NWRSS. Thus, we use sequential WLSEs based on $w_3(s)$ to fit an elementary chirp model to the ISAR data. Once the fitted ISAR data is obtained, we use small-bandwidth small-angle ISAR imaging algorithm to get the image of the fitted ISAR data, which is shown in Fig. 10(b). From Fig. 10, it can be observed that the proposed sequential WLSEs fit well the ISAR data. Thus, the proposed sequential WLSEs with appropriate weight function can be used to get back the ISAR image of a target from the observed data.

VIII. Conclusion

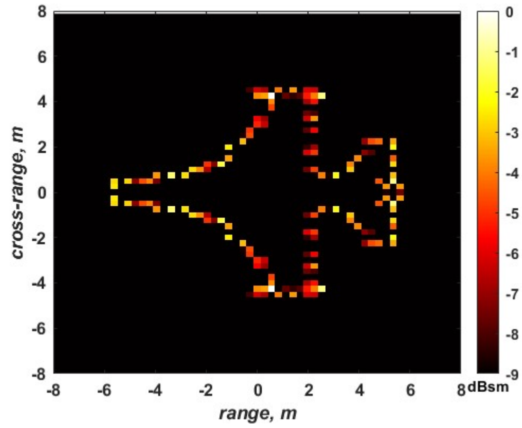
In this paper, we propose WLSEs to estimate the elementary chirp model parameters and prove that

Weight	NWRSS(27)	Weight	NWRSS(27)
w_1	4.901	w_4	4.714
w_2	5.438	w_5	6.788
w_3	4.243	w_6	4.341

TABLE II: Reporting NWRSS(27) based on different weight functions for the fitted ISAR data.



(a) ISAR image of a hypothetical airplane



(b) ISAR Image of fitted data with sequential WLSEs

Fig. 10: ISAR image of the observed and fitted data

they have the rate of convergence identical to the LSEs. We also propose sequential WLSEs for multiple component elementary chirp model, which are asymptotically equivalent to WLSEs and obtained with less computational complexity. It is observed that the WLSEs and sequential WLSEs are robust to the outliers with appropriate weight function. The sequential WLSEs take less time in computation than the other robust estimators and provide satisfactory results in the sense of the MSEs. To implement the proposed sequential WLSEs practically, we illustrate a simulated experiment based on the discussion of the choice of weight function. Additionally, we analyze synthetic and real data based on the proposed method. We observe that the proposed estimators are robust to the outliers, possess desired statistical properties under general set of assumptions and are easy to compute. The assumption of the real and imaginary parts of $\epsilon(t)$ are independently distributed can be relaxed.

There are some open problems that are of interest for future work. In this paper the objective function considers ℓ_2-norm. This is possible to enhance the robustness of the proposed WLSEs and sequential WLSEs by using the ℓ_p-norm $(1 \le p < 2)$ as suggested in [31]. We can also extend the robust weighted linear prediction estimator of [33] and [32], to estimate the parameters of multiple component elementary chirp model for the stronger resistance to the outliers. It will be interesting to develop these estimation methods for the elementary chirp model.

IX. Acknowledgements

The authors would like to thank the editor, the associate editor and the anonymous reviewers for their constructive and important suggestions which have helped to improve the manuscript significantly.

Appendix A
Preliminary Results

We will need the following results to establish the proofs of the theorems provided in this paper:

Lemma 1: If $\beta \in (0, 2\pi)$ and $w(s)$ satisfies assumption 1, then except for countable number of points, the following hold true:

$$(a) \lim_{N\to\infty} \frac{1}{N^{k+1}} \sum_{t=1}^{N} t^k w^2\left(\frac{t}{N}\right) \cos\left(\beta t^2\right) = \lim_{N\to\infty} \frac{1}{N^{k+1}} \sum_{t=1}^{N} t^k w^2\left(\frac{t}{N}\right) \sin\left(\beta t^2\right) = 0;$$

$$(b) \lim_{N\to\infty} \frac{1}{N^{k+1}} \sum_{t=1}^{N} t^k w^2\left(\frac{t}{N}\right) \cos\left(\beta t^2\right) \sin\left(\beta t^2\right) = 0;$$

$$(c) \lim_{N\to\infty} \frac{1}{N^{k+1}} \sum_{t=1}^{N} t^k w^2\left(\frac{t}{N}\right) \cos^2\left(\beta t^2\right) = \lim_{N\to\infty} \frac{1}{N^{k+1}} \sum_{t=1}^{N} t^k w^2\left(\frac{t}{N}\right) \sin^2\left(\beta t^2\right) = \frac{1}{2}\int_0^1 s^k w^2(s) ds = \frac{d_{k+1}}{2} > 0; \forall\ k = 0, 1, 2, \cdots.$$

Proof: It follows similarly as lemma C-2 of [12].

Lemma 2: If $w(s)$ satisfies assumption 1 and $\epsilon(t)$ satisfies assumption 2, then for $k \ge 0$;

$$\sup_{\beta} \left| \frac{1}{N^{k+1}} \sum_{t=1}^{N} t^k w\left(\frac{t}{N}\right) \epsilon(t) e^{i\beta t^2} \right| \xrightarrow{a.s.} 0.$$

Proof: It follows similarly as lemma B-6 of [12].

Appendix B
One Component Elementary Chirp Model

Lemma 3: Let us consider $S_\delta = \left\{\boldsymbol{\theta} : \left|\boldsymbol{\theta} - \boldsymbol{\theta}^0\right| \ge 3\delta; \boldsymbol{\theta} \in \boldsymbol{\Theta}\right\}$. If for any $\delta > 0$,

$$\liminf \inf_{\boldsymbol{\theta} \in S_\delta} \frac{1}{N}\left[Q(\boldsymbol{\theta}) - Q\left(\boldsymbol{\theta}^0\right)\right] > 0 \quad a.s., \tag{15}$$

then $\breve{\boldsymbol{\theta}} \xrightarrow{a.s.} \boldsymbol{\theta}^0$ as $N \to \infty$.

Proof: It follows similarly as lemma 1 of [17].

***Proof of Theorem** 1*: Let us consider

$$\frac{1}{N}\left[Q(\boldsymbol{\theta}) - Q\left(\boldsymbol{\theta}^0\right)\right] = f(\boldsymbol{\theta}) + g(\boldsymbol{\theta}), \text{ where,}$$

$$f(\boldsymbol{\theta}) = \frac{1}{N}\sum_{t=1}^{N} w\left(\frac{t}{N}\right)\left\{l_1^2 + l_2^2\right\},$$

$$g(\boldsymbol{\theta}) = \frac{2}{N}\sum_{t=1}^{N} w\left(\frac{t}{N}\right)\left\{\epsilon_R(t) l_1 + \epsilon_I(t) l_2\right\}, \text{ with}$$

$$l_1 = A_R^0 \cos\left(\beta^0 t^2\right) - A_R \cos\left(\beta t^2\right) - A_I^0 \sin\left(\beta^0 t^2\right) + A_I \sin\left(\beta t^2\right),$$

$$l_2 = A_R^0 \sin\left(\beta^0 t^2\right) - A_R \sin\left(\beta t^2\right) + A_I^0 \cos\left(\beta^0 t^2\right) - A_I \cos\left(\beta t^2\right).$$

Consider: $S_\delta = \left\{\boldsymbol{\theta} : \left|\boldsymbol{\theta} - \boldsymbol{\theta}^0\right| \ge 3\delta; \boldsymbol{\theta} \in \boldsymbol{\Theta}\right\}$.

$$S_\delta \subset S_{\delta 1} \cup S_{\delta 2} \cup S_{\delta 3} = S, \text{ where,}$$
$$S_{\delta 1} = \left\{\boldsymbol{\theta} : \left|A_R - A_R^0\right| \ge \delta; \boldsymbol{\theta} \in \boldsymbol{\Theta}\right\},$$
$$S_{\delta 2} = \left\{\boldsymbol{\theta} : \left|A_I - A_I^0\right| \ge \delta; \boldsymbol{\theta} \in \boldsymbol{\Theta}\right\} \text{ and}$$
$$S_{\delta 3} = \left\{\boldsymbol{\theta} : \left|\beta - \beta^0\right| \ge \delta; \boldsymbol{\theta} \in \boldsymbol{\Theta}\right\}.$$
$$\text{Thus,} \quad \liminf \inf_{\boldsymbol{\theta} \in S_\delta} f(\boldsymbol{\theta}) \ge \liminf \inf_{\boldsymbol{\theta} \in S} f(\boldsymbol{\theta}).$$

Now, the set $S_{\delta 1}$ is splitted as follows:

$$S_{\delta 1} \subset S_{\delta 1}^1 \cup S_{\delta 1}^2, \text{ where,}$$
$$S_{\delta 1}^1 = \left\{\boldsymbol{\theta} : \left|A_R - A_R^0\right| \ge \delta; \boldsymbol{\theta} \in \boldsymbol{\Theta}, \beta = \beta^0\right\} \text{ and}$$
$$S_{\delta 1}^2 = \left\{\boldsymbol{\theta} : \left|A_R - A_R^0\right| \ge \delta; \boldsymbol{\theta} \in \boldsymbol{\Theta}, \beta \ne \beta^0\right\}.$$

Now, using lemma C-2 of [12], it follows that

$$\liminf \inf_{\boldsymbol{\theta} \in S_{\delta 1}^1} f(\boldsymbol{\theta}) > 0, \text{ and } \liminf \inf_{\boldsymbol{\theta} \in S_{\delta 1}^2} f(\boldsymbol{\theta}) > 0.$$

Hence, $\liminf \inf_{\boldsymbol{\theta} \in S_{\delta 1}} f(\boldsymbol{\theta}) > 0$. Similarly, it can be shown for $S_{\delta 2}$ and $S_{\delta 3}$. Thus, $\lim_{N\to\infty} \inf \inf_{\boldsymbol{\theta} \in S_\delta} f(\boldsymbol{\theta}) > 0$ $a.s.$. Using lemma 2, we obtain:

$$\lim_{N\to\infty} \sup_{\boldsymbol{\theta} \in S_\delta} g(\boldsymbol{\theta}) = 0 \ a.s. \tag{16}$$

Therefore, using lemma 3, the result holds true.

***Proof of Theorem** 2*: Expanding $Q^{'}(\breve{\boldsymbol{\theta}})$ using Taylor series around the $\boldsymbol{\theta}^0$, as follows:

$$Q^{'}(\breve{\boldsymbol{\theta}}) - Q^{'}(\boldsymbol{\theta}^0) = (\breve{\boldsymbol{\theta}} - \boldsymbol{\theta}^0) Q^{''}(\bar{\boldsymbol{\theta}}). \tag{17}$$

Here, $\bar{\boldsymbol{\theta}}$ lies between $\boldsymbol{\theta}^0$ and $\breve{\boldsymbol{\theta}}$. Notice that $Q^{'}(\breve{\boldsymbol{\theta}}) = 0$

and therefore, rewrite (17) as follows:

$$(\breve{\boldsymbol{\theta}}-\boldsymbol{\theta}^0)\boldsymbol{D}^{-1}=-Q^{'}(\boldsymbol{\theta}^0)\boldsymbol{D}[\boldsymbol{D}Q^{''}(\bar{\boldsymbol{\theta}})\boldsymbol{D}]^{-1}. \quad (18)$$

Using lemma 1, lemma C-2 of [12] and Lindeberg Feller CLT, we obtain:

$$Q^{'}(\boldsymbol{\theta}^0)\boldsymbol{D}\xrightarrow{d}\mathcal{N}_3\left(0,\sigma^2\boldsymbol{\Sigma}\right). \quad (19)$$

Using Theorem 1 and $\bar{\boldsymbol{\theta}}$ lies between $\boldsymbol{\theta}^0$ and $\breve{\boldsymbol{\theta}}$, obtain:

$$\lim_{N\to\infty}\boldsymbol{D}Q^{''}(\bar{\boldsymbol{\theta}})\boldsymbol{D}=\lim_{N\to\infty}\boldsymbol{D}Q^{''}(\boldsymbol{\theta}^0)\boldsymbol{D}.$$

Using lemma 2 and lemma C-2 of [12], we get:

$$\lim_{N\to\infty}\boldsymbol{D}Q^{''}(\boldsymbol{\theta}^0)\boldsymbol{D}=\boldsymbol{G}. \quad (20)$$

Using (18), (19) and (20), we get the desired result.

Appendix C
Multiple Component Elementary Chirp Model

A. Proofs of the theorems of the WLSEs

Lemma 4: Let $S_\delta=\left\{\boldsymbol{u}:\left|\boldsymbol{u}-\boldsymbol{u}^0\right|\geq\delta;\boldsymbol{u}\in\boldsymbol{\mathcal{U}}\right\}$. If for any $\delta>0$,

$$\liminf\inf_{\boldsymbol{u}\in S_\delta}\frac{1}{N}\left[Q\left(\boldsymbol{u}\right)-Q\left(\boldsymbol{u}^0\right)\right]>0 \;\; a.s., \quad (21)$$

then $\breve{\boldsymbol{u}}\xrightarrow{a.s.}\boldsymbol{u}^0$ as $N\to\infty$.

Proof: It follows similarly as lemma 3.

Proof of Theorem 3: Let us consider

$$\frac{1}{N}\left[Q\left(\boldsymbol{u}\right)-Q\left(\boldsymbol{u}^0\right)\right]=f\left(\boldsymbol{u}\right)+g\left(\boldsymbol{u}\right),\text{where,}$$

$$f\left(\boldsymbol{u}\right)=\frac{1}{N}\sum_{t=1}^{N}w\left(\frac{t}{N}\right)\left\{l_1^2+l_2^2\right\},$$

$$g\left(\boldsymbol{u}\right)=\frac{2}{N}\sum_{t=1}^{N}w\left(\frac{t}{N}\right)\left\{\epsilon_R\left(t\right)l_1+\epsilon_I\left(t\right)l_2\right\},\text{ with}$$

$$l_1=\sum_{k=1}^{p}\left(A_{Rk}^0\cos\left(\beta_k^0t^2\right)-A_{Rk}\cos\left(\beta_kt^2\right)-A_{Ik}^0\right.$$
$$\left.\sin\left(\beta_k^0t^2\right)+A_{Ik}\sin\left(\beta_kt^2\right)\right),$$

$$l_2=\sum_{k=1}^{p}\left(A_{Rk}^0\sin\left(\beta_k^0t^2\right)-A_{Rk}\sin\left(\beta_kt^2\right)+A_{Ik}^0\right.$$
$$\left.\cos\left(\beta_k^0t^2\right)-A_{Ik}\cos\left(\beta_kt^2\right)\right).$$

Using lemma 2, $\lim\limits_{N\to\infty}\sup\limits_{\boldsymbol{u}\in S_\delta}g\left(\boldsymbol{u}\right)=0\;a.s.$ In the same manner as Theorem 1, it can be proved that $\lim\limits_{N\to\infty}\inf\inf\limits_{\boldsymbol{u}\in S_\delta}f\left(\boldsymbol{u}\right)>0\;a.s.$. Hence, using lemma 4, the result holds true.

Proof of Theorem 4: Expanding $Q^{'}(\breve{\boldsymbol{u}})$ using multivariate Taylor series around the point $\boldsymbol{u}^0$, as follows:

$$Q^{'}(\breve{\boldsymbol{u}})-Q^{'}(\boldsymbol{u}^0)=(\breve{\boldsymbol{u}}-\boldsymbol{u}^0)Q^{''}(\bar{\boldsymbol{u}}). \quad (22)$$

Here, $\bar{\boldsymbol{u}}$ lies between $\boldsymbol{u}^0$ and $\breve{\boldsymbol{u}}$. We rewrite (22) as:

$$(\breve{\boldsymbol{u}}-\boldsymbol{u}^0)\mathscr{D}^{-1}=-Q^{'}(\boldsymbol{u}^0)\mathscr{D}[\mathscr{D}Q^{''}(\bar{\boldsymbol{u}})\mathscr{D}]^{-1}. \quad (23)$$

Using lemma 1, lemma C-2 of [12] and Lindeberg Feller CLT, we have:

$$Q^{'}(\boldsymbol{u}^0)\mathscr{D}\xrightarrow{d}\mathcal{N}_{3p}\left(0,\sigma^2\boldsymbol{\mathcal{E}}\right);\text{ with} \quad (24)$$

$$\boldsymbol{\mathcal{E}}=\begin{bmatrix}\boldsymbol{\Sigma}_1&0&\dots&0\\0&\boldsymbol{\Sigma}_2&\dots&0\\\vdots&\vdots&\ddots&\vdots\\0&0&0&\boldsymbol{\Sigma}_p\end{bmatrix}. \quad (25)$$

Using Theorem 3 and $\bar{\boldsymbol{u}}$ lies between $\boldsymbol{u}^0$ and $\breve{\boldsymbol{u}}$, obtain:

$$\lim_{N\to\infty}\mathscr{D}Q^{''}(\bar{\boldsymbol{u}})\mathscr{D}=\lim_{N\to\infty}\mathscr{D}Q^{''}(\boldsymbol{u}^0)\mathscr{D}.$$

Using lemma 2 and lemma C-2 of [12], we have:

$$\lim_{N\to\infty}\mathscr{D}Q^{''}(\boldsymbol{u}^0)\mathscr{D}=\boldsymbol{\mathcal{G}};\text{ with} \quad (26)$$

$$\boldsymbol{\mathcal{G}}=\begin{bmatrix}\boldsymbol{G}_1&0&\dots&0\\0&\boldsymbol{G}_2&\dots&0\\\vdots&\vdots&\ddots&\vdots\\0&0&0&\boldsymbol{G}_p\end{bmatrix}. \quad (27)$$

Using (23), (24) and (26), the stated result holds true.

B. Proofs of the theorems of the Sequential WLSEs

Lemma 5: Let $S_\delta=\left\{\boldsymbol{\theta}:\left|\boldsymbol{\theta}-\boldsymbol{\theta}_1^0\right|\geq3\delta;\boldsymbol{\theta}\in\boldsymbol{\Theta}\right\}$. If for any $\delta>0$,

$$\liminf\inf_{\boldsymbol{\theta}\in S_\delta}\frac{1}{N}\left[Q_1\left(\boldsymbol{\theta}\right)-Q_1\left(\boldsymbol{\theta}_1^0\right)\right]>0\;a.s., \quad (28)$$

then $\tilde{\boldsymbol{\theta}}_1\xrightarrow{a.s.}\boldsymbol{\theta}_1^0$ as $N\to\infty$.

Proof: It follows by proceeding similarly as lemma 3.

Proof of Theorem 5:

$$\frac{1}{N}\left[Q_1\left(\boldsymbol{\theta}\right)-Q_1\left(\boldsymbol{\theta}_1^0\right)\right]=f\left(\boldsymbol{\theta}\right)+g\left(\boldsymbol{\theta}\right),\text{where,}$$

$$f\left(\boldsymbol{\theta}\right)=\frac{1}{N}\sum_{t=1}^{N}w\left(\frac{t}{N}\right)\left\{l_1^2+l_2^2\right\}+\frac{2}{N}\sum_{t=1}^{N}w\left(\frac{t}{N}\right)$$
$$\left\{l_1\left(\sum_{k=2}^{p}A_{Rk}^0\cos\left(\beta_k^0t^2\right)-A_{Ik}^0\sin\left(\beta_k^0t^2\right)\right)+l_2\right.$$
$$\left.\left(\sum_{k=2}^{p}A_{Rk}^0\sin\left(\beta_k^0t^2\right)+A_{Ik}^0\cos\left(\beta_k^0t^2\right)\right)\right\},$$

$$g\left(\boldsymbol{\theta}\right)=\frac{2}{N}\sum_{t=1}^{N}w\left(\frac{t}{N}\right)\left\{\epsilon_R\left(t\right)l_1+\epsilon_I\left(t\right)l_2\right\},\text{ with}$$

$$l_1=A_{R1}^0\cos\left(\beta_1^0t^2\right)-A_R\cos\left(\beta t^2\right)-A_{I1}^0\sin\left(\beta_1^0t^2\right)$$
$$+A_I\sin\left(\beta t^2\right),$$

$$l_2=A_{R1}^0\sin\left(\beta_1^0t^2\right)-A_R\sin\left(\beta t^2\right)+A_{I1}^0\cos\left(\beta_1^0t^2\right)$$
$$-A_I\cos\left(\beta t^2\right).$$

Using lemma 2, $\lim_{N\to\infty}\sup_{\boldsymbol{\theta}\in S_\delta} g(\boldsymbol{\theta}) = 0$ $a.s.$. In the same manner as Theorem 1, it can be proved that $\lim_{N\to\infty}\inf \inf_{\boldsymbol{\theta}\in S_\delta} f(\boldsymbol{\theta}) > 0$ $a.s.$. Hence, using lemma 5, the result.

Lemma 6: If assumptions 1, 2 and 3 hold, then $\left(\tilde{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1^0\right)\left(\sqrt{N}\boldsymbol{D}\right)^{-1} \xrightarrow{a.s.} 0$ as $N \to \infty$, where $\boldsymbol{D}$ is same as defined in Theorem 2.

Proof: Consider the following:

$$Q_1(\boldsymbol{\theta}) = \sum_{t=1}^{N} w\left(\frac{t}{N}\right)\left|y_1(t) - Ae^{i\beta t^2}\right|^2. \tag{29}$$

Here, $y_1(t) = y(t)$; the original data. Now, expanding $Q_1^{'}(\tilde{\boldsymbol{\theta}}_1)$ using Taylor series around the point $\boldsymbol{\theta}_1^0$:

$$Q_1^{'}(\tilde{\boldsymbol{\theta}}_1) - Q_1^{'}(\boldsymbol{\theta}_1^0) = (\tilde{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1^0)Q_1^{''}(\bar{\boldsymbol{\theta}}_1). \tag{30}$$

Here, $\bar{\boldsymbol{\theta}}_1$ lies between $\boldsymbol{\theta}_1^0$ and $\tilde{\boldsymbol{\theta}}_1$. We rewrite (30) as:

$$\begin{aligned}&(\tilde{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1^0)\left(\sqrt{N}\boldsymbol{D}\right)^{-1} = \\ &\left[-\frac{1}{\sqrt{N}}Q_1^{'}(\boldsymbol{\theta}_1^0)\boldsymbol{D}\right]\left[\boldsymbol{D}Q_1^{''}(\bar{\boldsymbol{\theta}}_1)\boldsymbol{D}\right]^{-1}.\end{aligned} \tag{31}$$

Using lemma 2, we obtain the following:

$$\frac{1}{\sqrt{N}}Q_1^{'}(\boldsymbol{\theta}_1^0)\boldsymbol{D} \xrightarrow{a.s.} 0 \quad \text{as} \quad N \to \infty. \tag{32}$$

Using Theorem 5 and $\bar{\boldsymbol{\theta}}_1$ lies between $\boldsymbol{\theta}_1^0$ and $\tilde{\boldsymbol{\theta}}_1$, get:

$$\lim_{N\to\infty}\boldsymbol{D}Q_1^{''}(\bar{\boldsymbol{\theta}}_1)\boldsymbol{D} = \lim_{N\to\infty}\boldsymbol{D}Q_1^{''}(\boldsymbol{\theta}_1^0)\boldsymbol{D}.$$

Using lemma 2 and lemma C-2 of [12], we have:

$$\lim_{N\to\infty}\boldsymbol{D}Q_1^{''}(\boldsymbol{\theta}_1^0)\boldsymbol{D} = \boldsymbol{G}_1. \tag{33}$$

Using (31), (32) and (33), the result holds true.

Proof of Theorem 6 : Using lemma 6, obtain:

$$\tilde{A}_{R1} = A_{R1}^0 + o(1),\ \tilde{A}_{I1} = A_{I1}^0 + o(1) \ \text{and}$$
$$\tilde{\beta}_1 = \beta_1^0 + o\left(N^{-2}\right).$$

$$\text{Hence,}\quad \tilde{A}_1 e^{i\tilde{\beta}_1 t^2} = A_1^0 e^{i\beta_1^0 t^2} + o(1). \tag{34}$$

Hence the result follows, using (34) and proceeding similarly as Theorem 5.

Using Lemma 7 the result of Theorem 8 follows.

Lemma 7: If $\tilde{A}_R, \tilde{A}_I$ and $\tilde{\beta}$ are obtained by minimizing the following:

$$Q_{(p+1)}(\boldsymbol{\theta}) = \frac{1}{N}\sum_{t=1}^{N} w\left(\frac{t}{N}\right)\left|\epsilon(t) - Ae^{i\beta t^2}\right|^2,$$

and assumption 2 holds, then $\tilde{A}_R \xrightarrow{a.s.} 0$ and $\tilde{A}_I \xrightarrow{a.s.} 0$.

Proof: Consider:

$$\begin{aligned}&Q_{(p+1)}(\boldsymbol{\theta}) = \\ &\frac{1}{N}\sum_{t=1}^{N} w\left(\frac{t}{N}\right)\Big[|\epsilon(t)|^2 + \left(A_R^2 + A_I^2\right)\left\{\cos^2\left(\beta t^2\right) + \right. \\ &\left.\sin^2\left(\beta t^2\right)\right\} - 2\left\{\epsilon_I(t)\left(A_I\cos\left(\beta t^2\right) + A_R\sin\left(\beta t^2\right)\right)\right. \\ &\left. + \epsilon_R(t)\left(A_R\cos\left(\beta t^2\right) - A_I\sin\left(\beta t^2\right)\right)\right\}\Big].\end{aligned}$$

Thus, we have the esimators:

$$\tilde{A}_R = \frac{\sum_{t=1}^{N} w\left(\frac{t}{N}\right)\left(\epsilon_R(t)\cos\left(\beta t^2\right) + \epsilon_I(t)\sin\left(\beta t^2\right)\right)}{\sum_{t=1}^{N} w\left(\frac{t}{N}\right)\left(\cos^2\left(\beta t^2\right) + \sin^2\left(\beta t^2\right)\right)}.$$

$$\tilde{A}_I = \frac{\sum_{t=1}^{N} w\left(\frac{t}{N}\right)\left(\epsilon_I(t)\cos\left(\beta t^2\right) - \epsilon_R(t)\sin\left(\beta t^2\right)\right)}{\sum_{t=1}^{N} w\left(\frac{t}{N}\right)\left(\cos^2\left(\beta t^2\right) + \sin^2\left(\beta t^2\right)\right)}.$$

Using lemma 2 and lemma C-2 of [12], the result.

***Proof of Theorem** 9* : Consider:

$$Q_1(\boldsymbol{\theta}) = \sum_{t=1}^{N} w\left(\frac{t}{N}\right)\left|y_1(t) - Ae^{i\beta t^2}\right|^2. \tag{35}$$

Using Taylor series, expand $Q_1^{'}(\tilde{\boldsymbol{\theta}}_1)$, then we obtain:

$$(\tilde{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_1^0)\boldsymbol{D}^{-1} = -Q_1^{'}(\boldsymbol{\theta}_1^0)\boldsymbol{D}[\boldsymbol{D}Q_1^{''}(\bar{\boldsymbol{\theta}}_1)\boldsymbol{D}]^{-1}. \tag{36}$$

Using lemma 1, Lindeberg Feller CLT, lemmas C-2 and C-3 of [12], we get:

$$Q_1^{'}(\boldsymbol{\theta}_1^0)\boldsymbol{D} \xrightarrow{d} \mathcal{N}_3\left(0, \sigma^2\boldsymbol{\Sigma}_1\right);\ \text{where,} \tag{37}$$

$$\boldsymbol{\Sigma}_1 = 2\begin{bmatrix} d_1 & 0 & -A_{I1}^0 d_3 \\ 0 & d_1 & A_{R1}^0 d_3 \\ -A_{I1}^0 d_3 & A_{R1}^0 d_3 & \left|A_1^0\right|^2 d_5 \end{bmatrix}. \tag{38}$$

Using $\tilde{\boldsymbol{\theta}}_1 \xrightarrow{a.s.} \boldsymbol{\theta}_1^0$ as $N \to \infty$,

$$\lim_{N\to\infty}\boldsymbol{D}Q_1^{''}(\bar{\boldsymbol{\theta}}_1)\boldsymbol{D} = \lim_{N\to\infty}\boldsymbol{D}Q_1^{''}(\boldsymbol{\theta}_1^0)\boldsymbol{D}.$$

Using lemma 2 and C-2 of [12], it can be obtained that:

$$\lim_{N\to\infty}\boldsymbol{D}Q_1^{''}(\boldsymbol{\theta}_1^0)\boldsymbol{D} = \boldsymbol{G}_1. \tag{39}$$

Using (36), (37) and (39), asymptotic distribution of $\tilde{\boldsymbol{\theta}}_1$ is obtained. The result for $\tilde{\boldsymbol{\theta}}_2$ can be obtained in the similar manner using lemma 6. Similarly, this result can be extended $\forall\ k \leq p$.

References

[1] Y. I. Abramovich and P. Turcaj, "Impulsive noise mitigation in spatial and temporal domains for surface-wave over-the horizon radar," *DTIC Document, Cooperative Research Centre for Sensor Signal and Information Processing, Mawson Lakes*, Australia 1999.

[2] K. Abratkiewicz. "Sonar pulse detection using chirp rate estimation and CFAR algorithms," *Hydroacoustics*, 20, 2017.

[3] K. Abratkiewicz, and P. Samczynski. " A block method using the chirp rate estimation for NLFM radar pulse reconstruction," *Sensors*, vol 19, no. 22, 2019.

[4] O. Aldimashki, A. Serbes. "LFM signal parameter estimation in the fractional Fourier domains: Analytical models and a high-performance algorithm," *Signal Processing*, 214, 109224, 2024.

[5] O. Aldimashki, A. Serbes. "Performance of chirp parameter estimation in the fractional Fourier domains and an algorithm for fast chirp-rate estimation," *IEEE Transactions on Aerospace and Electronic Systems*, vol 56, no. 5, 3685–3700, 2020.

[6] D. C. Alexander, G. J. Barker, and S. R. Arridge, "Detection and modeling of non- Gaussian apparent diffusion coefficient profiles in human brain data," *Magnetic Resonance in Medicine*, vol. 48, no. 2, pp. 331–340, 2002.

[7] G. Bai, Y. Cheng, W. Tang, and S. Li. "Chirp rate estimation for LFM signal by multiple DPT and weighted combination," *IEEE Signal Processing Letters*, vol 26, no. 1, 149–153, 2018.

[8] J. Cao, N. Zhang, and L. Song. "A fast algorithm for the chirp rate estimation," *4th IEEE International Symposium on Electronic Design, Test and Applications*, 45–48, 2008.

[9] P. C. Etter, "Underwater Acoustic Modeling and Simulation," *New York: Taylor & Francis*, 2003.

[10] D. Fourer, F. Auger, K. Czarnecki, S. Meignen, and P. Flandrin. "Chirp rate and instantaneous frequency estimation: application to recursive vertical synchrosqueezing," *IEEE Signal Processing Letters*, vol 24, no. 11, 1724–1728, 2017.

[11] I. Djurović, C. Ioana, L. Stanković, and P. Wang. "Adaptive algorithm for chirp-rate estimation," *EURASIP Journal on Advances in Signal Processing*, 1–9, 2009.

[12] D. Kundu, S. Nandi, and R. Grover. "On Weighted Least Squares Estimators for Chirp Like Model," *Sankhya A*, pp.1-40, 2023.

[13] A. Lahiri, D. Kundu, and A. Mitra. "On least absolute deviation estimators for one-dimensional chirp model," *Statistics*, vol 48, no. 2, 405-420, 2014.

[14] A. Lapidoth. "A foundation in digital communication," *Cambridge University Press*, 2017.

[15] N. Laurent, M. A. Colominas, and S. Meignen. "On Local Chirp Rate Estimation in Noisy Multicomponent Signals: With an Application to Mode Reconstruction," *IEEE Transactions on Signal Processing*, vol 70, 3429–3440, 2022.

[16] M. Mboup, and T. Adalı. "A generalization of the Fourier transform and its application to spectral analysis of chirp-like signals," *Applied and Computational Harmonic Analysis*, vol 32, no. 2, 305–312, 2012.

[17] A. Mittal, R. Grover, D. Kundu, and A. Mitra. "Estimation of the Elementary Chirp Model Parameters," *IEEE Transactions on Aerospace and Electronic Systems*, vol 59, no. 5, 5218–5234 , 2023.

[18] A. Mittal, R. Grover, D. Kundu, and A. Mitra. "On Efficient Parameter Estimation of Elementary Chirp Model," *IEEE Transactions on Signal Processing*, vol 71, 2352–2365, 2023.

[19] P. O'shea. "A fast algorithm for estimating the parameters of a quadratic FM signal," *IEEE Transactions on Signal Processing*, vol 52, no. 2, 385–393, 2004.

[20] C. Ozdemir. "Inverse synthetic aperture radar imaging with MATLAB algorithms," *John Wiley & Sons*, 2021.

[21] Y. Pantazis, , O. Rosec, and Y. Stylianou. "Chirp rate estimation of speech based on a time-varying quasi-harmonic model," *IEEE International Conference on Acoustics, Speech and Signal Processing*, 3985–3988, 2009.

[22] L. Qi, Z. Shen, Q. Guo, Y. Wang, and K. Mykola. "Chirp Rates Estimation for Multiple LFM Signals by DPT–SVD," *Circuits, Systems, and Signal Processing*, vol 42, no. 5, 2804-2827, 2023.

[23] N. Rabiee, H. Azad, and N. Parhizgar,. "Exploiting Chirp Rate Estimation Methods to Improve Image Formation Quality of Synthetic Aperture Radar," *International Journal of Aerospace Engineering*, vol. 2021, 1–13, 2021.

[24] H. Saleem, and Z. M. Hussain. "Instantaneous Frequency and Chirp Rate Estimation for Noisy Quadratic FM Signals by CNN," *Journal of Kufa for Mathematics and Computer*, vol 10, no. 2, 1-9, 2023.

[25] E. Swiercz, D. Janczak, and K. Konopko. "Detection of LFM radar signals and chirp rate estimation based on time-frequency rate distribution," *Sensors*, vol 21, no. 16, 5415, 2021.

[26] J. Song, Y. Xu, Y. Liu, and Y. Zhang. "Investigation on estimator of chirp rate and initial frequency of LFM signals based on modified discrete chirp Fourier transform," *Circuits, Systems, and Signal Processing*, vol 38, no. 12, 5861–5882, 2019.

[27] P. Wang, H. Li, I. Djurović, and B. Himed. "Integrated cubic phase function for linear FM signal analysis," *IEEE Transactions on Aerospace and Electronic Systems*, vol 46, no. 3, 963-977, 2010.

[28] P. Wang, and J. Yang. "Multicomponent chirp signals analysis using product cubic phase function," *Digital Signal Processing*, vol 16, no. 6, 654–669, 2006.

[29] X.G. Xia. "Discrete chirp-Fourier transform and its application to chirp rate estimation," *IEEE Transactions on Signal Processing*, vol 48, no. 11, 3122–3133, 2000.

[30] A. Young, D. Luong, B.Balaji, and S. Rajan. "Machine Learning Approach to Chirp Rate Estimation of Linear Frequency Modulated Radars," *Integrated Communications Navigation and Surveillance Conference (ICNS), IEEE*, 2020.

[31] W. Zeng, H. C. So, and L. Huang, "ℓ_p-MUSIC: Robust direction-of-arrival estimator for impulsive noise environments," *IEEE Transactions on Signal Processing*, vol. 61, no. 17, pp. 4296–4308, 2013.

[32] Z. Zhou, L. Huang, M. G. Christensen and S. Zhang, "Robust Spectral Analysis of Multi-Channel Sinusoidal Signals in Impulsive Noise Environments," *IEEE Transactions on Signal Processing*, vol. 70, pp. 919-935, 2022.

[33] Z. Zhou, Y. Liu, M. G. Christensen and K. Chen, "Linear Prediction Approach to the Robust Parameter Estimation for the Damped Sinusoids," *2020 15th IEEE International Conference on Signal Processing (ICSP)*, pp. 478-483, 2020.

[34] Z. Zhou and B. Liao. "Maximum a Posteriori Based Fundamental Frequency and Order Estimation in Impulsive Noise," *IEEE Transactions on Signal Processing*, vol. 71, pp. 3053-3066, 2023, doi: 10.1109/TSP.2023.3306200.

[35] A. M. Zoubir, V. Koivunen, Y. Chakhchoukh, and M. Muma, "Robust estimation in signal processing: A tutorial-style treatment of fundamental concepts," *IEEE Signal Processing Magazine*, vol. 29, no. 4, pp. 61–80, 2012.

[36] A. M. Zoubir, V. Koivunen, Y. Chakhchoukh, and M. Muma, "Robust statistics for signal processing," *Cambridge University Press*, 2018.