On Efficient Robust Estimation of Elementary Chirp Model Parameters

Anjali Mittal[a], Debasis Kundu[b], Amit Mitra[b]

[a]*Department of Mathematics, Indian Institute of Technology Indore, Indore, 453552, Madhya Pradesh, India*
[b]*Department of Statistics and Data Science, Indian Institute of Technology Kanpur, Kanpur, 208016, Uttar Pradesh, India*

Abstract

In this paper, we present an efficient algorithm which produces robust estimates of the parameters of the elementary chirp model. The crucial thing about the proposed algorithm is that it produces frequency rate estimates in a finite number of steps with desired theoretical properties. Recently, weighted least squares estimators (WLSEs) were proposed to estimate the parameters of the elementary chirp model; these estimators are robust to outliers and have desirable statistical properties. However, it is numerically challenging to compute WLSEs. In this paper, we put forward estimators that are robust to outliers, possess desired theoretical properties under the same error assumptions, and are computationally efficient. We establish that the proposed efficient robust estimators (EREs) are consistent and have same rate of convergence as WLSEs. For the multiple component elementary chirp model, we have proposed a sequential robust efficient algorithm. We also establish the asymptotic properties of the proposed sequential EREs. It has been reported from the comprehensive simulation studies that the proposed estimators perform satisfactorily in the presence of outliers while taking less computing time. The performance of the proposed estimators depends on the weight function; thereby, its selection has also been addressed. The performance of the proposed methodology has been validated using a real EEG data set.

Keywords: Frequency rate, chirp rate, elementary chirp, consistency, asymptotic normality, efficient robust estimators, sequential efficient robust estimators.

1. Introduction

In this paper, we consider the following elementary chirp model with p components:

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

where, $i = \sqrt{-1}$ and p is the number of components present in the model, which is assumed to be known in advance. The A_k^0's are the amplitude parameters and β_k^0's are the frequency rate parameters. The $\epsilon(t)$'s represent the error random variables present in the observed signal and satisfy Assumption 1. Here, for the given sample, the objective is to estimate the unknown parameters, i.e., A_k^0's and β_k^0's.

Observe that in the multiple component elementary chirp (MCEC) model (1), the frequency rates, i.e., β_k^0's are the non-linear parameters and the amplitude parameters, i.e., A_k^0's are the linear parameters. Once we have the estimators for the β_k^0's, then estimators for the A_k^0's can be found using linear regression. Therefore, estimation of the frequency rates is crucial here. In the signal processing literature, the frequency rate is popularly known as the chirp rate. The chirp rate estimation has applications in various areas, to name a few, in velocity estimation of moving targets [22], in radar [3, 4, 16, 27], in synchrosqueezing [11, 31], in sonar [1] and in speech signal [2]. Various estimation methods for chirp rate estimation are available in the literature. The estimators based on local polynomial Fourier transform [10], quadratic phase function [15], discrete polynomial-phase transform [26], cubic phase function [23, 30], modified discrete chirp-Fourier transform [29], regression quantile [8], fractional Fourier transform [5, 6], discrete polynomial transform and singular value decomposition [25], convolution neural network [28], least squares and approximate least squares [18], and finite step efficient procedure [19] are in the literature.

Different estimation methods for the unknown parameters related to elementary chirp models have been discussed by several authors. To name a few, the detection of chirp model based on Wigner distribution [17], estimation of chirp model based on various filters [21], estimation of chirplet [9], estimation of sinusoids based on amplitudes [7] and time-frequency analysis [24] have been discovered.

The least squares estimators (LSEs) are the intrinsic estimators to estimate the MCEC model (1) parameters. Recently, efficient estimators (EFFEs) based on a finite step iterative procedure [19] have been established to estimate the MCEC model (1) parameters. It has been proven that the EFFEs are consistent and have identical asymptotic distribution as that of the LSEs under the same error assumptions as the LSEs. However, LSEs and EFFEs are sensitive when outliers are present. The robust estimators for estimation in the presence of outliers are Huber's M-estimators (HMEs) and least absolute deviation estimators (LADEs). However, it is difficult to compute these robust estimators numerically, and it is difficult to determine their statistical properties under the same error assumption as LSEs and EFFEs. These issues were explicitly discussed in [20] and were overcome through the development of weighted least squares estimators (WLSEs) and sequential WLSEs [20] to estimate the MCEC model (1) parameters in the presence of outliers. It has been proven that WLSEs and sequential WLSEs have desired statistical properties, namely, consistency and asymptotic normality, under the same error assumption as LSEs and EFFEs. Furthermore, in comparison to other robust estimators such as HMEs and LADEs, it has been reported that WLSEs and sequential WLSEs require less time and are straightforward to employ in practice. However, numerically finding sequential WLSEs is a challenging problem due to the inherent non-linearity involved in the optimization. Our objective is to put forward estimators that are robust to outliers, possess desired statistical properties under the same error assumptions as WLSEs, LSEs and EFFEs, and are computationally less challenging.

In this paper, we introduce efficient robust estimators (EREs) based on a robust efficient algorithm to estimate the one component elementary chirp (OCEC) model (2) parameters, which provide consistent estimators with the identical rate of convergence as WLSEs, LSEs and EFFEs. The motivation for EREs comes from WLSEs, which were recently proposed to estimate the parameters of the MCEC model (1), the chirp model [13] and the chirp-like model [14] in the presence of outliers. The idea is to choose the weight function that provides less weight to the part of the data where outliers are present compared to the other parts of the data. For the MCEC model (1) parameter estimation, we propose sequential EREs, where estimates of the model are obtained sequentially. It has been established that sequential EREs are consistent

and have the identical rate of convergence as WLSEs, LSEs and EFFEs. It has been observed that the proposed robust efficient algorithm is a finite step iterative algorithm. This algorithm provides the frequency rate estimator in three steps if we start with an initial estimator of β^0 with the convergence rate $O_p\left(N^{-2}\right)$. It has been reported that EREs and sequential EREs are robust to outliers and perform better than LADEs and HMEs. Also, the proposed robust efficient algorithm takes less time to compute the estimates than the other robust estimators (please see Fig. 7).

Further, to validate the theoretical properties, we perform various types of numerical experiments for different sets of parameters. For the OCEC model (2), the performance of EREs is compared with LSEs, EFFEs, WLSEs, LADEs, HMEs and cubic phase function estimators (CPFEs). We observe that the performance of EREs is at par with WLSEs. We observe that, in the absence of outliers, LSEs and EFFEs perform best in terms of the mean squared error (MSE) among all the discussed estimators. On the other hand, in the presence of outliers, EREs and WLSEs with an appropriate weight function perform best among all the discussed estimators. For the MCEC model (1), the performance of sequential EREs is compared with sequential LSEs, sequential EFFEs, sequential WLSEs, LADEs, HMEs and product cubic phase function estimators (PCPFEs). In the presence of outliers, sequential EREs and sequential WLSEs with an appropriate weight function perform best among all the discussed estimators. Further, we discuss the choice of weight function based on the leave-one-out cross-validation with an illustration of a real data analysis based on the proposed methodology. This method provides better fitting of the data in terms of the normalized weighted residual sum of squares (NWRSS) in comparison to the choice of weight function method based on the pool of weight functions given in [20]. The NWRSS is computed by using the normalized weight function, $\tilde{w}(s)$, (19) in the weighted residual sum of squares (WRSS) instead of the weight function $w(s)$. This demonstrates that the new method provides a better fit to the data, as shown in Section 5.

The critical contributions of this paper are outlined as follows: In this paper, we put forward sequential robust efficient algorithm to estimate the parameters of the MCEC model (1). The proposed sequential EREs are obtained in finite steps with desired statistical properties, robust to the outliers and simple to use. The property of the finite step convergence is unique of the proposed algorithm, which is not common in the usual iterative algorithms. The proposed estimators are less costly numerically, less time consuming, easy to compute, and it can be used under a weaker assumption compared to the other robust estimators like HMEs and LADEs. Moreover, the proposed method works in more general setup. Specifically, if there are no outliers in the data, then also the proposed method works, whereas other robust estimators like HMEs and LADEs perform poorly. We also propose a method to choose the proper weight function, which is based on the cross validation method.

The remainder of the paper is structured as follows. In Section 2, we study the asymptotic properties of the proposed EREs for the OCEC model (2). In Section 3, we examine the asymptotic properties of the proposed sequential EREs for the MCEC model (1). Section 4 presents the simulation studies. In Section 5, we discuss the choice of the weight function and analyze a real data set. Finally, the conclusions are presented in Section 6. The derivations of all the theorems are provided in the appendices.

2. OCEC Model

Consider the OCEC model as follows:

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

where, $A^0 = A_R^0 + iA_I^0$ is the amplitude parameter. The A_R^0 and A_I^0 are the real and imaginary parts of A^0, respectively, such that at least one of them is non-zero. The β^0 is the frequency rate parameter with $\beta^0 \in (0, 2\pi)$. The $\epsilon(t)$'s are the error random variables embedded in $y(t)$. For the observed signal of size N, the objective is to estimate the unknown parameters A^0 and β^0.

Following are the assumptions on the weight function, i.e., $w(s)$ and error random variables under which statistical properties of EREs have been derived.

Assumption 1. *Let $\{\epsilon(t); t = 1, \ldots, N\}$ be a sequence of independent and identically distributed (i.i.d.) complex-valued random variables, where $\epsilon(t) = \epsilon_R(t) + i\epsilon_I(t)$. The $\epsilon_R(t)$ and $\epsilon_I(t)$ of $\epsilon(t)$ are independently distributed, with $E(\epsilon_R(t)) = 0, V(\epsilon_R(t)) = \frac{\sigma^2}{2}$, $E(\epsilon_I(t)) = 0, V(\epsilon_I(t)) = \frac{\sigma^2}{2}$. The fourth order moment of real and imaginary parts of $\epsilon(t)$ exist.*

Assumption 2. *Suppose $w(s)$ is a non-negative continuous function defined on* $[0, 1]$*, satisfying* $\min_{0 \le s \le 1} w(s) > \eta > 0$.

We need assumptions 1 and 2 on weight function and error random variables to prove the lemmas associated with these quantities. Further, these lemmas are used in proving the theoretical properties of the proposed estimators. Even in the case of LSEs, these are the minimum assumptions on $\epsilon(t)$'s that we need, please see [18].

In the following subsection, we propose an iterative method for parameter estimation of OCEC model (2) which provides estimate in finite steps. This algorithm is similar to the efficient algorithm given in [19]. However, the algorithm provided in [19] does not provide estimators that are robust to the outliers. To make the algorithm robust, we incorporate the weight function in the correction terms, given in (3) and (4), in such a way that the algorithm trims out the effect of the outliers present in the data. This algorithm provides estimators that converge at the same rate as the WLSEs, LSEs and EFFEs and robust over LSEs and EFFEs for estimating the parameters when outliers are present.

2.1. Robust Efficient Algorithm

We now propose the updating scheme of an algorithm that produces a frequency rate estimator with the identical rate of convergence as that of the WLSEs, LSEs and EFFEs. The theoretical justification of it comes from Theorem 1. Let us introduce some notations as follows:

$$P_N^\beta = \sum_{t=1}^{N} y(t)\, w\left(\frac{t}{N}\right)\left(t^2 - \frac{N^2}{3}\right) e^{-i\beta t^2}, \tag{3}$$

$$Q_N^\beta = \sum_{t=1}^{N} y(t)\, w\left(\frac{t}{N}\right) e^{-i\beta t^2}, \tag{4}$$

$$c_{k+1} = \int_0^1 s^k w(s) ds > 0, \tag{5}$$

$$d_{k+1} = \int_0^1 s^k w^2(s) ds > 0,\ k = 0, 1, 2, \ldots. \tag{6}$$

Here, $w(.)$ is a weight function satisfying Assumption 2.

If $\tilde{\beta}$ is a consistent estimator of β^0, such that $\tilde{\beta} - \beta^0 = O_p\left(N^{-2-\delta}\right)$, for some $\delta > 0$, then the updated estimator of β^0 is obtained as follows:

$$\breve{\beta} = \tilde{\beta} + \frac{3c_1}{(3c_5 - c_3)\, N^4} \operatorname{Im}\left(\frac{P_N^{\tilde{\beta}}}{Q_N^{\tilde{\beta}}}\right), \text{ provided}, 3c_5 - c_3 \neq 0. \tag{7}$$

Here, $\operatorname{Im}(.)$ denotes the imaginary part of a complex number and c_k's are obtained from equation (5). We initialize the algorithm with a consistent estimator $\tilde{\beta}$ and then update the estimator using (7). Please note that the correction terms, given in (3) and (4), involve the weight function, which takes care of the effect of the outliers present in the data. Moreover, for $w(s) = 1$, the proposed algorithm (7) is same as the algorithm given in [19].

Note: Please note that the set of coefficients of $w(s)$ satisfying $3c_5 - c_3 = 0$ is nonempty under the assumptions on the weight function. For example, for the linear weight function, i.e., $w(s) = a_0 + a_1 s$, we have $3c_5 - c_3 \neq 0$. However, for the quadratic weight function, i.e., $w(s) = a_0 + a_1 s + a_2 s^2$, there exist specific choices of the coefficients for which $3c_5 - c_3 = 0$. One such example is $w(s) = s^2 - \left(\frac{1872}{945}\right) s + 1, \quad s \in [0, 1]$ for which $3c_5 - c_3 = 0$. Although, it seems to be an restriction, in practice it is actually not. Even with these restrictions, the practitioner has ample choice to choose the weight function. The details will be discussed later.

Theorem 1. *If* $\tilde{\beta} - \beta^0 = O_p\left(N^{-2-\delta}\right)$, *where* $\delta \in \left(0, \frac{1}{2}\right]$, *then*

(*a*) $\breve{\beta} - \beta^0 = O_p\left(N^{-2-2\delta}\right)$ *if* $\delta \leq \frac{1}{4}$.

(*b*) $N^{\frac{5}{2}}\left(\breve{\beta} - \beta^0\right) \xrightarrow{d} \mathcal{N}\left(0, \sigma^2\left(\frac{d_1 + 9d_5 - 6d_3}{2(3c_5 - c_3)^2 |A^0|^2}\right)\right)$, *provided,* $3c_5 - c_3 \neq 0$ *if* $\delta > \frac{1}{4}$, *where,* c_k*'s and* d_k*'s are obtained from equations* (5) *and* (6)*, respectively. Here* $\xrightarrow{d}$ *stands for convergence in distribution and* $\mathcal{N}$ *denotes the normal distribution.*

Proof. See Appendix A.

Theorem 1 states that if we start with an estimator of order $O_p\left(N^{-2-\delta}\right)$ then we obtain the estimator with improved order $O_p\left(N^{-2-2\delta}\right)$ if $\delta \leq \frac{1}{4}$, else we obtain the estimator with desired rate of convergence $O_p\left(N^{-\frac{5}{2}}\right)$ and the normal asymptotic distribution. Thus, we have a faster rate of convergence for ERE of frequency rate than the usual rate of convergence which is $O_p\left(N^{-\frac{1}{2}}\right)$. Also note that the correction terms given in (3) and (4), play important role in obtaining the desired rate of convergence. This becomes clear by looking at the proof of Theorem 1. Later, we see how do we get an estimator of order $O_p\left(N^{-2-\delta}\right)$ in practice by proposing the robust efficient algorithm.

It has been established in [20] that WLSE of β^0, $\hat{\beta}$ follows normal distribution asymptotically as follows:

$$N^{\frac{5}{2}}\left(\hat{\beta} - \beta^0\right) \xrightarrow{d} \mathcal{N}\left(0, \sigma^2\left(\frac{c_1^2 d_5 + c_3^2 d_1 - 2c_1 c_3 d_3}{2\left(c_1 c_5 - c_3^2\right)^2 \left|A^0\right|^2}\right)\right), \text{provided, } c_1 c_5 - c_3^2 \neq 0. \tag{8}$$

It has been established in [18] that LSE of β^0, $\hat{\hat{\beta}}$ follows normal distribution asymptotically as follows:

$$N^{\frac{5}{2}}\left(\hat{\hat{\beta}}-\beta^0\right)\xrightarrow{d}\mathcal{N}\left(0,\sigma^2\frac{45}{8\left|A^0\right|^2}\right). \tag{9}$$

In subsection 4.1, we provide the comparison of the asymptotic variances of WLSE, ERE and LSE of β^0 for different weight functions.

Now we put forward a robust efficient algorithm based on the estimator obtained by minimizing the reduced weighted residual sum of squares, i.e., $Q\left(\hat{A}_R(\beta),\hat{A}_I(\beta),\beta\right)$ over the grid, $\beta_{grid}=\frac{2\pi l}{N^2}, l=1,\ldots,N^2-1$, which has the convergence rate $O_p\left(N^{-2}\right)$. That is, the initial estimator for the proposed algorithm is obtained as follows:

$$\tilde{\beta}=\arg\min_{\beta_{grid}} Q\left(\hat{A}_R(\beta),\hat{A}_I(\beta),\beta\right). \tag{10}$$

Here, $Q\left(\hat{A}_R(\beta),\hat{A}_I(\beta),\beta\right)$ is same as defined in [20].

Now, we propose the robust efficient algorithm which is a finite step iterative procedure. We begin with the initial estimator $\tilde{\beta}$, obtained from (10), and iteratively update it at every step using the following finite-step algorithm. The m^{th} step estimator $\breve{\beta}^{(m)}$ is computed from the $(m-1)^{th}$ step estimator $\breve{\beta}^{(m-1)}$, using the following formula:

$$\breve{\beta}^{(m)}=\breve{\beta}^{(m-1)}+\frac{3c_1}{(3c_5-c_3)N_m^4}\,\mathrm{Im}\left[\frac{P_{N_m}^{\breve{\beta}^{(m-1)}}}{Q_{N_m}^{\breve{\beta}^{(m-1)}}}\right],\text{provided, } 3c_5-c_3\neq 0, \tag{11}$$

where N_m represents the sample size at the m^{th} step. Here, $P_{N_m}^{\breve{\beta}^{(m-1)}}$ and $Q_{N_m}^{\breve{\beta}^{(m-1)}}$ can be obtained by replacing N and β with N_m and $\breve{\beta}^{(m-1)}$ in P_N^{β} and Q_N^{β}, respectively. In the following steps, $\lceil x\rceil$ represents the least integer function, which gives the smallest integer greater than or equal to x. We repeatedly apply the formula given in (11), selecting N_m at each step as follows:

Step 1: Choose $N_1=\lceil N^{\frac{8}{9}}\rceil$ for $m=1$ and $\breve{\beta}^{(0)}=\tilde{\beta}$, the initial estimator. Note that $\tilde{\beta}-\beta^0=O_p\left(N^{-2}\right)=O_p\left(N_1^{-2-\frac{1}{4}}\right)$. Then applying part (a) of theorem 1, we have;

$$\breve{\beta}^{(1)}-\beta^0=O_p\left(N_1^{-2-\frac{1}{2}}\right)=O_p\left(N^{-\frac{20}{9}}\right).$$

Step 2: Choose $N_2=\lceil N^{\frac{80}{81}}\rceil$ for $m=2$ and calculate $\breve{\beta}^{(2)}$ from $\breve{\beta}^{(1)}$ using (11). Since $\breve{\beta}^{(1)}-\beta^0=O_p\left(N^{-\frac{20}{9}}\right)=O_p\left(N_2^{-2-\frac{1}{4}}\right)$, therefore applying part (a) of theorem 1, we have;

$$\breve{\beta}^{(2)}-\beta^0=O_p\left(N_2^{-2-\frac{1}{2}}\right)=O_p\left(N^{-\frac{200}{81}}\right).$$

Step 3: Choose $N_3=N$ for $m=3$ and calculate $\breve{\beta}^{(3)}$ from $\breve{\beta}^{(2)}$ using (11). Since $\breve{\beta}^{(2)}-\beta^0=O_p\left(N^{-\frac{200}{81}}\right)=O_p\left(N_3^{-2-\frac{38}{81}}\right)$ and $\frac{38}{81}>\frac{1}{4}$, therefore, using part (b) of theorem 1, we have

$$N^{\frac{5}{2}}\left(\breve{\beta}-\beta^0\right)\xrightarrow{d}\mathcal{N}\left(0,\frac{\sigma^2\left(d_1+9d_5-6d_3\right)}{2\left(3c_5-c_3\right)^2\left|A^0\right|^2}\right).$$

Notice that the initial estimator has the order $O_p\left(N^{-2}\right)$ and to start the algorithm we need an estimator with order $O_p\left(N^{-2-\delta}\right)$. Therefore, we take a fraction of sample size at the initial step of this algorithm and gradually proceed towards the complete sample size. Thus, in three steps this algorithm provides the ERE of frequency rate with desired rate of convergence same as that of the WLSEs, LSEs and EFFEs.

Once $\breve{\beta}$ is determined, the amplitude parameters can be estimated using linear regression. Note that $\breve{A} = \breve{A}_R + i\breve{A}_I$, where,

$$\breve{A}_R = \frac{1}{Nc_1}\sum_{t=1}^{N} w\left(\frac{t}{N}\right)\left(y_R(t)\cos\left(\breve{\beta}t^2\right) + y_I(t)\sin\left(\breve{\beta}t^2\right)\right), \tag{12}$$

$$\breve{A}_I = \frac{1}{Nc_1}\sum_{t=1}^{N} w\left(\frac{t}{N}\right)\left(y_I(t)\cos\left(\breve{\beta}t^2\right) - y_R(t)\sin\left(\breve{\beta}t^2\right)\right). \tag{13}$$

Here, $y_R(t)$ and $y_I(t)$ represent the real and imaginary parts of the data $y(t)$, (2). The following theorem establishes the asymptotic distribution of the EREs for the amplitude parameters.

Theorem 2. *If Assumptions 1 and 2 are satisfied, asymptotic distributions of $\breve{A}_R$ and $\breve{A}_I$ are provided as follows:*

$$N^{\frac{1}{2}}\left(\breve{A}_R - A_R^0\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{\sigma^2}{2c_1^2}\left(d_1 + \frac{(d_1 - 6d_3 + 9d_5)\,c_3^2 A_I^{0^2}}{(3c_5 - c_3)^2\left|A^0\right|^2}\right)\right);$$

$$N^{\frac{1}{2}}\left(\breve{A}_I - A_I^0\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{\sigma^2}{2c_1^2}\left(d_1 + \frac{(d_1 - 6d_3 + 9d_5)\,c_3^2 A_R^{0^2}}{(3c_5 - c_3)^2\left|A^0\right|^2}\right)\right), \text{ provided, } 3c_5 - c_3 \neq 0.$$

Proof. See Appendix A.

Theorem 2 states that the EREs of the amplitude parameters follow normal distribution asymptotically with mean 0, and variances $\frac{\sigma^2}{2c_1^2}\left(d_1 + \frac{(d_1-6d_3+9d_5)c_3^2 A_I^{0^2}}{(3c_5-c_3)^2\left|A^0\right|^2}\right)$ and $\frac{\sigma^2}{2c_1^2}\left(d_1 + \frac{(d_1-6d_3+9d_5)c_3^2 A_R^{0^2}}{(3c_5-c_3)^2\left|A^0\right|^2}\right)$, respectively. These estimators converge at the rate $O_p\left(N^{-\frac{1}{2}}\right)$. This rate of convergence is same as the usual rate of convergence and is slower than that of the frequency rate. Consequently, the estimator of the frequency rate parameter has a smaller asymptotic variance than that of the amplitude parameters.

It has been established in [20] that the WLSE of A_R^0 and A_I^0, i.e., $\hat{A}_R$ and $\hat{A}_I$ follow normal distribution asymptotically as follows:

$$N^{\frac{5}{2}}\left(\hat{A}_R - A_R^0\right) \xrightarrow{d} \mathcal{N}\left(0, \sigma^2\left(\frac{\left|A^0\right|^2 c_1^2 d_1 c_5^2 + A_R^{0^2} c_3^2 d_1\left(c_3^2 - 2c_1c_5\right) + A_I^{0^2} c_1^2 c_3\left(c_3 d_5 - 2c_5 d_3\right)}{2c_1^2\left(c_1c_5 - c_3^2\right)^2\left|A^0\right|^2}\right)\right);$$

$$N^{\frac{5}{2}}\left(\hat{A}_I - A_I^0\right) \xrightarrow{d} \mathcal{N}\left(0, \sigma^2\left(\frac{\left|A^0\right|^2 c_1^2 d_1 c_5^2 + A_I^{0^2} c_3^2 d_1\left(c_3^2 - 2c_1c_5\right) + A_R^{0^2} c_1^2 c_3\left(c_3 d_5 - 2c_5 d_3\right)}{2c_1^2\left(c_1c_5 - c_3^2\right)^2\left|A^0\right|^2}\right)\right);$$

provided, $c_1c_5 - c_3^2 \neq 0$.

It has been established in [18] that the LSE of A_R^0 and A_I^0, i.e., $\hat{\hat{A}}_R$ and $\hat{\hat{A}}_I$ follow normal distribution asymptotically as follows:

$$N^{\frac{5}{2}}\left(\hat{\hat{A}}_R - A_R^0\right) \xrightarrow{d} \mathcal{N}\left(0, \sigma^2\left(\frac{1}{2} + \frac{5A_I^{0^2}}{8\left|A^0\right|^2}\right)\right).$$

$$N^{\frac{5}{2}}\left(\hat{\hat{A}}_I - A_I^0\right) \xrightarrow{d} \mathcal{N}\left(0, \sigma^2\left(\frac{1}{2} + \frac{5A_R^{0^2}}{8\left|A^0\right|^2}\right)\right).$$

Thus, Theorem 1 and Theorem 2 provides EREs for the parameters of the model (2) with the desired statistical properties. Note that results of Theorem 1 and Theorem 2 are identical with that of the EFFEs for $w(s) = 1$, please see [19].

3. MCEC Model

In this section, we introduce a sequential robust efficient algorithm for the MCEC model (1). Along with the Assumptions 2 and 1, we have the following assumption on the parameters that are needed to prove the asymptotic properties of the estimators. The true parameter vector, $\left(A_{R1}^0, A_{I1}^0, \beta_1^0, \ldots, A_{Rp}^0, A_{Ip}^0, \beta_p^0\right)$ for the model (1), is denoted by $\boldsymbol{\mu}^0$.

Assumption 3. *The $\boldsymbol{\mu}^0$ is an interior point of the parameter space $\boldsymbol{\Theta}^{(p)}$; $\boldsymbol{\Theta} = [-S, S]\times[-S, S]\times[0, 2\pi]$. Additionally, the frequency rates β_k^0's are distinct for $k = 1, \cdots, p$.*

Without loss of generality, to avoid the identifiability issue, we assume that A_k^0's satisfy the following relationship:

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

3.1. Sequential Robust Efficient Algorithm

We introduce a theorem that produces a frequency rate estimator with the identical rate of convergence as that of the WLSEs, LSEs and EFFEs in this section. First, we propose the updating scheme. The theoretical foundation of the scheme is supported by Theorem 3.

If $\tilde{\beta}_k$ is a consistent estimator of β_k^0, for $k = 1, \cdots, p$, such that $\tilde{\beta}_k - \beta_k^0 = O_p\left(N^{-2-\delta}\right)$, for some $\delta > 0$, then updated estimator of β_k^0 is computed as follows:

$$\breve{\beta}_k = \tilde{\beta}_k + \frac{3c_1}{(3c_5 - c_3)N^4}\,\mathrm{Im}\left(\frac{P_N^{\tilde{\beta}_k}}{Q_N^{\tilde{\beta}_k}}\right), \text{ provided, } 3c_5 - c_3 \neq 0, \text{ where,} \tag{14}$$

$$P_N^{\tilde{\beta}_k} = \sum_{t=1}^{N} y_k(t)\, w\left(\frac{t}{N}\right)\left(t^2 - \frac{N^2}{3}\right)e^{-i\tilde{\beta}_k t^2}, \tag{15}$$

$$Q_N^{\tilde{\beta}_k} = \sum_{t=1}^{N} y_k(t)\, w\left(\frac{t}{N}\right)e^{-i\tilde{\beta}_k t^2}. \tag{16}$$

Here, $y_k(t)$ is the data at the k^{th} step of the sequential algorithm and it is defined as follows:

$$\begin{aligned} &y_k(t) = y_{k-1}(t) - \breve{A}_{k-1}e^{i\breve{\beta}_{k-1}t^2};\\ &t = 1, \ldots, N;\ k = 2, \ldots, p;\ y_1(t) = y(t). \end{aligned} \tag{17}$$

We initialize the algorithm with a consistent estimator $\tilde{\beta}_k$ and then update the estimator using (14).

Further, we introduces a sequential robust efficient algorithm where we obtain the initial estimator using (10). In the sequential procedure, the components of the model (1) are estimated step by step, as described in [18]. We begin by estimating the first component of the model (1), as described in subsection 2.1, and then eliminate its effect from the original data. Next, we use the updated data to estimate the second chirp component, continuing this process until all p-components have been estimated. The following theorem provides the results of the sequential EREs of the parameters at each step.

Theorem 3. *If* $\tilde{\beta}_k - \beta_k^0 = O_p\left(N^{-2-\delta}\right)$*, for* $k = 1, \cdots, p$*, where* $\delta \in \left(0, \frac{1}{2}\right]$*, then*

(a) $\breve{\beta}_k - \beta_k^0 = O_p\left(N^{-2-2\delta}\right)$ *if* $\delta \leq \frac{1}{4}$.

(b) $N^{\frac{5}{2}}\left(\breve{\beta}_k - \beta_k^0\right) \xrightarrow{d} \mathcal{N}\left(0, \sigma^2\left(\frac{d_1+9d_5-6d_3}{2(3c_5-c_3)^2\left|A_k^0\right|^2}\right)\right)$*, provided,* $3c_5 - c_3 \neq 0$ *if* $\delta > \frac{1}{4}$*, where,* c_k*'s and* d_k*'s are obtained from equations* (5) *and* (6)*, respectively.*

Proof. See Appendix B.

Theorem 3 establishes that, for each component of the model, if we start with an estimator of order $O_p\left(N^{-2-\delta}\right)$, then the updated estimator has order $O_p\left(N^{-2-2\delta}\right)$ whenever $\delta \leq \frac{1}{4}$. Otherwise, the updated estimator is asymptotically normally distributed with mean 0 and variance $\sigma^2\left(\frac{d_1+9d_5-6d_3}{2(3c_5-c_3)^2|A_k^0|^2}\right)$. The sequential method works because of the asymptotic orthogonality of the components in model (1). This asymptotic orthogonality can be established using the Lemma 2 given in [12]. Consequently, the multiple-component case achieves the same rate of convergence as the one-component case.

Theorem 4. *If Assumptions 1, 2 and 3 are satisfied, the asymptotic distributions of* $\breve{A}_{Rk}$ *and* $\breve{A}_{Ik}$ $\forall k = 1, \ldots, p$*, are given by:*

$$N^{\frac{1}{2}}\left(\breve{A}_{Rk} - A_{Rk}^0\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{\sigma^2}{2c_1^2}\left(d_1 + \frac{(d_1 - 6d_3 + 9d_5)\, c_3^2 A_{Ik}^{0^2}}{(3c_5 - c_3)^2\left|A_k^0\right|^2}\right)\right);$$

$$N^{\frac{1}{2}}\left(\breve{A}_{Ik} - A_{Ik}^0\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{\sigma^2}{2c_1^2}\left(d_1 + \frac{(d_1 - 6d_3 + 9d_5)\, c_3^2 A_{Rk}^{0^2}}{(3c_5 - c_3)^2\left|A_k^0\right|^2}\right)\right),\ \textit{given},\ 3c_5 - c_3 \neq 0.$$

Here, $\breve{A}_{Rk}$ *and* $\breve{A}_{Ik}$ *represent the real and imaginary parts of* $\breve{A}_k$ *at the* k^{th} *step, respectively. They can be derived in a similar manner to* $\breve{A}_R$ (12) *and* $\breve{A}_I$ (13)*.*

Proof. See Appendix B.

Theorem 4 states that the sequential EREs of the amplitude parameters follow normal distribution asymptotically and have the rate of convergence $O_p\left(N^{-\frac{1}{2}}\right)$. Again the rate of convergence is same as the usual one and is slower than the sequential EREs of the frequency rate parameters.

Theorem 5. *If Assumptions 1, 2 and 3 hold true, then*

$$\breve{\boldsymbol{\theta}}_k \xrightarrow{p} \boldsymbol{\theta}_k^0 \text{ as } N \to \infty, \forall k = 1, \cdots, p.$$

Here $\xrightarrow{p}$ *stands for convergence in probability.*

Proof. See Appendix B.

Theorem 5 means that the proposed sequential EREs are consistent for the true values of the parameters.

Theorem 6. *If Assumptions 1, 2 and 3 hold true, then*

$$\breve{A}_{R(p+k)} \xrightarrow{p} 0, \breve{A}_{I(p+k)} \xrightarrow{p} 0 \text{ as } N \to \infty, \forall k = 1, 2, \cdots.$$

Proof. See Appendix B.

Theorem 6 proves that for the number of components greater than p the proposed sequential EREs of the amplitude parameters converge to 0 as sample size increases. Note that p is assumed to be known in advance in this paper.

The above theorems indicate that the sequential EREs are consistent and have the identical rate of convergence as the WLSEs, LSEs and EFFEs. They can be computed easily in finite number of iterations. Thus, sequential EREs can be utilized to estimate the unknown parameters of the model (1).

4. Simulation Studies

In this section, we observe the numerical performance of the proposed EREs and sequential EREs for the OCEC and the MCEC models, respectively. Our goal is to highlight the significance of choosing the appropriate weight function for the proposed method using the simulation experiments performed in this section. We also compare the performance of these estimators with the other well-known estimators that exist in the literature. We conduct numerical experiments for several different combinations of sample size N, error variance σ^2, weight function, outlier percentage p_{out} and outlier error variance σ^2_{out}.

4.1. OCEC Model Simulations

In this subsection, we compare the performance of the proposed ERE (2.1) with LADE, HME, WLSE [20], LSE [18], EFFE [19] and CPFE [23] method for an OCEC model (2) based on the MSE. We generate the data from OCEC model (2) with $A^0 = 8, \beta^0 = 1, N = 101, 301, 501, 701, 901$ and $\sigma^2 = 0.25, 0.5$. Here, $\epsilon(t)$ has the following form:

$$\epsilon(t) \sim \begin{cases} \mathcal{CN}(0, 30) & \text{if } t_1 \le t \le t_2 \\ \mathcal{CN}\left(0, \sigma^2\right) & \text{Otherwise,} \end{cases} \tag{18}$$

where, $t_1 = \frac{N}{2}(1 - p_{out}) + 1$, $t_2 = \frac{N}{2}(1 + p_{out})$, p_{out} denotes the percentage of outliers in the data and $\sigma^2 = 0.25$ or 0.5. That is the data is contaminated from the middle with p_{out} outliers. Here, $\mathcal{CN}$ denotes the circularly symmetric complex-valued normal distribution. We have contaminated the data with 10, 20, 30 and 40 percent of the middle part of the data. We repeat the simulations 1000 times for these varying combinations of parameters. We compute the estimate

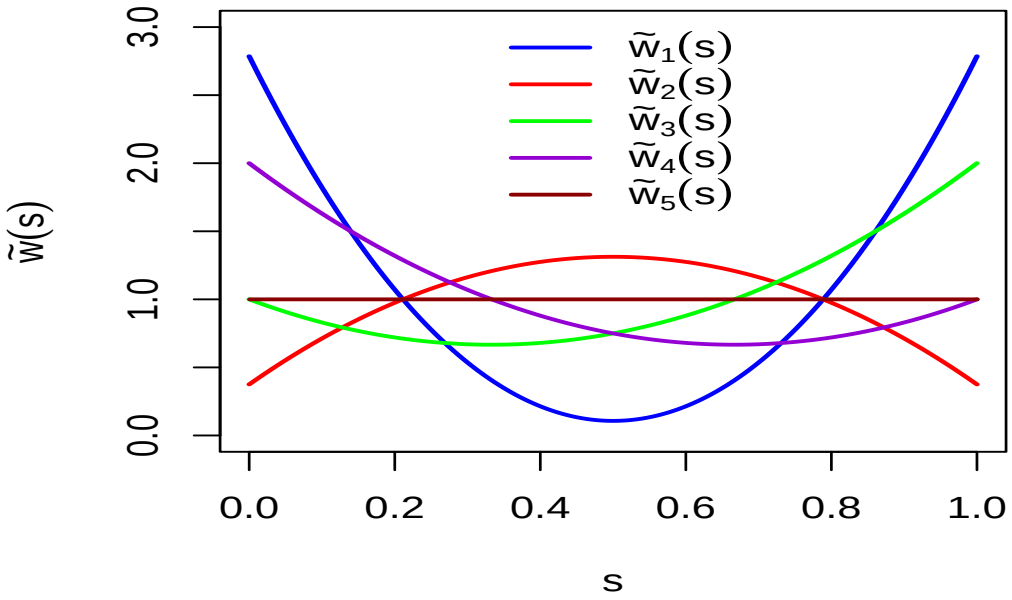


Figure 1: Plot of normalized weight functions.

of the parameters using ERE discussed in subsection 2.1, LADE, HME, WLSE, LSE, EFFE and CPFE method. We compute EREs and WLSEs based on the weight functions given as follows: $w_1(s) = s^2 - s + 0.26$, $w_2(s) = -s^2 + s + \frac{1}{10}$, $w_3(s) = \frac{2s^2}{3} + \frac{(1-s)^2}{3}$, $w_4(s) = \frac{s^2}{3} + \frac{2(1-s)^2}{3}$ and $w_5(s) = 1,\ s \in [0,1]$. Fig. 1 depicts the plot of these weight functions in the normalized form. The normalized form of the weight function is computed as follows:

$$\tilde{w}(s) = \frac{w(s)}{\int_0^1 w(s)\,ds};\ s \in [0,1]\,. \tag{19}$$

Notice that w_1 provides less weight in the middle portion of the data set, w_2 provides less weight on the edges, w_3 provides less weight on the left side of the data set, w_4 provides less weight on the right side of the data set than the other regions and w_5 provides equal weight to the whole data set. The weight function which provides less weight to regions of the data set where outliers are present, performs better than the other weight functions. In this case outliers lie in the middle portion of the data, therefore w_1 performs better in terms of the MSE and w_2 performs worse as it puts the largest weight in the middle portion.

We calculate the estimates of the frequency rate for all these methods and report $-10\log_{10}(MSE)$ for the best and worst performing robust estimators, namely, W1ERE, W2ERE, W1LSE, W2LSE, and LADE, HME, LSE, EFFE and CPFE. Note that we have to do 3-D optimization to compute LADE and HME, while to obtain WLSE, LSE and CPFE, we have to do 1-D optimization. We use Nelder-Mead algorithm to compute these estimates. For the initialization of the algorithm, we randomly choose the initial values among the neighbourhood of the true parameter values. To obtain EREs, we use the proposed three-step iterative algorithm as discussed in subsection 2.1. This indicates that the EREs have less computational complexity than the other approaches under discussion. We first calculate MSEs of these frequency rate estimates when there are no

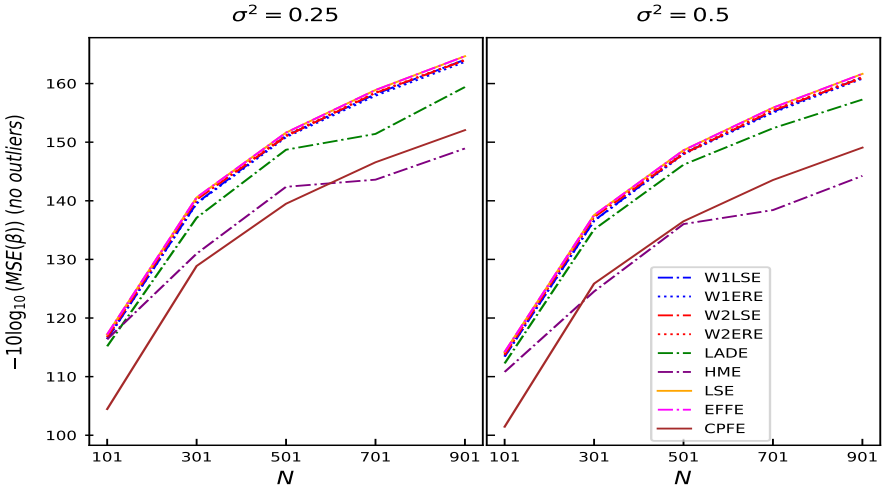


Figure 2: $-10\log_{10}(MSE)$ of estimates of frequency rate of the model (2) with $A^0 = 8$, $\beta^0 = 1$ for different error variances versus sample size.

outliers present in the data and report $-10\log_{10}(MSE)$ in Fig. 2. This figure displays that, in the absence of outliers, LSEs and EFFEs perform best among all the discussed estimators. The performance of the WLSEs and EREs with different weights is at par and close with that of the LSEs. Now we calculate MSEs of these frequency rate estimates in presence of the outliers and report $-10\log_{10}(MSE)$ in Fig. 3. Fig. 3 shows $-10\log_{10}(MSE)$ of estimates of frequency rate versus sample size for various p_{out} and σ^2 in the data. We also compute asymptotic variances of LSE, WLSE and ERE of the frequency rate estimates ("AVAR(β)") in presence of the outliers and report $-10\log_{10}(AVAR)$ in Fig. 4. The observations from these figures are as follows:

1. The $-10\log_{10}(MSE)$ increases as the N increases, this depicts the consistency of these estimators.
2. The $-10\log_{10}(MSE)$ of all the estimators decreases as σ^2 increases, i.e., as the signal-to-noise ratio (SNR) decreases $-10\log_{10}(MSE)$ of all the estimators decreases.
3. The $-10\log_{10}(MSE)$ of W1ERE and W1LSE are greater than LADE, LSE, EFFE, W2ERE, HME and CPFE in Fig. 3.
4. Fig. 4 depict that the asymptotic variance of W1ERE is greater than all the other estimators, whereas it is least for W2ERE.
5. Thus, ERE with an appropriate weight function provides the estimator with less cumbersome computations.

4.2. MCEC Model Simulations

In this subsection, we compare the performance of the proposed sequential ERE (3.1) with LADE, HME, sequential WLSE [20], sequential LSE [18], sequential EFFE [19] and PCPFE [30] method for an MCEC model (1) based on the MSE. We generate the data 1000 times from the model (1) with $p = 2$ and $A_1^0 = 10, \beta_1^0 = 1.2, A_2^0 = 8, \beta_2^0 = 1$. The error structure is the same as that taken in (18). The N, σ^2, p_{out} and σ_{out}^2 are same as those taken for OCEC

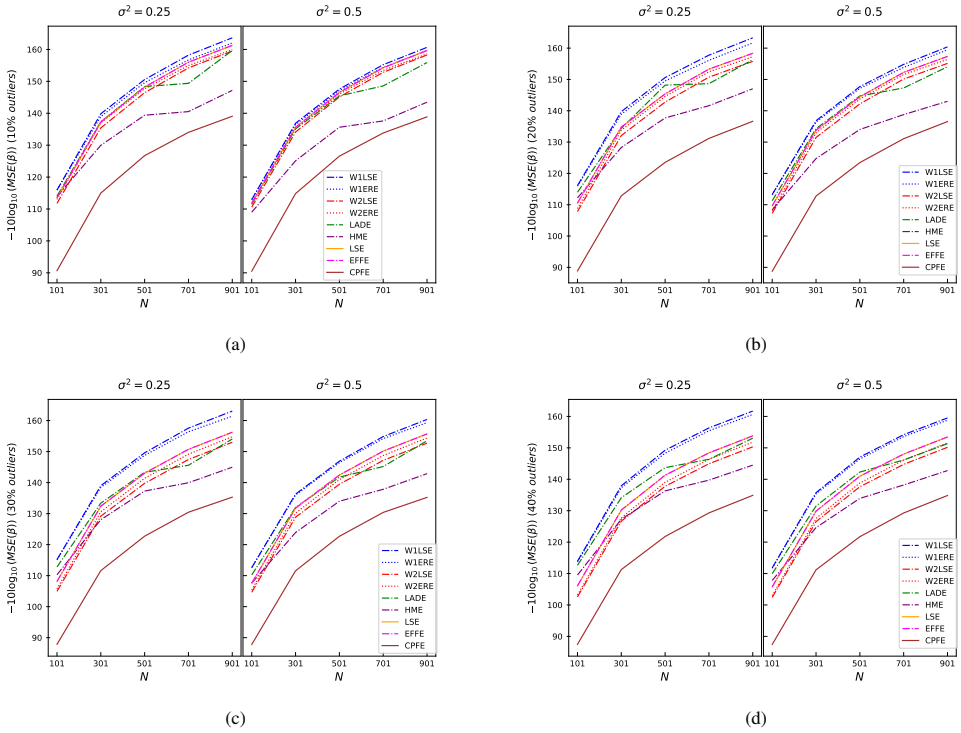


Figure 3: $-10\log_{10}(MSE)$ of estimates of frequency rate of the model (2) with $A^0 = 8, \beta^0 = 1$ and (a) 10% outliers (b) 20% outliers (c) 30% outliers (d) 40% outliers with $\sigma^2_{out} = 30$ for different error variances versus sample size.

simulation study in subsection 4.1. We estimate parameters of the model using the sequential ERE discussed in subsection 3.1, LADE, HME, sequential WLSE, sequential LSE, sequential EFFE, and PCPFE. Note that we have to do 6-D optimization to compute LADE and HME, while to obtain sequential WLSE, sequential LSE and PCPFE, we have to do two 1-D optimization. We use Nelder-Mead algorithm to compute these estimates. For the initialization of the algorithm, we randomly choose the initial values among the neighbourhood of the true parameter values. To obtain sequential ERE, we use the proposed three-step iterative algorithm discussed in subsection 3.1. This indicates that the sequential ERE have less computational complexity than the other approaches under discussion. We calculate MSEs of these frequency rate estimates and report $-10\log_{10}(MSE)$ in Figs. 5 and 6. Following are the observations from these figures:

1. The $-10\log_{10}(MSE)$ increases as the N increases, this depicts the consistency of these estimators.
2. The $-10\log_{10}(MSE)$ of all the estimators decreases as σ^2 increases.
3. The $-10\log_{10}(MSE)$ of sequential W1ERE and sequential W2ERE are at par with sequential W1LSE and sequential W2LSE, respectively.
4. The $-10\log_{10}(MSE)$ of sequential W1ERE and sequential W1LSE are greater than sequential LSE, sequential EFFE, sequential W2ERE, sequential W2LSE, LADE, HME and PCPFE in Figs. 5 and 6.

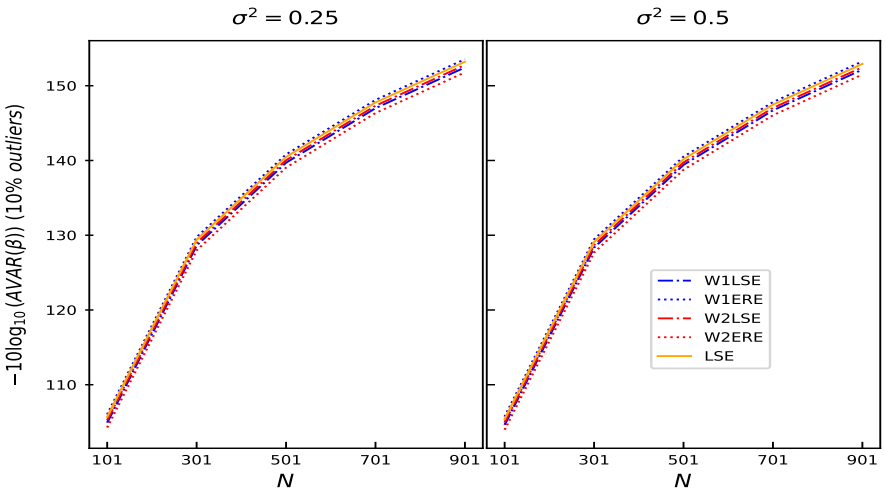
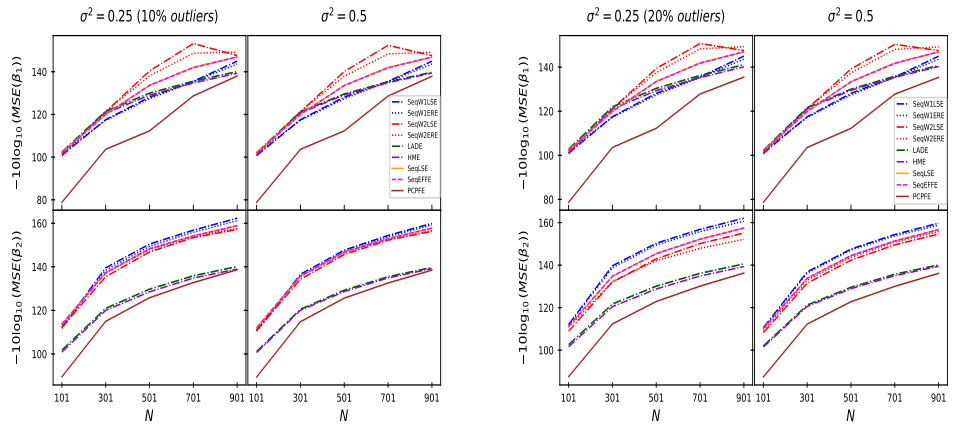


Figure 4: $-10\log_{10}(AVAR)$ of estimates of frequency rate of the model (2) with $A^0 = 8, \beta^0 = 1$ and 10% outliers with $\sigma^2_{out} = 30$ for different error variances versus sample size.

5. Therefore, for an appropriate choice of weight function sequential EREs provide estimators with less cumbersome computations than the other discussed approaches.

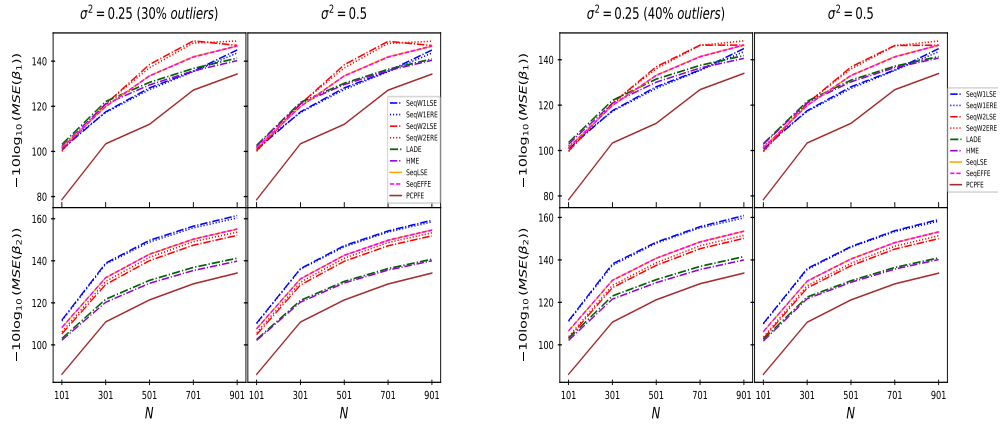
4.3. Time Comparison

In this subsection, we compare the performance of the proposed sequential ERE (3.1) with LADE, HME, sequential WLSE, sequential LSE, sequential EFFE and PCPFE method for an MCEC model (1) based on the computation time. We generate the data 1000 times from the model (1) with $p = 10$, and parameters $A^0_1 = 20, \beta^0_1 = 0.5, A^0_2 = 18, \beta^0_2 = 1.4, A^0_3 = 16, \beta^0_3 = 1.1, A^0_4 = 14, \beta^0_4 = 0.51, A^0_5 = 12, \beta^0_5 = 1.7, A^0_6 = 10, \beta^0_6 = 0.8, A^0_7 = 8, \beta^0_7 = 2.3, A^0_8 = 6, \beta^0_8 = 2.6, A^0_9 = 4, \beta^0_9 = 2, A^0_{10} = 2, \beta^0_{10} = 2.9$. The error structure is the same as that taken in (18). The $N = 101, 301, 501, 701, 901, \sigma^2 = 0.25, p_{out} = 10\%$ and $\sigma^2_{out} = 30$. The outliers lie in the middle region of the data. We compute average time involved in computing the estimates of the model parameters over 1000 simulations for the estimation techniques, namely, sequential W1ERE in subsection 3.1, LADE, HME, sequential WLSE, sequential LSE, sequential EFFE, and PCPFE. We randomly choose the initial values among the neighbourhood of the true parameter values. Fig. 7 (a) depicts the computation time in estimating the associated frequency rates and amplitude parameters versus sample size. Fig. 7 (b), represents the computation time ratio with respect to the computation time involved in sequential EFFE method. From this figure, it is clear that, in comparison to the other discussed approaches, sequential EFFE requires the least amount of computation time, whereas HME requires the most. It is also noticeable that sequential ERE requires less computing time than sequential WLSE, HME and LADE. This indicates that sequential ERE takes less time to compute, and with appropriate weight function, this performs better in terms of MSE than HME and LADE.



(a) (b)

Figure 5: $-10\log_{10}(MSE)$ of frequency rates estimates of the model (1) with $A_1^0 = 10, \beta_1^0 = 1.2, A_2^0 = 8, \beta_2^0 = 1$ and (a) 10% (b) 20% outliers with $\sigma_{out}^2 = 30$ for different error variances versus sample size.

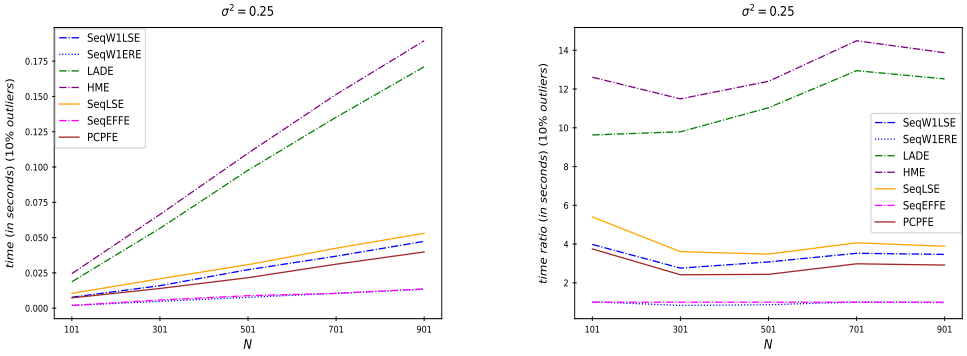


(a) (b)

Figure 6: $-10\log_{10}(MSE)$ of frequency rates estimates of the model (1) with $A_1^0 = 10, \beta_1^0 = 1.2, A_2^0 = 8, \beta_2^0 = 1$ and (a) 30% (b) 40% outliers with $\sigma_{out}^2 = 30$ for different error variances versus sample size.

4.4. When Frequency Rates are Close

We generate the data 1000 times from the MCEC model (1) with $p = 2$, and parameters $A_1^0 = 10, A_2^0 = 8, \beta_2^0 = 1$ and $\beta_1^0 = \beta_2^0 + \delta$, where, $\delta = 10^{-3}, 2\times10^{-3}, 3\times10^{-3}, 5\times10^{-3}, 2\times10^{-2}$ and 6×10^{-2}. The error structure is the same as that taken in (18). The $N = 901$, $\sigma^2 = 0.5$, $p_{out} = 20\%$ and $\sigma_{out}^2 = 30$. The outliers lie in the middle region of the data. In this case, we vary β_1^0 such that it becomes closer to $\beta_2^0 = 1$ while keeping fixed other parameters. We evaluate the performance of the sequential EREs based on the weight functions taken in Subsection 4.1, LADE, HME, sequential WLSEs, sequential LSEs, sequential EFFEs, and PCPFE, when the



(a) (b)

Figure 7: (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 $p = 10$, $A_1^0 = 20, \beta_1^0 = 0.5$, $A_2^0 = 18, \beta_2^0 = 1.4$, $A_3^0 = 16, \beta_3^0 = 1.1$, $A_4^0 = 14, \beta_4^0 = 0.51$, $A_5^0 = 12, \beta_5^0 = 1.7$, $A_6^0 = 10, \beta_6^0 = 0.8$, $A_7^0 = 8, \beta_7^0 = 2.3$, $A_8^0 = 6, \beta_8^0 = 2.6$, $A_9^0 = 4, \beta_9^0 = 2$, $A_{10}^0 = 2, \beta_{10}^0 = 2.9$ and 10% outliers with $\sigma_{out}^2 = 30$ versus sample size.

frequency rates become close to each other. In this case, we randomly choose the initial values among the neighbourhood of the true parameter values. We calculate MSEs of these estimators. Fig. 8 depicts $-10\log_{10}(MSE)$ of frequency rate estimates versus the distance between the frequency rates. From Fig. 8 it is evident that the sequential WLSEs and sequential EREs perform better than LADE, HME and PCPFE in terms of MSE when frequency rates are close to each other. Also, the MSEs of the sequential LSEs and sequential EFFEs are at par and worse than the MSEs of the sequential EREs and sequential WLSEs with the appropriate weight function. Thus, for real data analysis, we recommend using the proposed sequential WERE with an appropriate weight function.

4.5. When SNR varies

In this subsection, we compare the performance of the proposed ERE (2.1) with LADE, HME, WLSE [20], LSE [18], EFFE [19] and CPFE [23] method for the OCEC model (2) under varying SNR. The SNR is defined as $SNR = 10\log_{10}\left(\frac{\sum_{k=1}^{p}|A_k^0|^2}{p_{out}\sigma_{out}^2+(1-p_{out})\sigma^2}\right)$. Data are generated from the OCEC model (2) with $A^0 = 8, \beta^0 = 1, N = 701, 901$, while SNR varies from -5 to 10. The error structure follows (18). The data is contaminated from the middle with 20% outliers and $\sigma_{out}^2 = 30$. Each simulation is replicated 1000 times for different parameter settings. For estimation, we compute the frequency rate estimates using ERE 2.1, LADE, HME, WLSE, LSE, EFFE and CPFE method, while initial values randomly chosen in the neighbourhood of the true parameter values. Fig. 9 (a) shows $-10\log_{10}(MSE)$ of frequency rate estimates versus SNR and Fig. 9 (b) shows the zoomed in version. From Fig. 9, it is clear that the sequential WLSEs and sequential EREs perform better than LADE, HME and CPFE in terms of MSE for reasonably high SNR value. Also, the MSEs of the sequential LSEs and sequential EFFEs are at par and worse than the MSEs of the sequential EREs and sequential WLSEs with the appropriate weight function when SNR is sufficiently high.

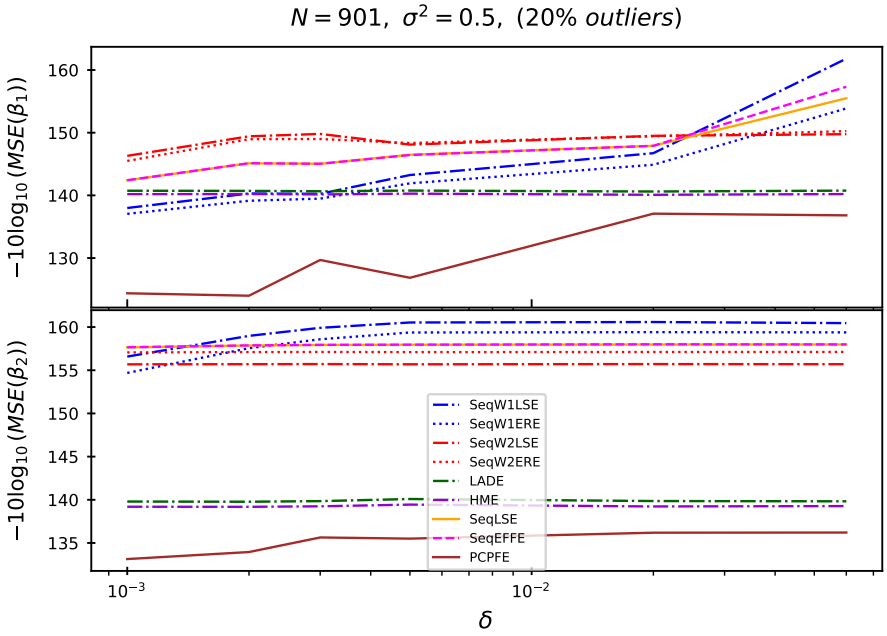
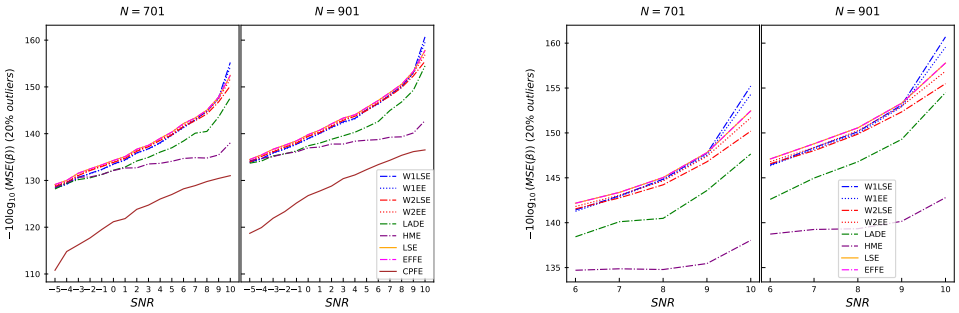


Figure 8: $-10\log_{10}(MSE)$ of frequency rates estimates of the model (1) with $A_1^0 = 10$, $A_2^0 = 8$, $\beta_2^0 = 1$ and $\beta_1^0 = \beta_2^0 + \delta$ with $\delta = 10^{-3}, 2\times10^{-3}, 3\times10^{-3}, 5\times10^{-3}, 2\times10^{-2}$, 6×10^{-2} and 20% outliers with $\sigma^2_{out} = 30$ versus δ.



(a) (b)

Figure 9: $-10\log_{10}(MSE)$ of estimates of frequency rate of the model (2) with $A^0 = 8$, $\beta^0 = 1$, 20% outliers with $\sigma^2_{out} = 30$ for different sample sizes versus SNR.

5. Choice of Weight Function and Data Analysis

5.1. Choice of Weight Function

The simulation results from the previous section demonstrate the importance of choosing the appropriate weight function. As we can see in the above section, the robust estimators whose weight function gives less weight to the region of the data where outliers lie than the other region

outperform the other estimators in terms of MSEs. In this section, we discuss a data driven approach 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. The intuition behind Algorithm

Algorithm 1 An algorithm to choose a weight function

1. Fix p, the number of components in MCEC model (1).
2. For the data of size N, we leave out one observation and then fit the MCEC model (1) to the remaining data of size $(N-1)$ using sequential LSE [18].
3. Using this fitted model, we predict the left out observation and obtain the square of the residual.
4. Repeat above three steps for all N observations.
5. Then we fit a polynomial to the square of the residuals of the predicted observations.
6. Then we use the inverse of the fitted square of the residuals as the weight function.

1 is as follows. In Algorithm 1, we predict each observation of the data by the model fitted to the remaining observations. If the observation that we are predicting is an outlier, then the square of the residuals of the predicted observations would be relatively higher than the other observations. That is how we put less weight on the outliers compared to other observations, by using the inverse of the fitted square of the residuals as the weight function. Moreover, note that in Algorithm 1, we have to estimate $3p$ parameters N number of times, and obtaining them would have been numerically expensive as well as time-consuming if we had used LSE instead of sequential LSE. Thus, using sequential LSE resolves the problem, making it convenient to use practically. We perform the following real-data analysis to illustrate this approach provided in Algorithm 1. The Psuedo-code for the Algorithm 1 is given in Algorithm 2.

5.2. Data Analysis

In this subsection, we consider a EEG data [32]. This data was also analyzed using sequential EFFE [19] and fitted with 41 elementary chirp components. Now, we contaminate the original EEG data by adding some outliers to the original data. We illustrate data analysis for the case when positions of the outliers are chosen randomly, such that it is difficult to differentiate the outlier data from the original data. We contaminate the original data by putting normal distributed errors at 2% random positions of it with mean 0 and variance 20 times that of the error variance of the fitted original data. Fig. 10 shows original EEG data and data with outliers. For the considered EEG data, $N = 256$. We take, $p = 41$, in Algorithm 1 and choose the weight function using the data driven approach discussed in Algorithm 1. Once obtaining the weight function, we fit MCEC model to this contaminated data and analyze it using sequential ERE. We fit the data at 41^{st} component. Thus, we fit an elementary chirp model to the data using sequential ERE based on the weight function obtained using the method discussed in Algorithm 1. Using sequential ERE, fitted data is displayed in Fig. 11. This figure illustrates how the fitted data using sequential ERE encloses the original data. The NWRSS at the 41^{st} component is 1.81.

Now, let us compare this method with the choice of weight function method based on the pool of weight functions given in [20]. We analyze the contaminated EEG data based on the method; by choosing a weight function among differently shaped weight functions. We fit an elementary chirp model to this contaminated data and analyze it using nine sequential ERE and

Algorithm 2 Pseudo-code for selecting the weight function

Require: Data $\{y(t)\}_{t=1}^{N}$, assumed number of components p
for $j = 1, \ldots, N$ **do**
Remove the jth observation to obtain the training sample.
$z^{(1)}(t) = \{y(t) : t \neq j\}$.
for $k = 1, \ldots, p$ **do**
Estimate $(\hat{A}_k, \hat{\beta}_k)$ from $z^{(k)}(t)$ using the sequential LSE as:

$$\left(\hat{A}_k, \hat{\beta}_k\right) = \arg\min \sum_{t=1}^{N-1} \left| z^{(k)}(t) - A e^{i\beta t^2} \right|^2$$

Compute

$$\hat{s}_k(t) = \hat{A}_k e^{i\hat{\beta}_k t^2}.$$

Update

$$z^{(k+1)}(t) = z^{(k)}(t) - \hat{s}_k(t).$$

end for
Predict the deleted observation

$$\hat{y}_j = \sum_{k=1}^{p} \hat{A}_k e^{i\hat{\beta}_k j^2}.$$

Compute $r_j = \left| y_j - \hat{y}_j \right|^2$.
end for
Set $d = 1$.
repeat
Fit a polynomial of degree d to $\{(j, r_j)\}_{j=1}^{N}$.
Obtain fitted squared residuals $\hat{r}_d(j)$.
if $\hat{r}_d(j)$ is not strictly positive for some j **then**
$d \leftarrow d + 1$.
end if
until $\hat{r}_d(j) > 0, \ j = 1, \ldots, N$
Construct the weight function

$$w(j) = \frac{1}{\hat{r}_d(j)},$$

and normalize it as

$$\tilde{w}(j) = \frac{w(j)}{\sum_{i=1}^{N} w(i)}.$$

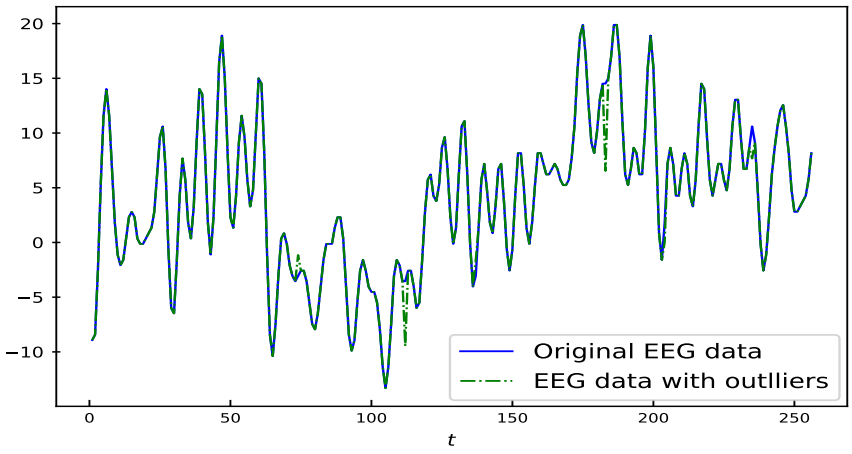


Figure 10: Plot of EEG data and contaminated data.

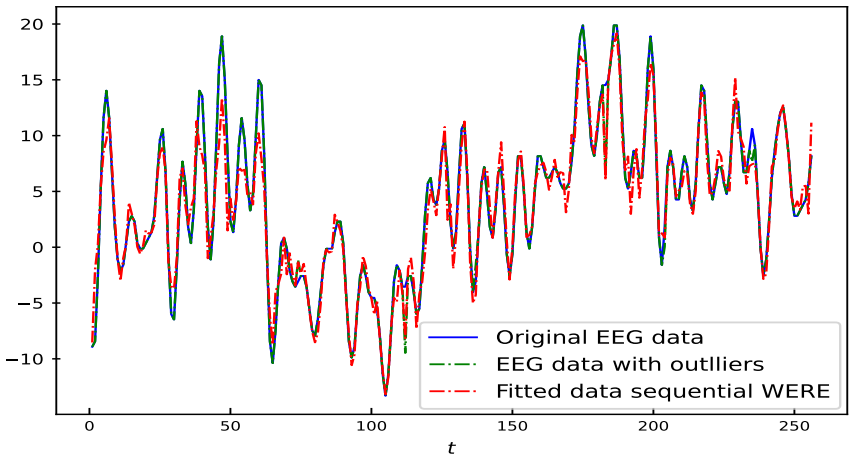


Figure 11: Fitted EEG data using sequential ERE based on Algorithm 1.

sequential EFFE. Since the locations of the outliers are random, we consider a class of weight functions with different shapes, the same as those taken in Section VI of [20]. We fit the data that at 41 elementary chirp components, at which we previously fitted the original data using the sequential EFFE. Table 1 contains the NWRSS at the 41^{st} component, i.e., NWRSS(41), which we calculated using the weight functions mentioned above. From Table 1, it can be observed that the sequential ERE based on $w_9(s)$ give the minimum NWRSS. Thus, we fit an elementary chirp model with 41 components to the data using sequential ERE based on $w_9(s)$. Using sequential

Weight	NWRSS(41)	Weight	NWRSS(41)
w_1	438.23	w_6	581.02
w_2	609.28	w_7	555.36
w_3	444.22	w_8	645.24
w_4	336.16	w_9	304.63
w_5	702.04	w_{10}	7405.91

Table 1: Reporting NWRSS(41) based on different weight functions for the fitted data.

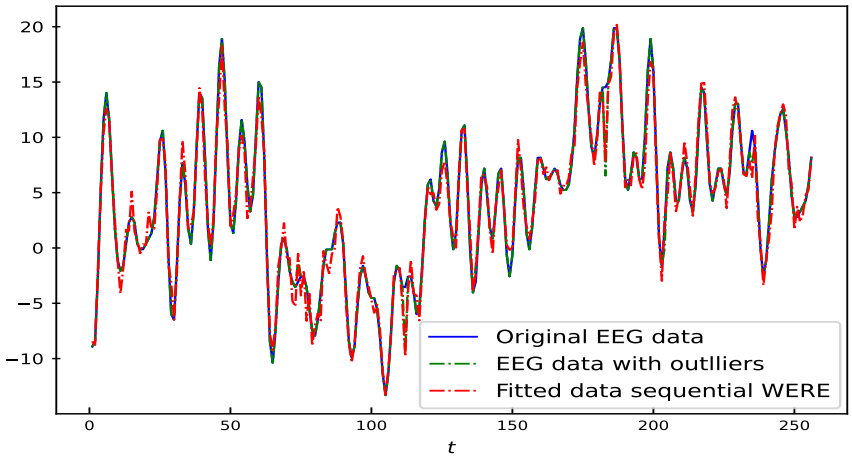


Figure 12: Fitted EEG data using sequential ERE based on the method given Section V of [20].

W9ERE, fitted data is shown in Fig. 12. The fitted data using sequential W9ERE encloses the original data, as shown in this figure. Observe that the NWRSS obtained using Algorithm 1 is 1.81 whereas it is 304.63 for the method based on pool of weight function [20]. The reason behind this huge difference is the weight function chosen from both the methods. This indicates that the Algorithm 1 provides better fit in terms of the NWRSS in comparison to the method based on the pool of weight functions given in [20]. Also notice that in general the weight functions are quite robust because of the way they are chosen. They provide less weight in the regions where the outliers are present than the other regions.

6. Conclusion

In this paper, we propose a sequential robust efficient algorithm to estimate the parameters of the elementary chirp model. The proposed sequential EREs with an appropriate weight function are robust to the outliers. It has been established that the sequential EREs produced by the sequential robust efficient algorithm are consistent and have the identical rate of convergence as sequential WLSEs. The sequential EREs can be easily computed in a finite number of iterations and take less computing time than the other prominent estimators, namely, sequential WLSEs, LADEs and HMEs. In order to practically perform the proposed sequential EREs, we discuss to

choose the weight function based on the cross validation method with a demonstration of a real data analysis.

Acknowledgement

The authors are thankful to the editor, the associate editor and the anonymous reviewers for their critical and constructive suggestions which have helped to improve the manuscript significantly.

Appendix A. OCEC Model

***Proof of Theorem** 1:* Notice that,

$$Q_N^{\tilde{\beta}} = A^0 \sum_{t=1}^{N} w\left(\frac{t}{N}\right) e^{i(\beta^0-\tilde{\beta})t^2} + \sum_{t=1}^{N} \epsilon(t)\, w\left(\frac{t}{N}\right) e^{-i\tilde{\beta}t^2}. \tag{A.1}$$

Using Taylor series and lemma C-2 of [14], we get:

$$\begin{aligned}&\sum_{t=1}^{N} w\left(\frac{t}{N}\right) e^{i(\beta^0-\tilde{\beta})t^2} = \sum_{t=1}^{N} w\left(\frac{t}{N}\right)\left[e^{i(\beta^0-\beta^0)t^2} + i\left(\beta^0-\tilde{\beta}\right)t^2 e^{i(\beta^0-\beta^*)t^2}\right]\\ &= Nc_1 - iO_p\left(N^{1-\delta}\right) + O_p\left(N^{1-\delta}\right), \text{where}, \beta^* \in \left(\beta^0, \tilde{\beta}\right).\end{aligned}$$

Using Taylor series, Lindeberg Feller CLT, lemma 2 of [20] and choosing a large L such that $1 - L\delta < 0$, we get:

$$\begin{aligned}&\sum_{t=1}^{N} \epsilon(t)\, w\left(\frac{t}{N}\right) e^{-i\tilde{\beta}t^2} = \sum_{t=1}^{N} \epsilon(t)\, w\left(\frac{t}{N}\right)\left[e^{-i\beta^0 t^2} + \sum_{l=1}^{L-1} \frac{\left(-i\left(\tilde{\beta}-\beta^0\right)t^2\right)^l}{l!} e^{-i\beta^0 t^2} + \right.\\ &\left.\frac{\left(-it^2\right)^L}{L!}\left(\tilde{\beta}-\beta^0\right)^L e^{-i\beta^* t^2}\right] = O_p\left(N^{\frac{1}{2}}\right), \text{where}, \beta^* \in \left(\beta^0, \tilde{\beta}\right).\end{aligned}$$

Therefore, $Q_N^{\tilde{\beta}}$, (A.1), can be expressed as follows:

$$Q_N^{\tilde{\beta}} = A^0\left(Nc_1 + O_p\left(N^{1-\delta}\right)\right); \delta \in \left(0, \frac{1}{2}\right].$$

Now, consider

$$P_N^{\tilde{\beta}} = A^0 \sum_{t=1}^{N} \left(t^2 - \frac{N^2}{3}\right) w\left(\frac{t}{N}\right) e^{i(\beta^0-\tilde{\beta})t^2} + \sum_{t=1}^{N} \epsilon(t)\, w\left(\frac{t}{N}\right)\left(t^2 - \frac{N^2}{3}\right) e^{-i\tilde{\beta}t^2}. \tag{A.2}$$

Using Taylor series and lemma C-2 of [14], we get:

$$\sum_{t=1}^{N}\left(t^2-\frac{N^2}{3}\right)w\left(\frac{t}{N}\right)e^{i(\beta^0-\tilde{\beta})t^2}=\sum_{t=1}^{N}\left(t^2-\frac{N^2}{3}\right)$$
$$w\left(\frac{t}{N}\right)\left[e^{i(\beta^0-\beta^0)t^2}+e^{i(\beta^0-\beta^0)t^2}it^2\left(\beta^0-\tilde{\beta}\right)+\frac{\left(it^2\left(\beta^0-\tilde{\beta}\right)\right)^2}{2!}e^{i(\beta^0-\beta^*)t^2}\right]$$
$$=N^3\left(c_3-\frac{c_1}{3}\right)+i\left(\beta^0-\tilde{\beta}\right)\left[N^5\left(c_5-\frac{c_3}{3}\right)-O_p\left(N^{5-\delta}\right)\right],\text{where},\beta^*\in\left(\beta^0,\tilde{\beta}\right).$$

Using Taylor series, Lindeberg Feller CLT, lemma 2 of [20] and choosing a large L such that $1-L\delta<0$, we get:

$$\sum_{t=1}^{N}\epsilon(t)\,w\left(\frac{t}{N}\right)\left(t^2-\frac{N^2}{3}\right)e^{-i\tilde{\beta}t^2}=\sum_{t=1}^{N}\epsilon(t)\,w\left(\frac{t}{N}\right)$$
$$\left(t^2-\frac{N^2}{3}\right)\left[e^{-i\beta^0t^2}+e^{-i\beta^0t^2}\sum_{l=1}^{L-1}\frac{\left(-i\left(\tilde{\beta}-\beta^0\right)t^2\right)^l}{l!}+\frac{\left(-it^2\right)^L}{L!}\left(\tilde{\beta}-\beta^0\right)^L e^{-i\beta^*t^2}\right]$$
$$=\sum_{t=1}^{N}\epsilon(t)\,w\left(\frac{t}{N}\right)\left(t^2-\frac{N^2}{3}\right)e^{-i\beta^0t^2}+$$
$$\sum_{l=1}^{L-1}\frac{(-i)^l}{l!}O_p\left(N^{\frac{5}{2}-l\delta}\right)+\frac{(-i)^L}{L!}O_p\left(N^2\right),\text{where},\beta^*\in\left(\beta^0,\tilde{\beta}\right).$$

Therefore, $P_N^{\tilde{\beta}}$, (A.2), becomes

$$P_N^{\tilde{\beta}}=\sum_{t=1}^{N}\epsilon(t)\,w\left(\frac{t}{N}\right)\left(t^2-\frac{N^2}{3}\right)e^{-i\beta^0t^2}+$$
$$A^0i\left(\beta^0-\tilde{\beta}\right)\left[N^5\left(c_5-\frac{c_3}{3}\right)-O_p\left(N^{5-\delta}\right)\right]+A^0N^3\left(c_3-\frac{c_1}{3}\right).$$

Thus, we obtain:

$$\breve{\beta}=\tilde{\beta}+\frac{3c_1}{(3c_5-c_3)N^4}\,\mathrm{Im}\left(\frac{P_N^{\tilde{\beta}}}{Q_N^{\tilde{\beta}}}\right)$$
$$\breve{\beta}=\beta^0+\left(\tilde{\beta}-\beta^0\right)O_p\left(N^{-\delta}\right)+\frac{3c_1}{(3c_5-c_3)N^5}\,\mathrm{Im}\left[\frac{\sum_{t=1}^{N}\epsilon(t)\,w\left(\frac{t}{N}\right)\left(t^2-\frac{N^2}{3}\right)e^{-i\beta^0t^2}}{A^0c_1}\right].$$

Using lemma 1 of [20] and lemma C-2 of [14], the asymptotic variance of

$$\frac{3c_1}{(3c_5-c_3)N^{\frac{5}{2}}}\,\mathrm{Im}\left[\frac{\sum_{t=1}^{N}\epsilon(t)\,w\left(\frac{t}{N}\right)\left(t^2-\frac{N^2}{3}\right)e^{-i\beta^0t^2}}{A^0c_1}\right]$$

is obtained as $\sigma^2\left(\frac{d_1+9d_5-6d_3}{2(3c_5-c_3)^2\left|A^0\right|^2}\right)$. Hence the result of theorem 1 follows using Lindeberg-Feller CLT.

***Proof of Theorem** 2*: Consider, $\breve{A}_R$ as given in equation (12). Now, expanding $\cos\left(\breve{\beta}t^2\right)$ and $\sin\left(\breve{\beta}t^2\right)$ using Taylor series around the point β^0, write $\breve{A}_R$ as:

$$\breve{A}_R = \frac{1}{Nc_1}\sum_{t=1}^{N} w\left(\frac{t}{N}\right)\Big[y_R(t)\left\{\cos\left(\beta^0 t^2\right) - t^2\left(\breve{\beta}-\beta^0\right)\sin\left(\bar{\beta}t^2\right)\right\} + y_I(t)\left\{\sin\left(\beta^0 t^2\right) + t^2\left(\breve{\beta}-\beta^0\right)\cos\left(\bar{\beta}t^2\right)\right\}\Big], \text{where}, \bar{\beta}\in\left(\beta^0,\breve{\beta}\right).$$

Using Theorem 1 and Chebyshev inequality, we get:

$$N^2\left(\breve{\beta}-\beta^0\right)\xrightarrow{p} 0 \text{ as } N\to\infty. \tag{A.3}$$

Using Lindeberg-Feller CLT, Theorem 1, (A.3), lemmas 1 and 2 of [20] and lemma C-2 of [14], we obtain asymptotic distribution of $\breve{A}_R$.The asymptotic distribution of $\breve{A}_I$ can be derived in the similar manner. Hence the result.

Appendix B. MCEC Model

***Proof of Theorem** 3*: First, this result is proved for $\breve{\beta}_1$. Consider:

$$Q_N^{\tilde{\beta}_1} = \sum_{k=1}^{p} A_k^0 \sum_{t=1}^{N} w\left(\frac{t}{N}\right) e^{i(\beta_k^0-\tilde{\beta}_1)t^2} + \sum_{t=1}^{N} w\left(\frac{t}{N}\right)\epsilon(t)\, e^{-i\tilde{\beta}_1 t^2}. \tag{B.1}$$

Using Taylor series expansion and lemma C-2 of [14], $\sum_{t=1}^{N} w\left(\frac{t}{N}\right) e^{i(\beta_1^0-\tilde{\beta}_1)t^2}$ is written as follows:

$$\begin{aligned}\sum_{t=1}^{N} w\left(\frac{t}{N}\right) e^{i(\beta_1^0-\tilde{\beta}_1)t^2} &= \sum_{t=1}^{N} w\left(\frac{t}{N}\right)\left[e^{i(\beta_1^0-\beta_1^0)t^2} + i\left(\beta_1^0-\tilde{\beta}_1\right)t^2 e^{i(\beta_1^0-\beta_1^*)t^2}\right]\\ &= Nc_1 - iO_p\left(N^{1-\delta}\right) + O_p\left(N^{1-\delta}\right), \text{where, } \beta_1^*\in\left(\beta_1^0,\tilde{\beta}_1\right).\end{aligned}$$

For $k = 2,\cdots,p$, using lemmas C-2 and C-3 of [14], Taylor series expansion and choosing L such that $1 - L\delta < 0$, $\sum_{t=1}^{N} w\left(\frac{t}{N}\right) e^{i(\beta_k^0-\tilde{\beta}_1)t^2}$ is expressed as follows:

$$\begin{aligned}\sum_{t=1}^{N} w\left(\frac{t}{N}\right) e^{i(\beta_k^0-\tilde{\beta}_1)t^2} &= \sum_{t=1}^{N} w\left(\frac{t}{N}\right)\Bigg[e^{i(\beta_k^0-\beta_1^0)t^2} + \\ &\sum_{l=1}^{L-1}\frac{\left(-i\left(\tilde{\beta}_1-\beta_1^0\right)t^2\right)^l}{l!}e^{i(\beta_k^0-\beta_1^0)t^2} + e^{i(\beta_k^0-\beta_1^*)t^2}\frac{\left(-i\left(\tilde{\beta}_1-\beta_1^0\right)t^2\right)^L}{L!}\Bigg]\\ &= o_p\left(N^{\frac{1}{2}}\right), \text{where, } \beta_1^*\in\left(\beta_1^0,\tilde{\beta}_1\right).\end{aligned}$$

Now, using Taylor series expansion, Lindeberg Feller CLT, lemma 2 of [20] and choosing a large L such that $1 - L\delta < 0$, we write 2^{nd} term of $Q_N^{\tilde{\beta}_1}$, (B.1), as follows:

$$\begin{aligned}
&\sum_{t=1}^{N} \epsilon(t)\, w\left(\frac{t}{N}\right) e^{-i\tilde{\beta}_1 t^2} = \sum_{t=1}^{N} \epsilon(t)\, w\left(\frac{t}{N}\right)\Big[e^{-i\beta_1^0 t^2} \\
&+ \sum_{l=1}^{L-1} \frac{\left(-i\left(\tilde{\beta}_1 - \beta_1^0\right)t^2\right)^l}{l!} e^{-i\beta_1^0 t^2} + \frac{\left(-it^2\right)^L}{L!} e^{-i\beta_1^* t^2}\left(\tilde{\beta}_1 - \beta_1^0\right)^L\Bigg] \\
&= O_p\left(N^{\frac{1}{2}}\right), \text{where}, \beta_1^* \in \left(\beta_1^0, \tilde{\beta}_1\right).
\end{aligned}$$

Therefore $Q_N^{\tilde{\beta}_1}$, (B.1), can be written as follows:

$$Q_N^{\tilde{\beta}_1} = A_1^0\left(Nc_1 + O_p\left(N^{1-\delta}\right)\right); \delta \in \left(0, \frac{1}{2}\right].$$

Further consider,

$$P_N^{\tilde{\beta}_1} = \sum_{k=1}^{p} A_k^0 \sum_{t=1}^{N} w\left(\frac{t}{N}\right)\left(t^2 - \frac{N^2}{3}\right) e^{i(\beta_k^0 - \tilde{\beta}_1)t^2} + \sum_{t=1}^{N} w\left(\frac{t}{N}\right)\epsilon(t)\left(t^2 - \frac{N^2}{3}\right) e^{-i\tilde{\beta}_1 t^2}. \tag{B.2}$$

Using Taylor series and lemma C-2 of [14], the following is written as:

$$\begin{aligned}
&\sum_{t=1}^{N} w\left(\frac{t}{N}\right)\left(t^2 - \frac{N^2}{3}\right) e^{i(\beta_1^0 - \tilde{\beta}_1)t^2} = \\
&\sum_{t=1}^{N} w\left(\frac{t}{N}\right)\left(t^2 - \frac{N^2}{3}\right)\left[e^{i(\beta_1^0 - \beta_1^0)t^2} + e^{i(\beta_1^0 - \beta_1^0)t^2} it^2\left(\beta_1^0 - \tilde{\beta}_1\right) + \frac{\left(it^2\left(\beta_1^0 - \tilde{\beta}_1\right)\right)^2}{2!} e^{i(\beta_1^0 - \beta_1^*)t^2}\right] \\
&= N^3\left(c_3 - \frac{c_1}{3}\right) + i\left(\beta_1^0 - \tilde{\beta}_1\right)\left[N^5\left(c_5 - \frac{c_3}{3}\right) - O_p\left(N^{5-\delta}\right)\right], \text{where, } \beta_1^* \in \left(\beta_1^0, \tilde{\beta}_1\right).
\end{aligned}$$

For $k = 2, \cdots, p$, using lemmas C-2 and C-3 of [14],Taylor series and choosing a large L such that $\frac{3}{2} - L\delta < 0$, the following is written as:

$$\begin{aligned}
&\sum_{t=1}^{N} w\left(\frac{t}{N}\right)\left(t^2 - \frac{N^2}{3}\right) e^{i(\beta_k^0 - \tilde{\beta}_1)t^2} = \sum_{t=1}^{N} w\left(\frac{t}{N}\right) \\
&\left(t^2 - \frac{N^2}{3}\right)\Bigg[e^{i(\beta_k^0 - \beta_1^0)t^2} + \sum_{l=1}^{L-1} \frac{\left(-i\left(\tilde{\beta}_1 - \beta_1^0\right)t^2\right)^l}{l!} e^{i(\beta_k^0 - \beta_1^0)t^2} + \\
&\frac{\left(-i\left(\tilde{\beta}_1 - \beta_1^0\right)t^2\right)^L}{L!} e^{i(\beta_k^0 - \beta_1^*)t^2}\Bigg] = o_p\left(N^{\frac{5}{2}}\right), \text{where, } \beta_1^* \in \left(\beta_1^0, \tilde{\beta}_1\right).
\end{aligned}$$

Now, using Taylor series expansion, Lindeberg Feller CLT, lemma 2 of [20] and choosing a large L such that $1 - L\delta < 0$, we write 2^{nd} term of $P_N^{\tilde{\beta}_1}$, (B.2), as:

$$\sum_{t=1}^{N} \epsilon(t)\, w\left(\frac{t}{N}\right)\left(t^2 - \frac{N^2}{3}\right) e^{-i\tilde{\beta}_1 t^2} = \sum_{t=1}^{N} w\left(\frac{t}{N}\right)\epsilon(t)\left(t^2 - \frac{N^2}{3}\right)$$
$$\left[e^{-i\beta_1^0 t^2} + \sum_{l=1}^{L-1} \frac{\left(-i\left(\tilde{\beta}_1 - \beta_1^0\right)t^2\right)^l}{l!} e^{-i\beta_1^0 t^2} + \frac{\left(-it^2\right)^L}{L!}\left(\tilde{\beta}_1 - \beta_1^0\right)^L e^{-i\beta_1^* t^2}\right]$$
$$= \sum_{t=1}^{N} \epsilon(t)\, w\left(\frac{t}{N}\right)\left(t^2 - \frac{N^2}{3}\right) e^{-i\beta_1^0 t^2} + \sum_{l=1}^{L-1} \frac{(-i)^l}{l!} O_p\left(N^{\frac{5}{2}-l\delta}\right) +$$
$$\frac{(-i)^L}{L!} O_p\left(N^2\right), \text{where}, \beta_1^* \in \left(\beta_1^0, \tilde{\beta}_1\right).$$

Therefore $P_N^{\tilde{\beta}_1}$, (B.2), can be expressed as follows:

$$P_N^{\tilde{\beta}_1} = \sum_{t=1}^{N} \epsilon(t)\, w\left(\frac{t}{N}\right)\left(t^2 - \frac{N^2}{3}\right) e^{-i\beta_1^0 t^2} +$$
$$A_1^0 i\left(\beta_1^0 - \tilde{\beta}_1\right)\left[N^5\left(c_5 - \frac{c_3}{3}\right) - O_p\left(N^{5-\delta}\right)\right] + A_1^0 N^3\left(c_3 - \frac{c_1}{3}\right).$$

Now,

$$\begin{aligned} \breve{\beta}_1 &= \tilde{\beta}_1 + \frac{3c_1}{(3c_5 - c_3)N^4} \operatorname{Im}\left(\frac{P_N^{\tilde{\beta}_1}}{Q_N^{\tilde{\beta}_1}}\right) \\ \breve{\beta}_1 &= \beta_1^0 + \left(\tilde{\beta}_1 - \beta_1^0\right) O_p\left(N^{-\delta}\right) + \frac{3c_1}{(3c_5 - c_3)N^5} \operatorname{Im}\left[\frac{\sum_{t=1}^{N} \epsilon(t)\, w\left(\frac{t}{N}\right)\left(t^2 - \frac{N^2}{3}\right) e^{-i\beta_1^0 t^2}}{A_1^0 c_1}\right]. \end{aligned} \tag{B.3}$$

Using lemma 1 of [20] and lemma C-2 of [14], the asymptotic variance of

$$\frac{3c_1}{(3c_5 - c_3)N^{\frac{5}{2}}} \operatorname{Im}\left[\frac{\sum_{t=1}^{N} \epsilon(t)\, w\left(\frac{t}{N}\right)\left(t^2 - \frac{N^2}{3}\right) e^{-i\beta_1^0 t^2}}{A_1^0 c_1}\right]$$

is $\sigma^2\left(\frac{d_1 + 9d_5 - 6d_3}{2(3c_5 - c_3)^2 |A_1^0|^2}\right)$. Thus using Lindeberg-Feller CLT the theorem holds for $\breve{\beta}_1$. Next, we derive the theorem for $\breve{\beta}_2$. Using Chebyshev inequality and asymptotic distribution for $\breve{\beta}_1$, we get:

$$N^2\left(\breve{\beta}_1 - \beta_1^0\right) \xrightarrow{p} 0 \text{ as } N \to \infty. \tag{B.4}$$

Now, expanding $\cos\left(\breve{\beta}_1 t^2\right)$ and $\sin\left(\breve{\beta}_1 t^2\right)$ using Taylor series, $\breve{A}_{R1}$ is written as follows:

$$\begin{aligned} \breve{A}_{R1} = \frac{1}{Nc_1} \sum_{t=1}^{N} w\left(\frac{t}{N}\right)\Big[y_{R1}(t)\left\{\cos\left(\beta_1^0 t^2\right) - t^2\left(\breve{\beta}_1 - \beta_1^0\right)\sin\left(\bar{\beta}_1 t^2\right)\right\} + \\ y_{I1}(t)\left\{\sin\left(\beta_1^0 t^2\right) + t^2\left(\breve{\beta}_1 - \beta_1^0\right)\cos\left(\bar{\beta}_1 t^2\right)\right\}\Big], \text{where, } \bar{\beta}_1 \in \left(\beta_1^0, \breve{\beta}_1\right). \end{aligned} \tag{B.5}$$

Using (B.4), lemma 2 of [20], lemma C-2 of [14], it can be shown that $\breve{A}_{R1} \xrightarrow{p} A^0_{R1}$ as $N \to \infty$. Similarly, we can prove $\breve{A}_{I1} \xrightarrow{p} A^0_{I1}$ as $N \to \infty$. Thus, we obtain:

$$\breve{A}_{R1} = A^0_{R1} + o_p(1),\ \breve{A}_{I1} = A^0_{I1} + o_p(1) \text{ and } \breve{\beta}_1 = \beta^0_1 + o_p\left(N^{-2}\right). \tag{B.6}$$

Consider $y_2(t) = y_1(t) - \breve{A}_1 e^{i\breve{\beta}_1 t^2}$. Now using (B.6) and then following the similar procedure as used for $\breve{\beta}_1$, we obtain the desired result. Similarly, the result can be obtained for $k = 3, \ldots, p$.

Proof of Theorem 4: Consider $\breve{A}_{R1}$ as given in equation (B.5). Using the Lindeberg-Feller CLT, Theorem 3, (B.4), lemmas 1 and 2 of [20] and lemmas C-2 and C-3 of [14], we obtain the asymptotic distribution of $\breve{A}_{R1}$. Similarly asymptotic distribution of $\breve{A}_{I1}$ can be obtained. Now using (B.6) and $y_2(t) = y_1(t) - \breve{A}_1 e^{i\breve{\beta}_1 t^2}$, the result can be proved for $k = 2$. Similarly, we can derive the result for $k = 3, \ldots, p$.

Proof of Theorem 5 : The result follows from theorems 3 and 4 using the Chebyshev inequality.

Proof of Theorem 6 : Consider $\breve{A}_{R(p+1)}$ and $\breve{A}_{I(p+1)}$; sequential EREs of $A^0_{R(p+1)}$ and $A^0_{I(p+1)}$, respectively:

$$\breve{A}_{R(p+1)} = \frac{1}{Nc_1} \sum_{t=1}^{N} w\left(\frac{t}{N}\right)\left(y_{R(p+1)}(t) \cos\left(\breve{\beta}_{p+1} t^2\right) + y_{I(p+1)}(t) \sin\left(\breve{\beta}_{p+1} t^2\right)\right),$$

$$\breve{A}_{I(p+1)} = \frac{1}{Nc_1} \sum_{t=1}^{N} w\left(\frac{t}{N}\right)\left(y_{I(p+1)}(t) \cos\left(\breve{\beta}_{p+1} t^2\right) - y_{R(p+1)}(t) \sin\left(\breve{\beta}_{p+1} t^2\right)\right), \text{ where,}$$

$$\begin{aligned} y_{R(p+1)}(t) &= y_{R1}(t) - \sum_{k=1}^{p}\left(\breve{A}_{Rk} \cos\left(\breve{\beta}_k t^2\right) - \breve{A}_{Ik} \sin\left(\breve{\beta}_k t^2\right)\right), \\ y_{I(p+1)}(t) &= y_{I1}(t) - \sum_{k=1}^{p}\left(\breve{A}_{Rk} \sin\left(\breve{\beta}_k t^2\right) + \breve{A}_{Ik} \cos\left(\breve{\beta}_k t^2\right)\right). \end{aligned} \tag{B.7}$$

Using theorems 3 and 4, (B.7) can be written as follows:

$$\begin{aligned} y_{R(p+1)}(t) &= \epsilon_R(t) + o_p(1), \\ y_{I(p+1)}(t) &= \epsilon_I(t) + o_p(1). \end{aligned} \tag{B.8}$$

Using (B.8) and lemma 2 of [20], we get: $\breve{A}_{R(p+1)} \xrightarrow{p} 0$, $\breve{A}_{I(p+1)} \xrightarrow{p} 0$ as $N \to \infty$. Thus, the result holds.

References

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

[2] K. Abratkiewicz, K. Czarnecki, D. Fourer, and F. Auger. "Estimation of time-frequency complex phase-based speech attributes using narrow band filter banks," *IEEE Signal Processing Symposium*, 1–6, 2017.

[3] K. Abratkiewicz, D Gromek, and P. Samczynski. "Chirp rate estimation and micro-Doppler signatures for pedestrian security radar systems," *IEEE Signal Processing Symposium*, 212–215, 2017.

[4] K. Abratkiewicz, and P. Samczynski. " A block method using the chirp rate estimation for NLFM radar pulse reconstruction," *Sensors*, vol 19, no. 22, 2019.
[5] 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.
[6] 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.
[7] F.A.C. Alegria. "Estimating sinusoidal amplitude from phase noise and jitter corrupted data points," *Journal of the Franklin Institute*, 108747, 2026.
[8] S. S. Banerjee, A. Mitra, and R. Mondal. "Asymptotic Analysis of Regression Quantile Estimators for Real-Valued Chirp Signal Model," *Circuits, Systems, and Signal Processing*, vol 43, no. 2, 1053–1100, 2024.
[9] R. Demirli, and J. Saniie. "Asymmetric Gaussian chirplet model and parameter estimation for generalized echo representation," *Journal of the Franklin Institute*, vol 351, no. 2, 907–921, 2014.
[10] I. Djurović, and V. N. Ivanović. "The second-order local polynomial Fourier transform as instantaneous frequency and chirp rate estimator," *Signal Processing*, vol 203, 108794, 2023.
[11] 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.
[12] R. Grover, D. Kundu, and A. Mitra. "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, 5421–5465, 2021.
[13] D. Kundu, S. Nandi, and R. Grover. "On Weighted Least Squares Estimators of Parameters of a Chirp Model," *Circuits, Systems, and Signal Processing*, vol 42, no. 1, 493–521, 2023.
[14] D. Kundu, S. Nandi, and R. Grover. "On Weighted Least Squares Estimators for Chirp Like Model," *Sankhya A*, vol 86, no. 1, 27–66, 2024.
[15] X. Liu, J. Han, C. Wang, and B. Xiao. "Parameters estimation for chirp signal based on QPF-FRFT," *Optik*, vol 182, 529–537, 2019.
[16] S. Mandal, B. Santhanam, and M. M. Hayat. ""On improved accuracy chirp parameter estimation using the dfrft with application to sar-based vibrometry," 52^{nd} *IEEE Asilomar Conference on Signals, Systems, and Computers*, 1098–1102, 2018.
[17] L. T. Minh. "New Wigner distribution and ambiguity function for the linear canonical transform: Convolution theorems and applications in the detection of linear frequency modulated signals," *Journal of the Franklin Institute*, 107970, 2025.
[18] 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.
[19] 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.
[20] A. Mittal, D. Kundu, and A. Mitra. "On Weighted Least Squares Estimation of Elementary Chirp Model," *IEEE Transactions on Aerospace and Electronic Systems*, 2025.
[21] R. Neagu, and I. Zurbenko. "Tracking and separating non-stationary multi-component chirp signals," *Journal of the Franklin Institute*, vol 339, no. 6-7, 499–520, 2002.
[22] A. H. Oveis, M. A. Sebt, A. Noroozi, and R. Saleh. "Efficient Chirp Parameters Estimation Based on the Ringing Effect With Application to the Velocity Estimation of Ground Moving Targets," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol 11, no. 12, 4826–4834, 2018.
[23] 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.
[24] J. W. Pitton. "The statistics of time–frequency analysis," *Journal of the Franklin Institute*, vol 337, no. 4, 379–388, 2000.
[25] 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.
[26] N. Rabiee, H. Azad, and N. Parhizgar. "Enhancement DPT method in terms of estimation chirp rate and central frequency parameters of the LFM signal," *Radioengineering*, vol 28, no. 1, 231, 2019.
[27] 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.
[28] 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.
[29] 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.

[30] P. Wang, and J. Yang. “Multicomponent chirp signals analysis using product cubic phase function,” *Digital Signal Processing*, vol 16, no. 6, 654–669, 2006.
[31] G. Zhai, J. Zhou, K. Yu, and J. Li. “Chirp Rate Estimation of LFM Signals Based on Second-Order Synchrosqueezing Transform,” *Electronics*, vol 12, no. 24, 4938, 2023.
[32] H. Begleiter. https://archive.ics.uci.edu/ml/datasets/EEG%20Database.