ESTIMATING PARAMETERS OF ELEMENTARY CHIRP MODEL USING MODIFIED NEWTON-RAPHSON ALGORITHM

SWAGATA NANDI[1] AND DEBASIS KUNDU[2]

ABSTRACT. In this manuscript we have proposed a modified Newton-Raphson algorithm to obtain efficient estimators of the elementary chirp signal parameters in presence of stationary noise. The asymptotic distribution of the least squares estimators which minimizes the residual sum of squares has been provided. The least squares optimization function is a highly non-linear function in its chirp parameters and an iterative procedure is needed to optimize the criterion function. The Newton-Raphson algorithm does not work well for this kind of models. We have used Newton-Raphson algorithm by a step factor modification. The proposed algorithm starts with an initial estimator of order $O_p(n^{-2})$. This modified Newton-Raphson method provides estimators with the same rate of convergence $O_p(n^{-5/2})$ as the least squares estimators. Further, the asymptotic variances of the proposed estimators are $\frac{4}{9}$ times smaller than those of the least squares estimators. Extensive numerical experiments are conducted to study the performances of the proposed estimators and compared with the least squares estimators and fixed iteration efficient estimators using mean squared errors. The analyses of two sonar data, one sonar rocks data and one sonar mines data are demonstrated using the elementary chirp model and estimating the parameters using the proposed modified Newton-Raphson algorithm.

1. Introduction

In this manuscript, we have considered the problem of estimating the elementary chirp model parameters in presence of stationary noise. Mathematically, the model can be expressed as follows;

$$y(t) = \sum_{j=1}^{p} \Big[A_j^0 \cos(\beta_j^0 t^2) + B_j^0 \sin(\beta_j^0 t^2) \Big] + X(t); \quad t = 1, \ldots, n. \tag{1}$$

Here $y(t)$ is the observed signal and it has p components, the j-th component corresponds to the chirp rate β_j^0. The linear parameters A_j^0, B_j^0, $j = 1, \ldots, p$ are unknown amplitude

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

Key words and phrases. Elementary chirp signal, Least squares estimators, Asymptotic distributions, Efficient estimators, Modified Newton-Raphson, Asymptotic normality.

parameters associated with the chirp rate or the frequency rate parameter β_j^0. The random variable $X(t)$ is a mean zero stationary process and its structure is explicitly stated in Assumption 1. In this paper, we address the problem of estimating the unknown parameters $\{(A_j, B_j, \beta_j); j = 1, \ldots, p\}$; given a sample of size n.

The model as given in equation (1) is a special case of the usual chirp signal model, see for example Lahiri et al. [14], Santhanam and Santhanam [26], Shukla et al. [28], Miao [17], Ding et al. [8] and the references cited there in. Mboup and Adali [16] demonstrated the sum of elementary chirp model where a generalization of Fourier transform was presented to study the spectral analysis of chirp-like signals. Estimation of chirp parameters, that is, the frequency rate parameters in multicomponent elementary chirp model is an important problem in statistical signal processing and has received considerable attention in recent times. Some of the examples where elementary chirp signal have been used include micro-Doppler signature analysis in radar systems [3], radar pulse reconstruction [4], systematic aperture radar images [11], mode reconstruction [15], sonar pulse detection [1] and in many other real life situations. One can see Mboup and Adali [16], Mittal et al. [18], [19], Casazza and Fickus [7] and references cited therein.

Estimation of the unknown chirp rates parameters is an important issue. It has received a significant amount of attention in the signal processing literature. Some of the available techniques mainly aim to estimate the instantaneous frequency rate (IFR) which basically estimates the twice the chirp rate. The method based on cubic phase function (CPF) and its variants, mostly motivated by the CPF, are available in the literature. These methods include integrated CPF [31], product CPF [32] and non-parametric chirp rate estimator based on CPF [10]. There are several other techniques used in the literature to estimate the chirp rate. For example, methods based on discrete polynomial phase transform [23] and weighted combination [5], fractional auto-correlation and fractional Fourier transform [20], modified discrete chirp Fourier transform [29], time-frequency rate distribution [30] etc. have been widely used in practice.

The least squares estimators (LSEs) provide optimal rate of convergence in sinusoidal and usual chirp type models. Mittal et al. [19] derived asymptotic properties of the LSEs and the approximate LSEs in an equivalent complex-valued model of the model as given in equation (1). In presence of Gaussian error the asymptotic variances of the LSEs reach the Cramer-Rao lower bound. Shukla et al. [27] studied the asymptotic properties of the LSEs and sequential combined estimator for a multicomponent chirp model with equal rates. Although, the LSEs have the optimal rate of convergence, finding them is a computationally challenging problem. The model as given in equation (1) is a non-linear model in chirp rate parameters and the criterion function needs to be minimized, is a highly non-linear function in its parameters. Thus, very good (close enough to the true value) initial estimators are needed for any iterative procedure to converge. Mittal et al. [18] proposed two efficient algorithms based on choosing two different initial estimators of the chirp rate parameters. Their estimators have the same rates of convergence as the LSEs.

In this paper, we first discuss the method of least squares (LS) which minimizes the residual sum of squares and obtain the asymptotic distribution. As the main aim is to estimate the chirp rate parameters in an efficient way, we propose an iterative algorithm by modifying the well known Newton-Raphson (NR) method. The elementary chirp model, considered in this paper, is a more complicated model than the sinusoidal model. It has already been observed by several authors, see for example Kundu [13], that the standard Newton-Raphson algorithm does not work for the sinusoidal model. Kundu et al [13] proposed an estimation technique by modifying the NR algorithm for the sum of sinusoidal model. A similar approach has been used by Nandi and Kundu [21] for a fundamental frequency model with harmonics. The elementary chirp model, considered in this paper, is a more complicated model than the sinusoidal model. So it is expected that it would not work well in practice.

Our aim is to find the LSEs of the unknown parameters involved in the model as given in equation (1). We observe that the model as given in equation (1) is a highly non-linear model and the LS surface may have several local minima. Finding LSEs needs numerical algorithms to be implemented. The usual Gauss-Newton or NR algorithm does not work well for this kind of models. In many cases, they do not converge or converge to local minima. Therefore, the aim is to use the NR algorithm with a modification by reducing the step factor in each iteration. The idea is with smaller step in each iteration so that the algorithm will not miss the global minimum.

It should be mentioned that there are some deep learning and related methods available today to estimate sinusoidal frequency parameters, see for example Pan et al. [22]. It seems with some efforts these methods can be modified to estimate the chirp rate parameters of the model (1) also. It has not been attempted here because even if these methods can be modified for the model (1), it will not be easy to establish any theoretical properties of these estimators. Moreover, it is well known that any deep learning related method is computationally quite challenging, see [22]. The aim of this paper is different. The main contributions of this paper are the following; (a) provide a simple numerical method which can used quite conveniently in practice, (b) establish theoretically that these estimators have the same convergence rate as the most efficient least squares estimators and (c) compare numerically with the most efficient least squares estimators. Recently it has been observed by Mittal et al. [18, 19] that for multiple elementary chirp model, the sequnetial procedure works very well. Hence, here also we have adopted sequential process for multiple elementary chirp model.

The rest of the paper is organized as follows. In Section 2, we state the assumptions required. In Section 3, we describe the method of LS for model as given in equation (1) and state its asymptotic distribution. In Section 4, we propose the modified NR (MNR) method for a single component model and its implementation. In Section 5, we extend the algorithm to the multi-component case. Numerical results are presented in Section 6. One sonar rocks and another sonar mines data are analyzed in Section 7 and the paper in concluded in Section 8. All the proofs are provided in the Appendices.

2. Assumptions

In this section, we state the assumptions required for the LSEs as well as for the proposed method for the establishment of theoretical results.

Assumption 1. *The random variable $\{X(t)\}$ has the following linear structure*

$$X(t) = \sum_{k=-\infty}^{\infty} a(k)e(t-k), \tag{2}$$

where $\{e(t)\}$ is a sequence of independent and identically distributed (i.i.d.) random variables with mean zero and finite variance σ^2. The arbitrary real valued sequence $\{a(t)\}$ satisfies the following condition of absolute summability;

$$\sum_{k=-\infty}^{\infty} |a(k)| < \infty. \tag{3}$$

Assumption 2. *For $k = 1, \ldots, p$, A_k^0 and B_k^0 are not simultaneously equal to zero.*

Assumption 3. *$\beta_k^0 \in (0, \pi)$ and $\beta_k^0 \neq \beta_j^0$ for $k, j = 1, \ldots, p$.*

Assumption 1 is a general assumption of a weakly stationary process. Any stationary autoregressive (AR), and autoregressive moving average (ARMA) process can be expressed as equation(2), when the coefficients are absolutely summable. Assumptions 2 and 3 ensure that the model as given in equation (1) is a p-component model. Note that these are some basic assumptions one is needed for model identifiability, see for example Mittal et al. [18, 19].

3. Least Squares Estimators

In this section, we define the LSEs of the unknown parameters of model as given in equation (1). We denote the parameter vector corresponding to the k-th component as $\boldsymbol{\theta}_k = (A_k, B_k, \beta_k)^\top$ and $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^\top, \ldots, \boldsymbol{\theta}_p^\top)^\top$ is the parameter vector of order $3p$. Then the LSE of $\boldsymbol{\theta}$ minimizes

$$Q(\boldsymbol{\theta}) = \sum_{t=1}^{n} \left[y(t) - \sum_{j=1}^{p} \Big(A_j \cos(\beta_j t^2) + B_j \sin(\beta_j t^2) \Big) \right]^2 \tag{4}$$

with respect to A_j, B_j and β_j, $j = 1, \ldots, p$. Write $\mathbf{Y} = (y(1), \ldots, y(n))^T$ and

$$\mathbf{Z}(\boldsymbol{\beta}) = \Big[\mathbf{Z}_1(\beta_1), \ldots, \mathbf{Z}_p(\beta_p) \Big],$$

$$\mathbf{Z}_k(\beta_k) = \begin{bmatrix} \cos(\beta_k) & \sin(\beta_k) \\ \cos(2^2\beta_k) & \sin(2^2\beta_k) \\ \vdots & \vdots \\ \cos(n^2\beta_k) & \sin(n^2\beta_k) \end{bmatrix} \tag{5}$$

and $\boldsymbol{\xi} = (\boldsymbol{\xi}_1^\top, \ldots, \boldsymbol{\xi}_p^\top)^\top$, $\boldsymbol{\xi}_k = (A_k, B_k)^\top$, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^\top$, $k = 1, \ldots, p$. Then $Q(\boldsymbol{\theta})$ can be written as

$$\begin{aligned} Q(\boldsymbol{\theta}) &= (\mathbf{Y} - \mathbf{Z}_1(\beta_1)\boldsymbol{\xi}_1 - \ldots - \mathbf{Z}_p(\beta_p)\boldsymbol{\xi}_p)^\top (\mathbf{Y} - \mathbf{Z}_1(\beta_1)\boldsymbol{\xi}_1 - \ldots - \mathbf{Z}_p(\beta_p)\boldsymbol{\xi}_p) \\ &= (\mathbf{Y} - \mathbf{Z}(\boldsymbol{\beta})\boldsymbol{\xi})^\top (\mathbf{Y} - \mathbf{Z}(\boldsymbol{\beta})\boldsymbol{\xi}). \end{aligned} \tag{6}$$

For a given $\boldsymbol{\beta}$, the LSEs of $\boldsymbol{\xi}$ is $\widehat{\boldsymbol{\xi}}(\boldsymbol{\beta}) = \big(\mathbf{Z}(\boldsymbol{\beta})^\top \mathbf{Z}(\boldsymbol{\beta})\big)^{-1} \mathbf{Z}(\boldsymbol{\beta})^\top \mathbf{Y}$. Replacing this in $Q(\boldsymbol{\theta})$, we have

$$Q(\widehat{\boldsymbol{\xi}}(\boldsymbol{\beta}), \boldsymbol{\beta}) = \mathbf{Y}^\top (\mathbf{I} - \mathbf{P}_{\mathbf{Z}(\boldsymbol{\beta})}) \mathbf{Y},$$

$$\mathbf{P}_{\mathbf{Z}(\boldsymbol{\beta})} = \mathbf{Z}(\boldsymbol{\beta}) \big[\mathbf{Z}(\boldsymbol{\beta})^\top \mathbf{Z}(\boldsymbol{\beta}) \big]^{-1} \mathbf{Z}(\boldsymbol{\beta})^\top,$$

where $\mathbf{P}_{\mathbf{Z}(\boldsymbol{\beta})}$ is the projection matrix on the column space of $\mathbf{Z}(\boldsymbol{\beta})$. Then minimizing $Q(\widehat{\boldsymbol{\xi}}(\boldsymbol{\beta}), \boldsymbol{\beta})$ with respect to $\boldsymbol{\beta}$ boils down to maximizing $R(\boldsymbol{\beta})$ with respect to $\boldsymbol{\beta}$, where

$$R(\boldsymbol{\beta}) = \mathbf{Y}^T \mathbf{Z}(\boldsymbol{\beta}) \big[\mathbf{Z}(\boldsymbol{\beta})^\top \mathbf{Z}(\boldsymbol{\beta}) \big]^{-1} \mathbf{Z}(\boldsymbol{\beta})^\top \mathbf{Y}. \tag{7}$$

Using Lemma 1, stated in Appendix A, the following simplification can be done for large n.

$$\widehat{\boldsymbol{\xi}}_k(\beta_k) = \big(\mathbf{Z}_k(\beta_k)^\top \mathbf{Z}_k(\beta_k) \big)^{-1} \mathbf{Z}_k(\beta_k)^\top \mathbf{Y}, \tag{8}$$

$$R(\boldsymbol{\beta}) = \sum_{k=1}^{p} \mathbf{Y}^T \mathbf{Z}_k(\beta_k) \big[\mathbf{Z}_k(\beta_k)^\top \mathbf{Z}_k(\beta_k) \big]^{-1} \mathbf{Z}_k(\beta_k)^\top \mathbf{Y}. \tag{9}$$

Let $\widehat{\boldsymbol{\beta}} = (\widehat{\beta}_1, \ldots, \widehat{\beta}_p)^\top$ minimizes $R(\boldsymbol{\beta})$ and then the LSEs of the linear parameters are obtained as $\widehat{\boldsymbol{\xi}}_k(\widehat{\boldsymbol{\beta}}_k)$, $k = 1, \ldots, p$. This is the well known separable regression technique of Richards [25].

In the following, we state the asymptotic distribution of the LSE of $\boldsymbol{\theta}$. Let $\widehat{\boldsymbol{\theta}}_k$ and $\boldsymbol{\theta}_k^0$ be the LSE and the true value of $\boldsymbol{\theta}_k$, respectively. Define a diagonal matrix $\mathbf{D}$ as $\mathbf{D} = \text{diag}\{n^{1/2}, n^{1/2}, n^{5/2}\}$. Then the asymptotic distribution of $\widehat{\boldsymbol{\theta}}_k$ is given in the Theorem 3.1.

Theorem 3.1. *Under Assumptions 1-2, as $n \to \infty$*

$$\mathbf{D}(\widehat{\boldsymbol{\theta}}_k - \boldsymbol{\theta}_k^0) \xrightarrow{d} \mathcal{N}_3(\mathbf{0}, 2\sigma^2 c_k \boldsymbol{\Sigma}_k^{-1})$$

where c_k, $\boldsymbol{\Sigma}_k$ and $\boldsymbol{\Sigma}_k^{-1}$, $k = 1, \ldots, p$ are as follows;

$$c_k = \Big[\sum_{j=-\infty}^{\infty} a(j)\cos(\beta_k^0 j^2)\Big]^2 + \Big[\sum_{j=-\infty}^{\infty} a(j)\sin(\beta_k^0 j^2)\Big]^2 = \Big|\sum_{j=-\infty}^{\infty} a(j) e^{-\beta_k^0 j^2}\Big|^2,$$

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1 & 0 & \frac{B_k^0}{3} \\ 0 & 1 & -\frac{A_k^0}{3} \\ \frac{B_k^0}{3} & -\frac{A_k^0}{3} & \frac{1}{5}({A_k^0}^2 + {B_k^0}^2) \end{pmatrix},$$

$$\boldsymbol{\Sigma}^{-1} = \frac{1}{4({A_k^0}^2 + {B_k^0}^2)} \times \begin{pmatrix} (4{A_k^0}^2 + 9{B_k^0}^2) & -5A_k^0 B_k^0 & -15B_k^0 \\ -5A_k^0 B_k^0 & (9{A_k^0}^2 + 4{B_k^0}^2) & 15A_k^0 \\ -15B_k^0 & 15A_k^0 & 45 \end{pmatrix}.$$

For $k \neq m$, $\mathbf{D}(\widehat{\boldsymbol{\theta}}_k - \boldsymbol{\theta}_k^0)$ and $\mathbf{D}(\widehat{\boldsymbol{\theta}}_m - \boldsymbol{\theta}_m^0)$ are asymptotically independently distributed. ∎

The LSEs provide the optimal rates of convergence for the amplitudes and the chirp parameter as $O_p\Big(n^{-1/2}\Big)$ and $O_p\Big(n^{-5/2}\Big)$, respectively. Here, $x_n = O_p\Big(n^{-\delta}\Big)$ means that $n^\delta x_n$ is bounded in probability.

Remark 1. The asymptotic distribution of the chirp rate parameter corresponding to the k-th component is given by

$$n^{5/2}(\widehat{\beta}_k - \beta_k^0) \xrightarrow{d} \mathcal{N}\Big(0, \frac{45\sigma^2 c_k}{2({A_k^0}^2 + {B_k^0}^2)}\Big).$$

The rate of convergence $O_p(n^{-5/2})$ is same as the LSE of chirp rate parameter in the usual chirp model, see reference [14].

4. Modified Newton-Raphson Algorithm

The NR algorithm is a widely used numerical method in non-linear optimization. The aim is to find an efficient estimator using an NR type algorithm. The elementary chirp model is a highly non-linear model and belongs to the class of sum of sinusoidal frequency type models. It is well known that the NR algorithm does not work well in finding the LSEs of sinusoidal frequency model. Due to this reason the NR algorithm has not been used in the literature to find LSEs of sinusoidal models. So it is not advisable to use NR algorithm in its original form to estimate the chirp rate parameters in elementary multiple chirp models. Our aim is to modify this simple algorithm which will perform better in maximizing $R(\boldsymbol{\beta})$, defined in (7).

In this section, we concentrate on a single component model given by

$$y(t) = A^0 \cos(\beta^0 t^2) + B^0 \sin(\beta^0 t^2) + X(t); \qquad t = 1, \ldots, n. \tag{10}$$

The corresponding function to be maximized to find the LSE of β is given by

$$R(\beta) = \mathbf{Y}^T \mathbf{Z}(\beta)\big[\mathbf{Z}(\beta)^\top \mathbf{Z}(\beta)\big]^{-1} \mathbf{Z}(\beta)^\top \mathbf{Y}, \tag{11}$$

where $\mathbf{Z}(\beta)$ is a $n \times 2$ matrix equal to $\mathbf{Z}_1(\beta_1)$ with β_1 replaced by β. Now we first briefly describe the NR algorithm in case of maximizing $R(\beta)$. Let $\widehat{\beta}_1$ be the initial estimate of β^0 and $\widehat{\beta}_m$ be the estimate at the m-th iteration. Then the estimate at the $m+1$ iteration, $\widehat{\beta}_{m+1}$ is obtained as

$$\widehat{\beta}_{m+1} = \widehat{\beta}_m - \frac{R'(\widehat{\beta}_m)}{R''(\widehat{\beta}_m)}, \tag{12}$$

where $R'(\widehat{\beta}_m)$ and $R''(\widehat{\beta}_m)$ are the first and second order derivatives of $R(\beta)$, respectively, evaluated at $\widehat{\beta}_m$.

This standard NR method is modified with a reduction in the step factor in the following way.

$$\widehat{\beta}_{m+1} = \widehat{\beta}_m - \frac{4}{9} \times \frac{R'(\widehat{\beta}_m)}{R''(\widehat{\beta}_m)}, \tag{13}$$

A smaller correction factor in each iteration forces the estimator to converge to the global maximum. In the process, it actually prevents the algorithm to diverge or converge to a local maxima of $R(\beta)$. At a particular iteration, if the estimate is close enough to the global maximum, a large correction factor may move the estimate far away from the global maximum. Theorem 4.1, stated below, leads us to modify the NR algorithm in this fashion.

Theorem 4.1. *Let $\widetilde{\beta}$ be a consistent estimator of β^0 and $\widetilde{\beta} - \beta^0 = O_p(n^{-2-\delta})$, $\delta \in \left(0, \frac{1}{2}\right]$. Suppose $\tilde{\beta}$ is updated as $\widehat{\beta} = \widetilde{\beta} - \frac{4}{9} \times \frac{R'(\widetilde{\beta})}{R''(\widetilde{\beta})}$, then*

(a) $\widehat{\beta} - \beta^0 = O_p(n^{-2-3\delta})$ *if* $\delta \leq \frac{1}{6}$

(b) $n^{5/2}(\widehat{\beta} - \beta^0) \xrightarrow{d} N\left(0, \frac{10\sigma^2 c}{A^{02} + B^{02}}\right)$ *if* $\delta > \frac{1}{6}$,

where $c = \left| \sum_{j=-\infty}^{\infty} a(j) e^{-i\beta^0 j^2} \right|^2$.

Proof of Theorem 4.1: See in Appendix B.

According to theorem 4.1, if we start with a consistent estimator with a reasonable rate of convergence, then the MNR algorithm estimates the frequency rate parameter with the same rate of convergence as the LSE. In addition, the asymptotic variance of the LSE of β is more than twice than the asymptotic distribution of the MNR estimate. Specifically, the asymptotic variance of the LSE of β is $\frac{9}{4}$ times the asymptotic distribution of the MNR estimate.

Now we explain that starting from initial estimator $\widetilde{\beta}$ of β^0 with rate of convergence $O_p(n^{-2})$ how the above theorem is used to obtain an efficient estimator. Consider $\widetilde{\beta}$, the maximizer of $R(\beta)$ over a grid $\left\{\frac{\pi k}{n^2}, k = 1, \ldots, n^2 - 1\right\}$ as the initial estimator, then following Rice and Rosenblatt [24] the existence of such an initial guess can be shown. One can also show that $n^2(\widetilde{\beta} - \beta^0) \xrightarrow{a.s.} 0$.

Implementation: The key idea of the proposed algorithm is to use Theorem 4.1 judiciously. The estimator $\widetilde{\beta}$ is calculated using a subset of the observed data vector of size n and not to use the whole sample at the beginning so that the required order can be achieved. This approach was first applied by Bai et al. [6]. As the observations are dependent, the subset is selected in such a way that the dependence structure is not destroyed. Therefore, a subset of predefined size of consecutive data points is chosen to start the algorithm and we gradually increase the sample size to the whole sample. To initiate the algorithm, we obtain the argument maximum of $R(\beta)$ over $\left\{\frac{\pi k}{n^2}, k = 1, \ldots, n^2 - 1\right\}$ and denote it $\widetilde{\beta}_{(0)}$. Then $\widetilde{\beta}_{(0)} = O_p(n^{-2})$. The following algorithm is used to implement Theorem 4.1. Here $[x]$ denotes the integer part of x. We use the following notation just for the Algorithm. We denote $R'(\beta)$ and $R''(\beta)$ as $R'_m(\beta)$ and $R''_m(\beta)$, respectively. Here, $R'_m(\beta)$ and $R''_m(\beta)$ have been calculated based on subsample $\{y(1),\ldots,y(m)\}$ of $\{y(1),\ldots,y(n)$, for $m \leq n$.

Algorithm:

1. Choose $n_1 = [n^{14/15}]$, then $n = n_1^{15/14}$. Calculate $\widetilde{\beta}_{(1)}$ as

$$\widetilde{\beta}_{(1)} = \widetilde{\beta}_{(0)} - \frac{4}{9} \times \frac{R'_{n_1}(\widetilde{\beta}_{(0)})}{R''_{n_1}(\widetilde{\beta}_{(0)})}.$$

   Note that $\widetilde{\beta}_{(0)} - \beta^0 = O_p(n^{-2}) = O_p(n_1^{-2-1/7})$, therefore $\delta = \frac{1}{7} < \frac{1}{6}$ and after the first iteration using part (a) of Theorem 4.1, we have

$$\widetilde{\beta}_{(1)} - \beta^0 = O_p(n_1^{-2-3/7}) = O_p(n^{-34/15}) = O_p(n^{-2-4/15}).$$

2. As $\widetilde{\beta}_{(1)} - \beta^0 = O_p(n^{-2-4/15})$, $\delta = \frac{4}{15} > \frac{1}{6}$, we can use part (b) of Theorem 4.1. Take $n_{m+1} = n$ and repeat

$$\widetilde{\beta}_{(m+1)} = \widetilde{\beta}_{(m)} - \frac{4}{9} \times \frac{R'_{n_{m+1}}(\widetilde{\beta}_{(m)})}{R''_{n_{m+1}}(\widetilde{\beta}_{(m)})},$$

   for $m = 1, 2, \ldots$ until a suitable stopping criterion is satisfied. ∎

The algorithm suggests to start the iteration process with a sub-sample of size $n^{14/15}$ of the original sample of size n. Any sub-sample will work provided the correlation/dependence structure is not destroyed. The choice of the initial sub-sample is not important and size of the sub-sample is not unique. We can start with any n_1 consecutive data points to keep the dependence structure intact.

Remark 2. Kundu et al. [13] proposed an MNR algorithm for the sinusoidal frequency model by using a smaller step factor in each iteration of the algorithm. In sinusoidal model the MNR step factor is required to reduce $\frac{1}{4}$ times the NR step factor to obtain an efficient estimator whereas in case of the elementary chirp model, the correction factor is $\frac{4}{9}$.

Remark 3. As the linear parameters are separable and LSEs or MNR estimators of A^0 and B^0 are function of the corresponding estimator of β^0, we write

$$(\widehat{A}, \widehat{B}) = (A(\widehat{\beta}), B(\widehat{\beta})), \quad (\widehat{A}_{LS}, \widehat{B}_{LS}) = (A(\widehat{\beta}_{LS}), B(\widehat{\beta}_{LS})).$$

Here $\widehat{A}, \widehat{B}$ and $\widehat{\beta}$ denote the MNR estimate where as $\widehat{A}_{LS}, \widehat{B}_{LS}$ and $\widehat{\beta}_{LS}$ denote the LSEs of A^0, B^0 and β^0, respectively.

Expanding $A(\widehat{\beta})$ using Taylor expansion around the true value β_0, we can approximate

$$A(\widehat{\beta}) - A(\beta^0) = (\widehat{\beta} - \beta^0)[A'(\bar{\beta})],$$

where $A'(\bar{\beta})$ is the first order derivative of $A(\beta)$ at $\bar{\beta}$; $\bar{\beta}$ is a point on the line joining $\widehat{\beta}$ and β^0. As $\widehat{\beta}$ is asymptotically unbiased, for large n

$$\begin{aligned} & \mathrm{Var}[A(\widehat{\beta}) - A(\beta^0)] = \mathrm{Var}[\widehat{\beta} - \beta^0][A'(\bar{\beta})]^2 \\ \Rightarrow \quad & \mathrm{Var}[A(\widehat{\beta}) - A^0] = \frac{4}{9} \times \mathrm{Var}[\widehat{\beta}_{LS}][A'(\bar{\beta})]^2 = \frac{4}{9} \times \mathrm{Var}[\widehat{A}_{LS}]. \end{aligned}$$

Similarly, we can show that $\mathrm{Var}[B(\widehat{\beta}) - B^0] = \frac{4}{9} \times \mathrm{Var}[\widehat{B}_{LS}]$. Therefore, the asymptotic variances of $\widehat{A}$ and $\widehat{B}$ are also $\frac{4}{9}$ times those of the LSEs of A and B,respectively. It can be shown that $\mathrm{Cov}(\widehat{A}, \widehat{B}) = \frac{4}{9} \times \mathrm{Cov}(\widehat{A}_{LS}, \widehat{B}_{LS})$. Therefore, the asymptotic variance-covariance matrix of MNR estimators is $\frac{4}{9}$ times that of the LSEs.

5. Multi-component elementary model

In this section, we consider the multi-component elementary chirp model. Based on the idea of the MNR algorithm proposed in Section 4 for a single component model, we propose a sequential MNR algorithm for parameters of model as given in equation (1). Under Assumption 3, the following restriction is imposed on the amplitudes for identifiability issues.

$$A_1^{0\,2} + B_1^{0\,2} \geq \cdots \geq A_p^{0\,2} + B_p^{0\,2}. \tag{14}$$

Also, $1 \leq k \leq p-1$, $A_k^{0\,2} + B_k^{0\,2} = {A_{k+1}^0}^2 + {B_{k+1}^0}^2$, then it is assumed $\omega_k < \omega_{k+1}$. The sequence $\{X(t)\}$ satisfies Assumption 1.

5.1. Sequential Modified Newton-Raphson Algorithm.

We estimate the unknown parameters of model as given in equation (1) in a sequential manner component-wise. We estimate the first component using exactly the same method as the one implemented for one component model, discussed in Section 4. We start the algorithm from the maximizer of $R(\beta)$ and estimate β_1 as $\widehat{\beta}_1$. The corresponding linear parameters are estimated as $\widehat{A}_1$ and $\widehat{B}_1$ using $\widehat{\boldsymbol{\xi}}_1(\widehat{\beta}_1)$ where $\widehat{\boldsymbol{\xi}}_k(\beta_k)$ is defined in equation (8) for $k = 1, \ldots, p$. Then remove the effect of the estimated chirp rate $\widehat{\beta}_1$ from the observed data and define

$$y_1(t) = y(t) - \widehat{A}_1 \cos(\widehat{\beta}_1 t^2) - \widehat{B}_1 \sin(\widehat{\beta}_1 t^2).$$

Next, we use $[y_1(1), \ldots, y_n(n)]$ to estimate the second chirp component parameter, that is, A_2, B_2 and β_2. We continue to do so until all the p components parameters are estimated. We have assumed that p is known. But this sequential procedure automatically provides us an estimate of p. In case of real data, if this procedure is continued for q times and the residual series after q repetitions is a completely random sequence, then q can be taken as an estimate of p. This approach is implemented in Section 7 in Data Analysis.

The estimator $\widehat{\beta}_k$ of β_k^0 at the k-th step of the above sequential algorithm is the MNR estimate and similar to one component model is asymptotically normally distributed, stated in Theorem 5.1.

Theorem 5.1. *Let $\widetilde{\beta}_k$ be a consistent estimator of β_k^0, the chirp rate corresponding to the k-th component and $\widetilde{\beta}_k - \beta_k^0 = O_p(n^{-2-\delta})$, for some $\delta \in \left(0, \frac{1}{2}\right]$. Suppose $\widetilde{\beta}_k$ is updated as $\widehat{\beta}_k = \widetilde{\beta}_k - \frac{4}{9} \times \frac{R'(\widetilde{\beta}_k)}{R''(\widetilde{\beta}_k)}$, then*

$$\text{(a)}\ \widehat{\beta}_k - \beta_k^0 = O_p(n^{-2-3\delta}) \qquad \text{if } \delta \le \frac{1}{6}$$

$$\text{(b)}\ n^{5/2}(\widehat{\beta}_k - \beta_k^0) \xrightarrow{d} N\left(0, \frac{10\sigma^2 c_k}{{A_k^0}^2 + {B_k^0}^2}\right) \qquad \text{if } \delta > \frac{1}{6},$$

where $c_k = \left|\sum_{j=-\infty}^{\infty} a(j) e^{-i\beta_k^0 j^2}\right|^2$. Moreover, $n^{5/2}(\widehat{\beta}_k - \beta_k^0)$ and $n^{5/2}(\widehat{\beta}_l - \beta_l^0)$ for $k \neq l$ are asymptotically independently distributed.

6. Numerical Experiments

In this section, we present results from numerical experiments conducted using elementary chirp models for illustrative purposes. We have considered two models, a single component model and a two component model to assess the finite-sample performance of the proposed estimators with the following parameter values.

$$\text{Model 1:}\quad A^0 = B^0 = 2.0, \beta^0 = 0.2.$$

$$\text{Model 2:}\quad A_1^0 = B_1^0 = 2.5, \beta_1^0 = 0.8,$$

$$A_2^0 = B_2^0 = 2.0, \beta_2^0 = 1.2.$$

The data have been generated using these parameters and then contaminated by the following autoregressive of order one process.

$$X(t) = \phi X(t-1) + e(t), \quad \phi = 0.75.$$

We have taken the error variance, the variance of $X(t)$, as σ_X^2 and for reporting purposes three values of $\sigma_X^2 = 1.0, 2.0$ and 3.0 have been considered. The sequence of i.i.d. random variables $\{e(t)\}$ is from a Gaussian process with mean zero and variance $\sigma^2 = (1-\phi^2)\sigma_X^2$. The chosen sample sizes in these experiments are $100, 200, \ldots, 900$ and 1000. In case of both the models, the data are generated using the chosen parameter values and different values of σ_X^2, the variance of the error process $\{X(t)\}$ and sample sizes n. For every value of n and σ_X^2, 2000 realizations are generated and estimates are obtained using the proposed MNR algorithm. In case of Model 1, in each iteration, we generate a sample and find the initial estimate from $R(\beta)$ using a grid search over the grid $\big\{\frac{\pi k}{n^2}, k = 1, \ldots, n^2-1\big\}$. The iterative process is terminated if the absolute difference between two consecutive iterates is less than $n^{-5/2}$. In most of the cases, the iterations terminate in 4-6 iterations. We have computed the average estimates (AEs), mean squares errors (MSEs), asymptotic variances of LSEs (ASYV) and asymptotic variances of the MNR estimators (ASYV-P). In addition, we have also calculated the MSEs of LSEs (LSE-MSE) and efficient estimators (EFF-MSE) proposed by Mittal et al. [18] for comparison purposes. In case of Model 2, we have ${A_1^0}^2 + {B_1^0}^2 > {A_2^0}^2 + {B_2^0}^2$, so we first implement the same steps as Model 1, to estimate β_1. Then we repeat the same process after removing the effect of the first chirp rate parameter, that is β_1, to estimate the parameters of the second component. We are mainly concerned in estimation of the chirp rate parameters, therefore the results related to the chirp rate parameter estimates are reported. The MSEs of the MNR method, LSEs and efficient estimators along with the asymptotic variances of both the LSE and the proposed MNR estimator of β in case of Model 1 are plotted in Figure 1 for different sample sizes on negative logarithmic scales. The three sub-plots are corresponding to values of σ_X^2. The results for β_1 and β_2, present in Model 2, are provided in Figures 2 and 3, respectively. The following observations can be made based on the above experiments.

(1) MSEs of the MNR estimators increase as the error variance increases.
(2) In the other way, MSEs of the proposed estimators rapidly decrease as the sample size increases. This verifies the consistency property of the estimators.
(3) MSEs of the MNR estimators are generally smaller than those of the LSEs and efficient estimators which is expected as asymptotic variances of LSEs as well as the efficient estimators are larger than the asymptotic variances of the proposed estimator.

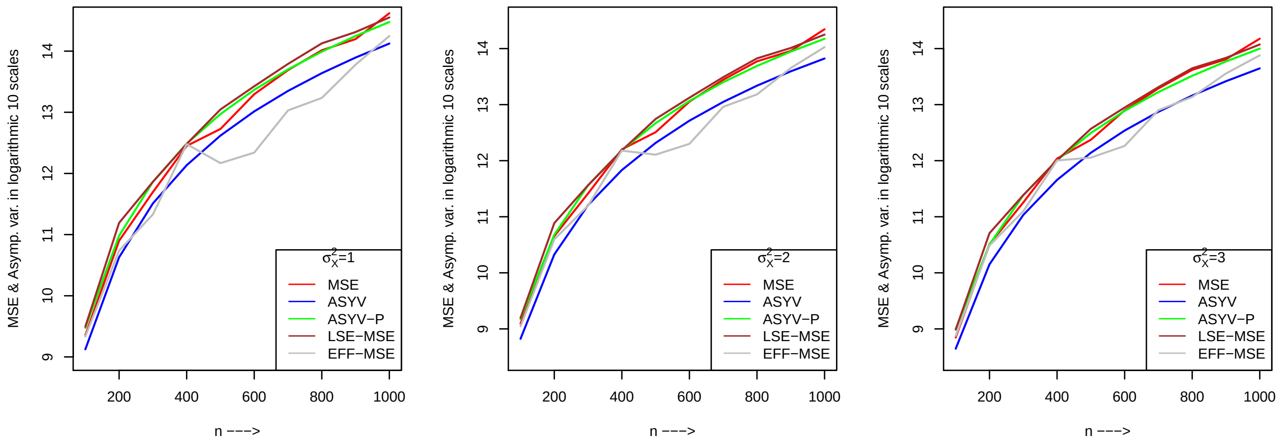

FIGURE 1. In each sub-plot, the graphed lines represent the MSEs of the MNR estimate, the asymptotic variance of LSE and the asymptotic variance of MNR estimate of β.

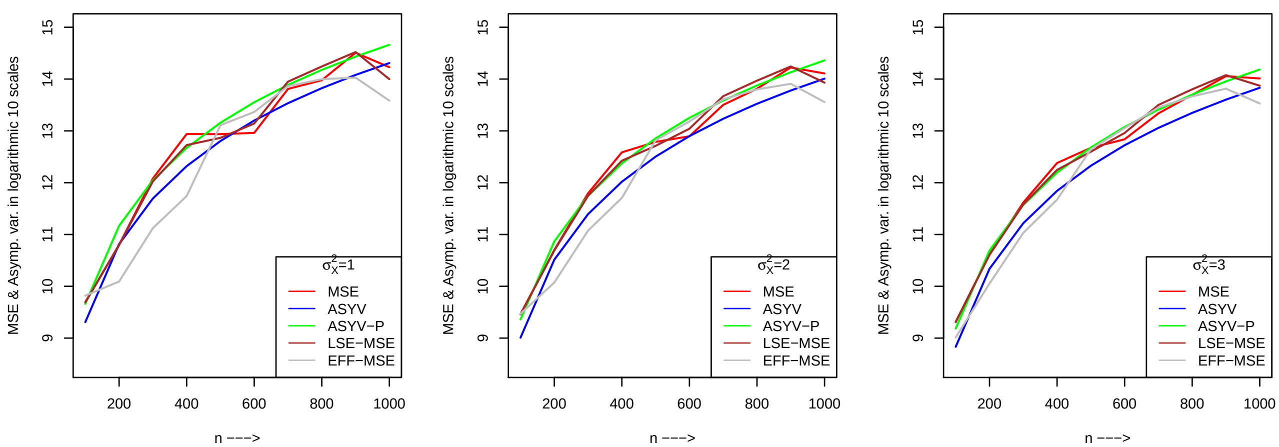

FIGURE 2. In each sub-plot, the graphed lines represent the MSEs of the MNR estimate, the asymptotic variance of LSE and the asymptotic variance of MNR estimate of β_1.

(4) MSEs of the proposed MNR estimators are usually smaller than the asymptotic variances of the LSEs.

(5) MSEs of the proposed MNR estimators are very close to the asymptotic variances of the proposed estimators.

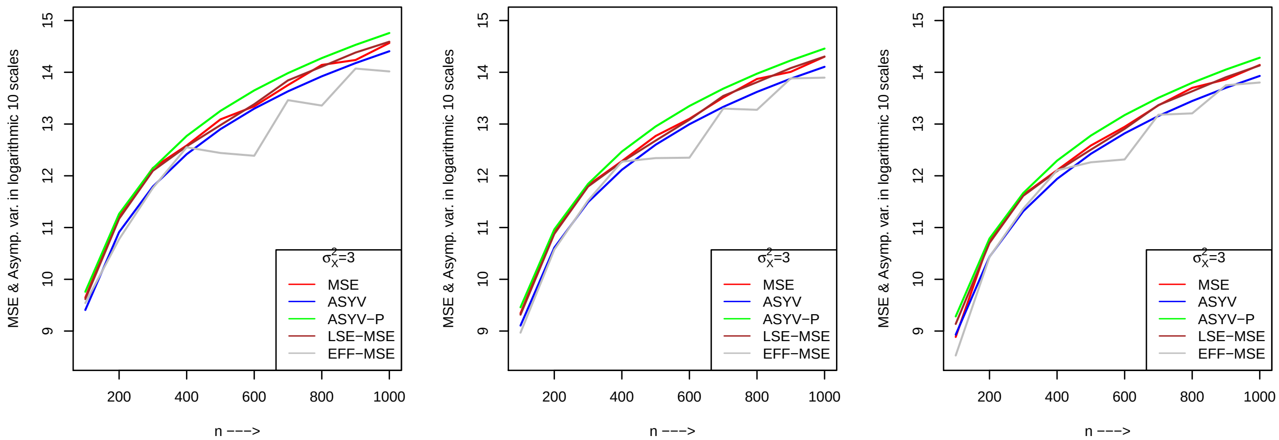

FIGURE 3. In each sub-plot, the graphed lines represent the MSEs of the MNR estimate, the asymptotic variance of LSE and the asymptotic variance of MNR estimate of β_2.

7. Data Analysis

In this section, we analyze two real datasets. The datasets are taken from UCI Machine Learning Repository, one is of type sonar rocks and other is of type sonar mines. Each dataset contains 60 observations. We have used the sonar rocks data CR036 and sonar mines data CM078 for illustrative purposes and observe that elementary chirp model can very effectively be used to analyze and compress such data.

The elementary chirp model is fitted to the mean corrected data. Exploratory analysis suggests that a seven component model is suitable for sonar rocks data CR036 and a eight component model for sonar mines data CM078. The mean-corrected raw data is plotted in Figure 4(a) and in 4(b) the same is plotted along with the fitted data. The residuals in fitting an elementary chirp model of order seven to data CR036 is plotted in Figure 4(c). Here one of the aim is to check whether the residuals from this fit satisfies the assumption of a stationary error. It has been observed that the residuals are random based on run test and they are i.i.d. based on minimum AIC and BIC in the class of stationary ARMA processes. For the other dataset sonar mines CM078, similar findings are observed and provided in Figure 5. In sub-plot (a) of Figure 5, the mean-corrected raw data is plotted and in sub-plot (b), the mean-corrected observed data along with the fitted values from a eight component chirp model are presented. Similar to Figure 4, in sub-plot (c), the residuals are plotted. We find similar structure of the estimated residuals as it was in data CR036. This analysis

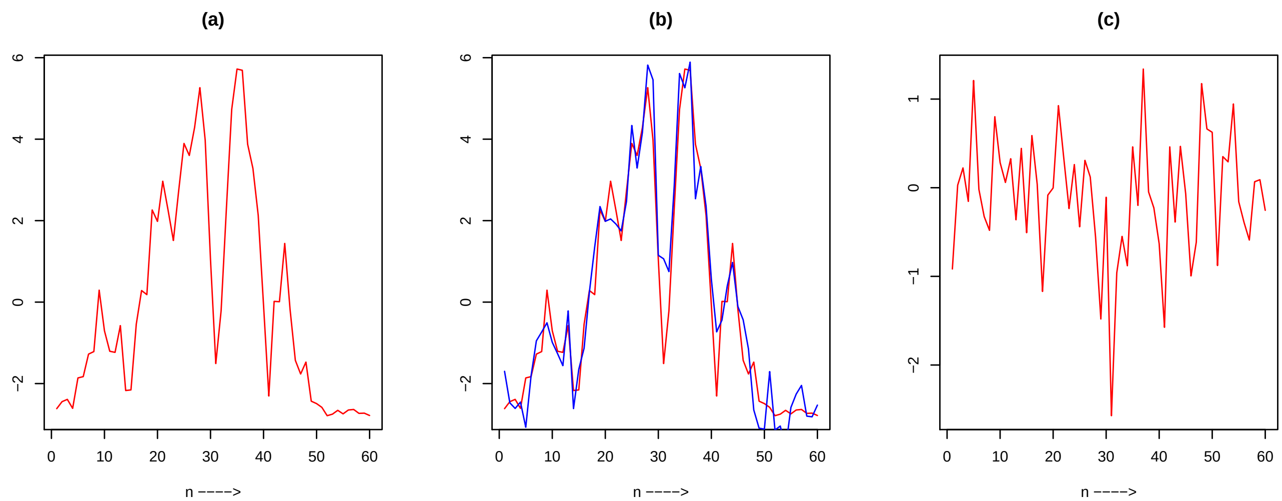

FIGURE 4. (a) The mean-corrected CR036 data. (b) The mean-corrected CR036 data (red) and the fitted values (blue) from a seven component model. (c) The residuals from this fit. Here the y-axis represents energy.

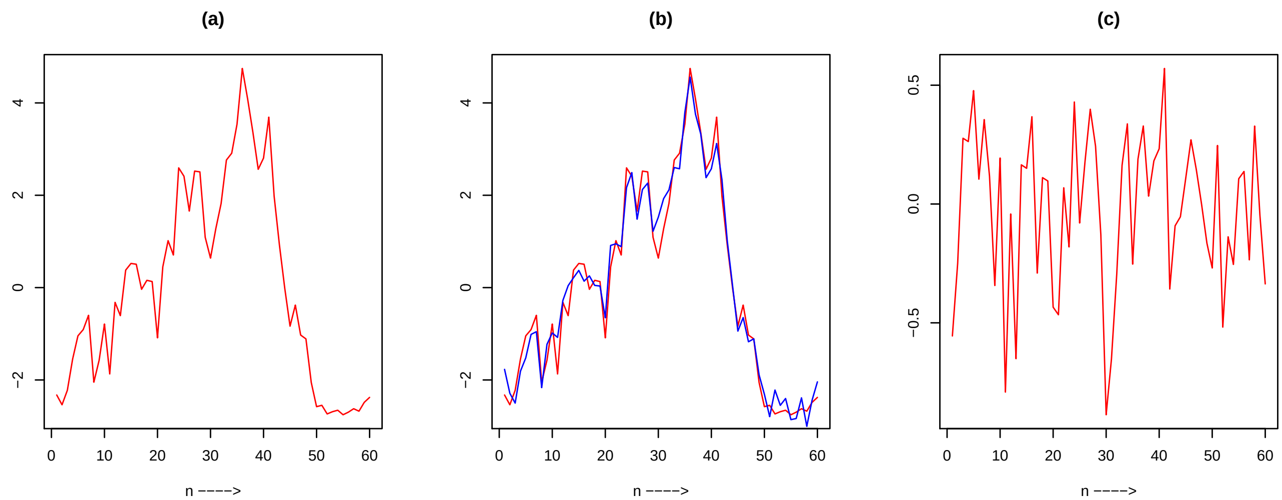

FIGURE 5. (a) The mean-corrected CM078 data. (b) The mean-corrected CM078 data (red) and the fitted values (blue)from a eight component model. (c) The residuals from this fit. Here the y-axis represents energy.

suggests that the elementary chirp model is suitable in analyzing these sonar data sets and the sub-plots (b) in both Figure 4 and Figure 5 shows that using MNR algorithm, the fitted data look good.

8. Conclusions

In this paper we have addressed the problem of efficient estimation of unknown parameters in an elementary chirp model. In elementary chirp models, chirp rates are the only non-linear parameters and the estimation of these parameters is a problem of interest. The NR Algorithm is an important method in non-linear regression problem. But it does not work well for sinusoidal/chirp type class of models in finding LSEs. We have proposed a modification of this algorithm by reducing the correction factor in each step of the algorithm. This modification results in efficient estimation of the chirp parameters and works very well in practice. The asymptotic variances of the proposed estimators are smaller than those of the LSEs. The proposed method can handle estimation of parameters in multiple elementary chirp models using a sequential procedure. Extensive simulations have been performed to compare the performances of the proposed estimators with the LSEs and the efficient estimators in terms of the MSEs. The MSEs of the proposed estimators are in general less than the MSEs of LSEs and the efficient estimators as well as the asymptotic variances of the LSEs and efficient estimators. The performances of the proposed estimators are quite satisfactory. The algorithm does not require any optimization, computationally less intensive and the iterations converge very fast in few steps.

Acknowledgments

The authors would also like to thank Ms. Anjali Mittal, IIT Kanpur, India for calculating the LSEs and efficient estimators reported in Numerical Analysis section. The authors would like to thank the reviewers and the Associate Editor for their constructive comments which have helped to improve the paper significantly.

Data Availability and Conflict of Interest Statements

The data are available publicly in the UCI Machine Learning Repository. The authors do not have any conflict of interest in the preparation of this manuscript.

Appendix A

Lemma 1. *If $\alpha, \beta \in (0, \pi)$, and $\alpha \neq \beta$, then except for countable number of points, the following results hold.*

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

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

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

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

Lemma 2. *If $\beta \in (0, \pi)$ then except for countable number of points, the following results hold.*

$$\sum_{t=1}^{n} t^k\cos^2(\beta t^2) = \frac{n^{k+1}}{2(k+1)} + O(n^k) = \sum_{t=1}^{n} t^k\sin^2(\beta t^2),$$

for $k = 0, 1, \ldots$

Lemma 3. *Let $\tilde{\beta} - \beta = O_p(n^{-2-\delta})$, then*

$$\sum_{t=1}^{n} t^k\ e^{-i\tilde{\beta}t^2} = O_p(n^k), \quad k = 1, 2, \ldots$$

$$\sum_{t=1}^{n} y(t)\cos(\tilde{\beta}t^2) = \frac{n}{2}\left(A + O_p(n^{-\delta})\right) \qquad \text{and} \qquad \sum_{t=1}^{n} y(t)\sin(\tilde{\beta}t^2) = \frac{n}{2}\left(B + O_p(n^{-\delta})\right),$$

where $y(t)$ is given in equation (10).

Proof. of Lemma 3: Consider

$$
\begin{aligned}
\sum_{t=1}^{n} y(t)e^{-i\tilde{\beta}t^2} &= \sum_{t=1}^{n}\Big[A^0\cos(\beta^0t^2)+B^0\sin(\beta^0t^2)+X(t)\Big]e^{-i\tilde{\beta}t^2}\\
&= \sum_{t=1}^{n}\left[\frac{A^0}{2}\left(e^{i\beta^0t^2}+e^{-i\beta^0t^2}\right)+\frac{B^0}{2i}\left(e^{i\beta^0t^2}-e^{-i\beta^0t^2}\right)+X(t)\right]e^{-i\tilde{\beta}t^2}\\
&= \left(\frac{A^0}{2}+\frac{B^0}{2i}\right)\sum_{t=1}^{n}e^{i(\beta^0-\tilde{\beta})t^2}+\left(\frac{A^0}{2}-\frac{B^0}{2i}\right)\sum_{t=1}^{N}e^{-i(\beta^0+\tilde{\beta})t^2}+\sum_{t=1}^{n}X(t)e^{-i\tilde{\beta}t^2}.
\end{aligned}
\tag{15}
$$

Observe that when $\tilde{\beta}-\beta^0 = O_p(n^{-2-\delta})$ following Bai et al. [6] and Lemma 1 we have

$$
\sum_{t=1}^{n} e^{-i(\beta^0+\tilde{\beta})t^2} = O_p(1) \tag{16}
$$

$$
\sum_{t=1}^{n} e^{i(\beta^0-\tilde{\beta})t^2} = O_p(1). \tag{17}
$$

Now consider

$$
\begin{aligned}
\sum_{t=1}^{n} e^{i(\beta-\tilde{\beta})t^2} &= \sum_{t=1}^{n} e^{i(\beta-\beta)t^2} + i(\beta-\tilde{\beta})\sum_{t=1}^{n} t^2 e^{i(\beta-\beta^*)t^2}\\
&= n + O_p(n^{-2-\delta})O_p(n^3)\\
&= n + O_p(n^{1-\delta}).
\end{aligned}
$$

Here, β^* is a point between β and $\tilde{\beta}$ and we will be using β^* for this purpose. Now to approximate the third term in equation (15), we choose L large enough, such that $L\delta > 1$. Using Taylor series approximation of $e^{-i\tilde{\beta}t^2}$ up to L-th order terms

$$
\begin{aligned}
\sum_{t=1}^{n} X(t)e^{-i\tilde{\beta}t^2} &= \sum_{t=1}^{n} X(t)\Big[e^{-i\beta t^2}+\sum_{l=1}^{L-1}\frac{(-it^2(\tilde{\beta}-\beta))^l}{l!}e^{-i\beta t^2}+\frac{(-t^2(\tilde{\beta}-\beta))^L}{L!}e^{-i\beta^* t^2}\Big]\\
&= O_p(n^{1/2})+\sum_{l=1}^{L-1}\frac{O_p(n^{-2l-\delta l})}{l!}O_p(n^{2l+1/2})+O_p\left((n^2.n^{-2-\delta})^L.n\right)\\
&= O_p(n^{1/2})+O_p(n^{1/2}-L\delta+\delta)O_p(n^{1-L\delta}) = O_p(n^{1/2}).
\end{aligned}
\tag{18}
$$

Using equations (16)-(18) in equation (15), we have the lemma. ∎

Appendix B

In order to prove Theorem 4.1, we write $\mathbf{Z}(\beta) = \mathbf{Z}$. Then for model as given in equation (10), define Define

$$\mathbf{D} = \mathrm{diag}\{1, 2, \ldots, n\} \quad \mathbf{E} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix},$$

$$\dot{\mathbf{Z}} = \frac{d}{d\beta}\mathbf{Z} = \mathbf{D}^2\mathbf{Z}\mathbf{E}, \quad \ddot{\mathbf{Z}} = \frac{d^2}{d\beta^2}\mathbf{Z} = -\mathbf{D}^4\mathbf{Z}.$$

Note that $\mathbf{E}\mathbf{E} = -\mathbf{I}$, $\mathbf{E}\mathbf{E}^T = \mathbf{I} = \mathbf{E}^T\mathbf{E}$ and since

$$\begin{aligned} \mathbf{0} = \frac{d}{d\beta}\mathbf{I} &= \frac{d}{d\beta}\left[(\mathbf{Z}^T\mathbf{Z})^{-1}(\mathbf{Z}^T\mathbf{Z})\right] \\ &= \left[\frac{d}{d\beta}(\mathbf{Z}^T\mathbf{Z})^{-1}\right](\mathbf{Z}^T\mathbf{Z}) + (\mathbf{Z}^T\mathbf{Z})^{-1}\left[\frac{d}{d\beta}(\mathbf{Z}^T\mathbf{Z})\right] \\ &= \left[\frac{d}{d\beta}(\mathbf{Z}^T\mathbf{Z})^{-1}\right](\mathbf{Z}^T\mathbf{Z}) + (\mathbf{Z}^T\mathbf{Z})^{-1}[\dot{\mathbf{Z}}^T\mathbf{Z} + \mathbf{Z}^T\dot{\mathbf{Z}}], \end{aligned}$$

hence,

$$\frac{d}{d\beta}(\mathbf{Z}^T\mathbf{Z})^{-1} = -(\mathbf{Z}^T\mathbf{Z})^{-1}[\dot{\mathbf{Z}}^T\mathbf{Z} + \mathbf{Z}^T\dot{\mathbf{Z}}](\mathbf{Z}^T\mathbf{Z})^{-1}.$$

The first and second order derivatives $R'(\beta)$ and $R''(\beta)$ of $R(\beta)$ are given by

$$\frac{1}{2}\,R'(\beta) = \mathbf{Y}^T\dot{\mathbf{Z}}(\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T\mathbf{Y} - \mathbf{Y}^T\mathbf{Z}(\mathbf{Z}^T\mathbf{Z})^{-1}\dot{\mathbf{Z}}^T\mathbf{Z}(\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T\mathbf{Y}.$$

Similarly,

$$\begin{aligned} \frac{1}{2}\,R''(\beta) = {} & \mathbf{Y}^T\ddot{\mathbf{Z}}(\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T\mathbf{Y} + \mathbf{Y}^T\dot{\mathbf{Z}}(\mathbf{Z}^T\mathbf{Z})^{-1}\dot{\mathbf{Z}}^T\mathbf{Y} \\ & -\mathbf{Y}^T\dot{\mathbf{Z}}(\mathbf{Z}^T\mathbf{Z})^{-1}(\dot{\mathbf{Z}}^T\mathbf{Z} + \mathbf{Z}^T\dot{\mathbf{Z}})(\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T\mathbf{Y} \\ & -\mathbf{Y}^T\dot{\mathbf{Z}}(\mathbf{Z}^T\mathbf{Z})^{-1}\dot{\mathbf{Z}}^T\mathbf{Z}(\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T\mathbf{Y} \\ & +\mathbf{Y}^T\mathbf{Z}(\mathbf{Z}^T\mathbf{Z})^{-1}(\dot{\mathbf{Z}}^T\mathbf{Z} + \mathbf{Z}^T\dot{\mathbf{Z}})(\mathbf{Z}^T\mathbf{Z})^{-1}\dot{\mathbf{Z}}^T\mathbf{Z}(\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T\mathbf{Y} \\ & -\mathbf{Y}^T\mathbf{Z}(\mathbf{Z}^T\mathbf{Z})^{-1}(\ddot{\mathbf{Z}}^T\mathbf{Z})(\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T\mathbf{Y} - \mathbf{Y}^T\mathbf{Z}(\mathbf{Z}^T\mathbf{Z})^{-1}(\dot{\mathbf{Z}}^T\dot{\mathbf{Z}})(\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T\mathbf{Y} \\ & +\mathbf{Y}^T\mathbf{Z}(\mathbf{Z}^T\mathbf{Z})^{-1}\dot{\mathbf{Z}}^T\mathbf{Z}(\mathbf{Z}^T\mathbf{Z})^{-1}(\dot{\mathbf{Z}}^T\mathbf{Z} + \mathbf{Z}^T\dot{\mathbf{Z}})(\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T\mathbf{Y} \\ & -\mathbf{Y}^T\mathbf{Z}(\mathbf{Z}^T\mathbf{Z})^{-1}\dot{\mathbf{Z}}^T\mathbf{Z}(\mathbf{Z}^T\mathbf{Z})^{-1}\dot{\mathbf{Z}}^T\mathbf{Y}. \end{aligned}$$

Assume that $\tilde{\beta} - \beta = O_p(n^{-2-\delta})$, hence, for large n, $\left(\frac{1}{n}\mathbf{Z}(\tilde{\beta})^T\mathbf{Z}(\tilde{\beta})\right)^{-1} = 2\ I + o_p(1)$. Then for large n,

$$\begin{aligned}
\frac{1}{2}\ R''(\tilde{\beta}) \;=\;& \frac{2}{n}\mathbf{Y}^T\ddot{\mathbf{Z}}\mathbf{Z}^T\mathbf{Y} - \frac{4}{n^2}\mathbf{Y}^T\dot{\mathbf{Z}}(\dot{\mathbf{Z}}^T\mathbf{Z} + \mathbf{Z}^T\dot{\mathbf{Z}})\mathbf{Z}^T\mathbf{Y} + \frac{2}{n}\mathbf{Y}^T\dot{\mathbf{Z}}\dot{\mathbf{Z}}^T\mathbf{Y} - \frac{4}{n^2}\mathbf{Y}^T\dot{\mathbf{Z}}\dot{\mathbf{Z}}^T\mathbf{Z}\mathbf{Z}^T\mathbf{Y} \\
&+\frac{8}{n^3}\mathbf{Y}^T\mathbf{Z}(\dot{\mathbf{Z}}^T\mathbf{Z} + \mathbf{Z}^T\dot{\mathbf{Z}})\dot{\mathbf{Z}}^T\mathbf{Z}\mathbf{Z}^T\mathbf{Y} - \frac{4}{n^2}\mathbf{Y}^T\mathbf{Z}\ddot{\mathbf{Z}}^T\mathbf{Z}\mathbf{Z}^T\mathbf{Y} - \frac{4}{n^2}\mathbf{Y}^T\mathbf{Z}\dot{\mathbf{Z}}^T\dot{\mathbf{Z}}\mathbf{Z}^T\mathbf{Y} \\
&+\frac{8}{n^3}\mathbf{Y}^T\mathbf{Z}\dot{\mathbf{Z}}^T\mathbf{Z}(\dot{\mathbf{Z}}^T\mathbf{Z} + \mathbf{Z}^T\dot{\mathbf{Z}})\mathbf{Z}^T\mathbf{Y} - \frac{4}{n^2}\mathbf{Y}^T\mathbf{Z}\dot{\mathbf{Z}}^T\mathbf{Z}\dot{\mathbf{Z}}^T\mathbf{Y} + o_p(1).
\end{aligned}$$

Substituting $\dot{\mathbf{Z}}$ and $\ddot{\mathbf{Z}}$ in terms of $\mathbf{D}$, $\mathbf{Z}$ AND $\mathbf{E}$, we obtain

$$\begin{aligned}
\frac{1}{2}\ R''(\tilde{\beta}) \;=\;& -\frac{2}{n}\mathbf{Y}^T\mathbf{D}^4\mathbf{Z}\mathbf{Z}^T\mathbf{Y} - \frac{4}{n^2}\mathbf{Y}^T\mathbf{D}^2\mathbf{Z}\mathbf{E}(\mathbf{E}^T\mathbf{Z}^T\mathbf{D}^2\mathbf{Z} + \mathbf{Z}^T\mathbf{D}^2\mathbf{Z}\mathbf{E})\mathbf{Z}^T\mathbf{Y} \\
&+\frac{2}{n}\mathbf{Y}^T\mathbf{D}^2\mathbf{Z}\mathbf{E}\mathbf{E}^T\mathbf{Z}^T\mathbf{D}^2\mathbf{Y} - \frac{4}{n^2}\mathbf{Y}^T\mathbf{D}^2\mathbf{Z}\mathbf{E}\mathbf{E}^T\mathbf{Z}^T\mathbf{D}^2\mathbf{Z}\mathbf{Z}^T\mathbf{Y} \\
&+\frac{8}{n^3}\mathbf{Y}^T\mathbf{Z}(\mathbf{E}^T\mathbf{Z}^T\mathbf{D}^2\mathbf{Z} + \mathbf{Z}^T\mathbf{D}^2\mathbf{Z}\mathbf{E})\mathbf{E}^T\mathbf{Z}^T\mathbf{D}^2\mathbf{Z}\mathbf{Z}^T\mathbf{Y} \\
&+\frac{4}{n^2}\mathbf{Y}^T\mathbf{Z}\mathbf{Z}^T\mathbf{D}^4\mathbf{Z}\mathbf{Z}^T\mathbf{Y} - \frac{4}{n^2}\mathbf{Y}^T\mathbf{Z}\mathbf{E}^T\mathbf{Z}^T\mathbf{D}^4\mathbf{Z}\mathbf{E}\mathbf{Z}^T\mathbf{Y} \\
&+\frac{8}{n^3}\mathbf{Y}^T\mathbf{Z}\mathbf{E}^T\mathbf{Z}^T\mathbf{D}^2\mathbf{Z}(\mathbf{E}^T\mathbf{Z}^T\mathbf{D}^2\mathbf{Z} + \mathbf{Z}^T\mathbf{D}^2\mathbf{Z}\mathbf{E})\mathbf{Z}^T\mathbf{Y} \\
&-\frac{4}{n^2}\mathbf{Y}^T\mathbf{Z}\mathbf{E}^T\mathbf{Z}^T\mathbf{D}^2\mathbf{Z}\mathbf{E}^T\mathbf{Z}^T\mathbf{D}^2\mathbf{Y} + o_p(1).
\end{aligned}$$

Using Lemma 2, we can approximate the following;

$$\begin{aligned}
&\frac{1}{n^3}\mathbf{Y}^T\mathbf{D}^2\mathbf{Z} = \begin{pmatrix}\frac{1}{6}A & \frac{1}{6}B\end{pmatrix} + O_p(\frac{1}{n}), \\
&\frac{1}{n^5}\mathbf{Y}^T\mathbf{D}^4\mathbf{Z} = \begin{pmatrix}\frac{1}{10}A & \frac{1}{10}B\end{pmatrix} + O_p(\frac{1}{n}), \\
&\frac{1}{n}\mathbf{Z}^T\mathbf{Y} = \frac{1}{2}(A \quad B)^T + O_p(\frac{1}{n}) \\
&\frac{1}{n^3}\mathbf{Z}^T\mathbf{D}^2\mathbf{Z} = \frac{1}{6}I + O_p(\frac{1}{n}), \\
&\frac{1}{n^5}\mathbf{Z}^T\mathbf{D}^4\mathbf{Z} = \frac{1}{10}I + O_p(\frac{1}{n}).
\end{aligned}$$

Therefore, for large n

$$\begin{aligned}
\frac{1}{2n^5}\ R''(\tilde{\beta}) \;&=\; (A^2+B^2)\left[-\frac{1}{10} - 0 + \frac{1}{18} - \frac{1}{18} + 0 + \frac{1}{10} - \frac{1}{10} + 0 + \frac{1}{18}\right] + O_p(\frac{1}{n}) \\
&=\; -\frac{2}{45}(A^2+B^2) + O_p(\frac{1}{n}).
\end{aligned}$$

In order to understand the terms of $\frac{1}{2}\ R'(\tilde{\beta})$, we consider the first term initially ignoring the lower order terms.

$$\begin{aligned}\mathbf{Y}^T\dot{\mathbf{Z}}(\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T\mathbf{Y} = \frac{2}{n}\mathbf{Y}^T\mathbf{D}^2\mathbf{Z}\mathbf{E}\mathbf{Z}^T\mathbf{Y} \ &= \ \frac{2}{n}\left(\sum_{t=1}^{n} y(t)t^2\cos(\tilde{\beta}t^2)\right)\left(\sum_{t=1}^{n} y(t)\sin(\tilde{\beta}t^2)\right) \\ &\quad -\frac{2}{n}\left(\sum_{t=1}^{n} y(t)t^2\sin(\tilde{\beta}t^2)\right)\left(\sum_{t=1}^{n} y(t)\cos(\tilde{\beta}t^2)\right).\end{aligned}$$

Using Lemma 3, we have

$$\sum_{t=1}^{n} y(t)\cos(\tilde{\beta}t^2) = \frac{n}{2}\left(A + O_p(n^{-\delta})\right) \quad \text{and} \quad \sum_{t=1}^{n} y(t)\sin(\tilde{\beta}t^2) = \frac{n}{2}\left(B + O_p(n^{-\delta})\right).$$

Similarly as the proof of Lemma 3, we can write

$$\begin{aligned}\sum_{t=1}^{n} y(t)t^2 e^{-i\tilde{\beta}t^2} \ &= \ \sum_{t=1}^{n}\left(A\cos(\beta t^2) + B\sin(\beta t^2) + X(t)\right)t^2 e^{-i\tilde{\beta}t^2} \\ &= \ \frac{1}{2}(A - iB)\sum_{t=1}^{n} t^2\ e^{i(\beta-\tilde{\beta})t^2} + \frac{1}{2}(A + iB)\sum_{t=1}^{n} t^2\ e^{-i(\beta+\tilde{\beta})t^2} + \sum_{t=1}^{n} X(t)t^2 e^{-i\tilde{\beta}t^2}.\end{aligned}$$

It can be observed that

$$\begin{aligned}\sum_{t=1}^{n} t^2\ e^{-i(\beta+\tilde{\beta})t^2} \ &= \ O_p(n^2), \\ \sum_{t=1}^{n} t^2\ e^{i(\beta-\tilde{\beta})t^2} \ &= \ \sum_{t=1}^{n} t^2 + i(\beta-\tilde{\beta})\sum_{t=1}^{n} t^4 - \frac{1}{2}(\beta-\tilde{\beta})^2\sum_{t=1}^{n} t^6 \\ &\quad -\frac{1}{6}i(\beta-\tilde{\beta})^3\sum_{t=1}^{n} t^8 + \frac{1}{24}(\beta-\tilde{\beta})^4\sum_{t=1}^{n} t^{10}e^{i(\beta-\beta^*)t^2}.\end{aligned}$$

The last term in the above expansion is of order $O_p(n^{2-4\delta})$. Using Taylor series expansion of $e^{-i\tilde{\beta}t^2}$, we can approximate as

$$\begin{aligned}\sum_{t=1}^{n} X(t)t^2 e^{-i\tilde{\beta}t^2} \ &= \ \sum_{t=1}^{n} X(t)t^2\Big[e^{-i\beta t^2} + \sum_{l=1}^{L-1}\frac{(-it^2(\tilde{\beta}-\beta))^l}{l!}e^{-i\beta t^2} + \frac{(-t^2(\tilde{\beta}-\beta))^L}{L!}e^{-i\beta^* t^2}\Big] \\ &= \ \sum_{t=1}^{n} X(t)t^2 e^{-i\beta t^2} + \sum_{l=1}^{L-1}\frac{O_p(n^{-2l-\delta l})}{l!}O_p(n^{2l+5/2}) + O_p((n^2.n^{-2-\delta})^L.n^3) \\ &= \ \sum_{t=1}^{n} X(t)t^2 e^{-i\beta_j t^2} + O_p(n^{7/2-L\delta}).\end{aligned}$$

Now consider

$$\begin{aligned}\sum_{t=1}^{n} y(t)t^2\cos(\tilde{\beta}t) &= \frac{1}{2}\left[A\left(\sum_{t=1}^{n} t^2 - \frac{1}{2}(\beta-\tilde{\beta})^2\sum_{t=1}^{n} t^6\right) + B\left(\sum_{t=1}^{n}(\beta-\tilde{\beta})t^4 - \frac{1}{6}(\beta-\tilde{\beta})^3\sum_{t=1}^{n} t^8\right)\right] \\ &+ \sum_{t=1}^{n} X(t)t^2\cos(\beta t^2) + O_p(n^{7/2-L\delta}) + O_p(n^2) + O_p(n^{2-4\delta}),\end{aligned}$$

and

$$\begin{aligned}\sum_{t=1}^{n} y(t)t^2\sin(\tilde{\beta}t) &= \frac{1}{2}\left[B\left(\sum_{t=1}^{n} t^2 - \frac{1}{2}(\beta-\tilde{\beta})^2\sum_{t=1}^{n} t^6\right) - A\left(\sum_{t=1}^{n}(\beta-\tilde{\beta})t^4 - \frac{1}{6}(\beta-\tilde{\beta})^3\sum_{t=1}^{n} t^8\right)\right] \\ &+ \sum_{t=1}^{n} X(t)t^2\sin(\beta t^2) + O_p(n^{7/2-L\delta}) \\ &+ O_p(n^2) + O_p(n^{2-4\delta}).\end{aligned}$$

Therefore,

$$\begin{aligned}\widehat{\beta} &= \tilde{\beta} - \frac{4}{9}\,\frac{R'(\tilde{\beta})}{R''(\tilde{\beta})} = \tilde{\beta} - \frac{4}{9}\,\frac{\frac{1}{2n^5}R'(\tilde{\beta})}{-\frac{2}{45}(A^2+B^2)+O_p(\frac{1}{n})} \\ &= \tilde{\beta} - \frac{4}{9}\,\frac{\frac{2}{n^6}\mathbf{Y}^T\mathbf{D}^2\mathbf{ZEZ}^T\mathbf{Y}}{-\frac{2}{45}(A^2+B^2)+O_p(\frac{1}{n})} = \tilde{\beta} + 20\frac{\frac{1}{n^6}\mathbf{Y}^T\mathbf{D}^2\mathbf{ZEZ}^T\mathbf{Y}}{(A^2+B^2)+O_p(\frac{1}{n})} \\ &= \tilde{\beta} + 20\frac{\frac{1}{4n^5}(A^2+B^2)}{(A^2+B^2)+O_p(\frac{1}{n})} \times \left\{(\beta-\tilde{\beta})\sum_{t=1}^{n} t^4 - \frac{1}{6}(\beta-\tilde{\beta})^3\sum_{t=1}^{n} t^8\right\} \\ &\quad + \frac{10}{(A^2+B^2)n^5+O_p(\frac{1}{n})}\left[B\sum_{t=1}^{n} X(t)t^2\cos(\beta t^2) + A\sum_{t=1}^{n} X(t)t^2\sin(\beta t^2)\right] \\ &+ O_p(n^{-\frac{3}{2}-L\delta}) + O_p(n^{-3}) + O_p(n^{-1-4\delta}) \\ &= \beta + (\beta-\tilde{\beta})O_p(n^{-2\delta}) \\ &+ \frac{10}{(A^2+B^2)n^5+O_p(\frac{1}{n})}\left[B\sum_{t=1}^{n} X(t)t^2\cos(\beta t) + A\sum_{t=1}^{n} X(t)t^2\sin(\beta t)\right] \\ &+ O_p(n^{-\frac{3}{2}-L\delta}) + O_p(n^{-3}) + O_p(n^{-1-4\delta}).\end{aligned}$$

Hence, if $\delta \le \frac{1}{6}$, then clearly $\widehat{\beta} - \beta = O_p(n^{-2-3\delta})$ as $(\beta-\tilde{\beta})O_p(n^{-2\delta})$ is the dominating term and rest of the terms are of smaller order. If $\delta > \frac{1}{6}$, then the behavior of $\widehat{\beta}$ is governed by the third term and

$$
\begin{aligned}
n^{\frac{5}{2}}(\widehat{\beta}-\beta) &\stackrel{d}{=} \frac{10n^{-\frac{5}{2}}}{(A^2+B^2)}\left[B\sum_{t=1}^{n}X(t)t^2\cos(\beta t^2)+A\sum_{t=1}^{n}X(t)t^2\sin(\beta t^2)\right] \\
&= \frac{10n^{-\frac{5}{2}}}{(A^2+B^2)}\left[B\sum_{k=-\infty}^{\infty}a(k)\sum_{t=1}^{n}e(t-k)t^2\cos(\beta t^2)+A\sum_{k=-\infty}^{\infty}a(k)\sum_{t=1}^{n}e(t-k)t^2\sin(\beta t^2)\right] \\
&\xrightarrow{d} \mathcal{N}\left(0,\frac{10\sigma^2 c}{A^2+B^2}\right).
\end{aligned}
$$

This is the distribution of NR estimator of the chirp parameter β.

References

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

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

[3] Abratkiewicz, K., Gromek, D., Samczynski, P.: Chirp rate estimation and micro-Doppler signatures for pedestrian security radar systems. 2019 IEEE Signal Processing Symposium, pp. 212–215, (2019)

[4] Abratkiewicz, K. and Samczynski, P.: A block method using the chirp rate estimation for NLFM radar pulse reconstruction. Sensors. 19, (2019)

[5] Bai, G., Cheng, Y., Tang, W., Li, S.: Chirp rate estimation for LFM signal by multiple DPT and weighted combination. IEEE Signal Processing Letters. 26, 149–153 (2018)

[6] Bai, Z.D., Rao, C.R., Chow, M., Kundu, D.: Efficient algorithm for estimating the parameters of superimposed complex exponential model. Journal of Statistical Planning and Inference. 110, 23-34 (2003)

[7] Casazza, P.G. and Fickus, M: Fourier transforms of finite chirps. Eurasip Journal on Advances in Signal Processing. 1–7 (2006)

[8] Ding, M., Wei, Y., Yu, L.: Parameter Estimation of a Single Chirp in the Presence of Wiener Phase Noise With Unknown Variance. IEEE Transactions on Signal Processing. 72, 3171–3186 (2024)

[9] Djurović, I.: Viterbi algorithm for chirp-rate and instantaneous frequency estimation. Signal Processing. 91, 1308–1314 (2011)

[10] Djurović, I., Ioana, C., Stanković, L., Wang, P.: Adaptive algorithm for chirp-rate estimation. EURASIP Journal on Advances in Signal Processing. 1–9 (2009)

[11] Djurović, I., Thayaparan, T., Stankovic, L.: SAR imaging of moving targets using polynomial Fourier transform. IET Signal Processing. 2, 237–246 (2008)

[12] Fourer, D., Auger, F., Czarnecki, K., Meignen, S., Flandrin, P.: Chirp rate and instantaneous frequency estimation: application to recursive vertical synchrosqueezin. IEEE Signal Processing Letters. 24, 1724–1728 (2017)

[13] Kundu, D., Bai, Z.D., Nandi, S., Bai, L.: Super efficient frequency estimator. Journal of Statistical Planning and Inference. 141, 2576 - 2588 (2011)

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

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

[16] Mboup, M. and Adali, T.: A generalization of the Fourier transform and its application to spectral analysis of chirp-like signals. Applied and Computational Harmonic Analysis. 32, 305–312 (2012

[17] Miao, H.: Generalized spectrum analysis of Chirp Cyclostationary signals associate with linear canonical transform. Signal Processing. 218, paper # 109385 (2024)

[18] Mittal, A., Grover, R., Kundu, D., Mitra, A.: Efficient methodology for estimating elementary chirp model parameters. IEEE Transactions on Signal Processing. 71, 2352 - 2365 (2023)

[19] Mittal, A., Kundu, D., Mitra, A., Grover, R.: Estimation of the Elementary Chirp Model Parameters. IEEE Transactions on Aerospace and Electronic Systems. 59, 5218 - 5234 (2023)

[20] Moghadasian, S.S.: A Fast and Accurate Method for Parameter Estimation of Multi-Component LFM Signals. IEEE Signal Processing Letters. 29, 1719 - 1723 (2022)

[21] Nandi, S. and Kundu, D.: Estimating the fundamental frequency using modified Newton-Raphson algorithm. Statistics. 53, 440 - 458 (2019)

[22] Pan, P., Zhang, Y., Deng, Z., Qi, W.: Deep Learning-Based 2-D Frequency Estimation of Multiple Sinusoidals. IEEE Transactions on Neural Networks and Learning Systems. 33, 5429 – 5440 (2022)

[23] Pang, C., Liu, S., Han, Y.: Coherent detection algorithm for radar maneuvering targets based on discrete polynomial-phase transform. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 12, 3412–3422 (2019)

[24] Rice, J.A. and Rosenblatt, M.: On frequency estimation. Biometrika. 75, 477 - 484 (1988)

[25] Richards, F.S.G.: A method of maximum likelihood estimation. Journal of the Royal Statistical Society B. 23, 469-475 (1961)

[26] Santhanam, B. and Santhanam, T.S.: On refinements to QMFD based chirp parameter estimation. Signal Processing. 213, paper # 109217 (2023)

[27] Shukla, A., Kundu, D., Mitra, A., Grover, R.: On estimating parameters of a multi-component Chirp Model with equal chirp rates. IEEE Transactions on Aerospace and Electronic Systems. 59, 5027 - 5038 (2023)

[28] Shukla, A., Kundu, D., Mitra, A.: Model order estimation in multi-component chirp signals. Signal Processing. 221, paper # 109503 (2024)

[29] Song, J., Xu, Y. Liu, Y., Zhang, Y.: Investigation on estimator of chirp rate and initial frequency of LFM signals based on modified discrete chirp Fourier transform. Circuits, Systems, and Signal Processing. 38, 5861–5882 (2019)

[30] Swiercz, E., Janczak, D., Konopko, K.: Detection of LFM Radar Signals and Chirp Rate Estimation Based on Time-Frequency Rate Distribution. Sensors. 16, 5415, (2021)

[31] Wang, P., Li, H., Djurović, I., Himed, B.: Integrated cubic phase function for linear FM signal analysis. IEEE Transactions on Aerospace and Electronic Systems. 46, 963-977 (2010)

[32] Wang, P. and Yang, J.: Multicomponent chirp signals analysis using product cubic phase function. Digital Signal Processing. 16, 654–669 (2006)

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

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