Noname manuscript No.
(will be inserted by the editor)

Asymptotic properties of least squares estimators and sequential least squares estimators of a chirp-like signal model parameters

Rhythm Grover[1,2] · Debasis Kundu[1] · Amit Mitra[1]

Received: date / Accepted: date

Abstract Sinusoidal model and chirp model are the two fundamental models in digital signal processing. Recently, a chirp-like model was introduced by Grover et al. [12]. A chirp-like model is a generalization of a sinusoidal model and provides an alternative to a chirp model. We derive, in this paper, the asymptotic properties of least squares estimators and sequential least squares estimators of the parameters of a chirp-like signal model. It is observed, theoretically as well as through extensive numerical computations that the sequential least squares estimators perform at par with the usual least squares estimators. The computational complexity involved in the sequential algorithm is significantly lower than that involved in calculating the least squares estimators. This is achieved by exploiting the orthogonality structure of the different components of the underlying model. The performances of both the estimators for finite sample sizes are illustrated by simulation results. In the specific real-life data analyses of signals, we show that a chirp-like signal model is capable of modeling phenomena that can be otherwise modeled by a chirp signal model, in a computationally more efficient manner.

Keywords sinusoidal model · chirp model · chirp-like model · least squares · sequential least squares · asymptotic properties

1 Introduction

Underlying a great deal of signal processing applications such as speech and music processing [28], electrocardiography [30], seismology [32], astronomy [41] and economics [13], are sinusoidal signals embedded in noise. In a general form, a sinusoidal signal can be written as:

$$y(t) = \sum_{j=1}^{p} \{A_j^0 \cos(\alpha_j^0 t) + B_j^0 \sin(\alpha_j^0 t)\} + X(t),\ t = 1, \ldots, n. \tag{1}$$

Here, A_j^0s, B_j^0s are the amplitudes, α_j^0s are the frequencies and $X(t)$ is the random error component of the observed signal $y(t)$. Due to the widespread applicability of this model, many methods have been proposed for its parameter estimation. In this respect, one may look into the monograph of Kundu and Nandi [19]. We also invoke readers to the interesting articles by Kay and Marple [16], Prasad et al. [34] and Stoica [40] for more contributions in this area.

Another important model in digital signal processing is a chirp signal model, encountered in many natural as well as man-made phenomena such as navigational chirps emitted by bats [8], bird sounds [15], human voice [5], radar and sonar systems [25], [42] and communications [11]. Mathematically, a chirp signal is expressed as follows:

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

Here β_j^0s are the frequency rates and again, A_j^0s, B_j^0s are the amplitudes, α_j^0s are the frequencies and $X(t)$ is the random error component of the observed signal $y(t)$. In the last few decades, numerous algorithms have been developed for estimating the unknown parameters of this model. For some of the earliest references on the joint estimation of frequency and frequency rate, see Bello [2], Kelly [17] and Abatzoglou [1]. Thereafter several other estimation methods have been proposed as well, such as methods based on phase unwrapping [6], suboptimal

[1,]Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Kanpur - 208016, India
[2]Corresponding author. E-mail: : rhythm@iitk.ac.in

FFT [33], quadratic phase transform [14], maximum likelihood [37], nonlinear least squares [31], least absolute deviation [21], MCMC based bayesian sampling [26], linear prediction approach [10], Sigmoid transform [23], modified discrete chirp Fourier transform [39] and many more.

Of particular interest to us here, is the method of least squares principle. It is the most commonly used method and is one of the first methods in classical estimation theory. For the chirp model in presence of stationary noise, the least squares estimators (LSEs) are strongly consistent and asymptotically normally distributed. In fact, if the errors are assumed to be Gaussian, LSEs achieve the Cramer Rao lower bound [18]. However, despite these optimal statistical properties, finding them in practice is computationally challenging. The reason behind this is the highly non-linear nature of the least squares surface. Recently, Lahiri et al. [22] proposed the sequential LSEs which have the same statistical properties as the usual LSEs but they reduce the complexity involved in finding the LSEs to a great extent. This is obtained by breaking the multi-dimensional search into multiple two-dimensional searches, using the orthogonality structure of different chirp components present in the model. Nevertheless, there is still a huge computational cost involved in finding the sequential LSEs and there is a need to develop computationally more efficient algorithms for practical implementation.

The subject of the present paper is to introduce a novel model, a chirp-like model, mathematically expressed as follows:

$$y(t) = \sum_{j=1}^{p} \{A_j^0 \cos(\alpha_j^0 t) + B_j^0 \sin(\alpha_j^0 t)\} + \sum_{k=1}^{q} \{C_k^0 \cos(\beta_k^0 t^2) + D_k^0 \sin(\beta_k^0 t^2)\} + X(t),\ t = 1, \ldots, n, \tag{3}$$

where, A_j^0s, B_j^0s, C_k^0s and D_k^0s are the amplitudes, α_j^0s are the frequencies and β_k^0s are the frequency rates. $X(t)$ accounts for the noise present in the signal. This new model is a linear combination of a sinusoidal model (1) and an elementary chirp model[1]. This choice is made mainly because of the following three reasons:

– First, it is observed that this model exhibits same type of behavior as the chirp model (2) and is capable of modeling similar physical phenomena. To demonstrate, we analyze a speech signal data set "UUU" using both a chirp model and a chirp-like model. The corresponding "best" fittings are plotted together in the following figure:

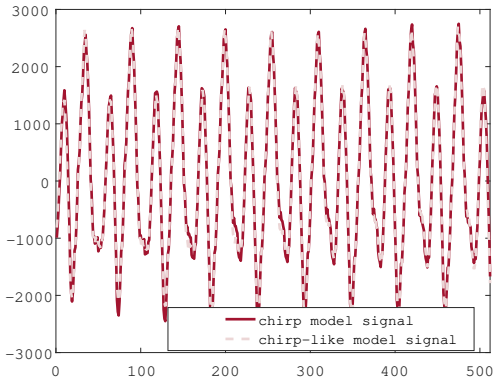

Fig. 1 Fitted chirp signal (red solid line) and fitted chirp-like signal (pink dashed line) to the "UUU" sound data

It is evident from the above figure that the two signals are well-matched.

– Second, this model not only provides an alternative to a chirp model but can be seen as a generalization of a sinusoidal model also. For the special case of $C_k^0 = D_k^0 = 0$ for all $k = 1, \ldots, q$, the proposed model (3) reduces to the sinusoidal model (1).
– Lastly, parameter estimation of this model using a sequential algorithm is computationally simpler and faster compared to the sequential LSEs of a chirp model.

[1]A simple elementary chirp model has the following mathematical expression:

$$y(t) = C^0 \cos(\beta^0 t^2) + D^0 \sin(\beta^0 t^2) + X(t);\ t = 1, \ldots, n. \tag{4}$$

Here C^0, D^0 are the amplitudes, β^0 is the chirp rate , and $X(t)$ is the noise. Although in recent years a lot of work has been done on a chirp model (2), not much attention has been paid on an elementary chirp model. For the reference on model (4), one may refer to Casazza and Fickus [4] and Mboup and Adali [27].

Parameter estimation of a chirp-like model is first formulated as a multidimensional non-linear least squares estimation problem in this paper. We theoretically develop the statistical properties of the LSEs such as strong consistency and asymptotic normality. For a practical solution with computational simplicity, we propose a sequential algorithm. The proposed method turns the multidimensional optimization search into a string of one-dimensional optimization problems. We derive the large-sample properties of the sequential LSEs as well and observe that they have the same properties as the usual LSEs. The theoretical results are then corroborated through extensive simulation studies and a few data analyses.

The rest of the paper is organized as follows. In the next section, we define a one-component chirp-like model and study the asymptotic properties of the LSEs and the sequential LSEs of the parameters of this model. In section 3, we study the asymptotic properties of a more generalised model, a multiple-component chirp-like model (3). In section 4, we perform simulations to validate the asymptotic results and in section 5, we analyze four speech signal data sets and a simulated data set to see how the proposed model performs in practice. We conclude the paper in section 6.

2 One Component Chirp-like Model

In this section we consider a one-component chirp-like model, expressed mathematically as follows:

$$y(t) = A^0 \cos(\alpha^0 t) + B^0 \sin(\alpha^0 t) + C^0 \cos(\beta^0 t^2) + D^0 \sin(\beta^0 t^2) + X(t). \tag{5}$$

Our problem is to estimate the unknown parameters of the model namely A^0, B^0, C^0, D^0, α^0 and β^0 under the following assumption on the noise component:

Assumption 1 *Let Z be the set of integers. $\{X(t)\}$ is a stationary linear process of the form:*

$$X(t) = \sum_{j=-\infty}^{\infty} a(j)e(t-j), \tag{6}$$

where $\{e(t); t \in Z\}$ is a sequence of independently and identically distributed (i.i.d.) random variables with $E(e(t)) = 0$, $V(e(t)) = \sigma^2$, and $a(j)$s are real constants such that

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

For a stationary linear process, this is a standard assumption. This assumption covers a large class of stationary processes. For instance, any finite-dimensional stationary MA, AR, or ARMA process can be formulated in the above-stated representation.

We will use the following notations for further development: $\boldsymbol{\theta} = (A, B, \alpha, C, D, \beta)$, the parameter vector, $\boldsymbol{\theta}^0 = (A^0, B^0, \alpha^0, C^0, D^0, \beta^0)$, the true parameter vector, $\widehat{\boldsymbol{\theta}} = (\widehat{A}, \widehat{B}, \widehat{\alpha}, \widehat{C}, \widehat{D}, \widehat{\beta})$, the LSE of $\boldsymbol{\theta}^0$ and $\boldsymbol{\Theta} = [-M, M] \times [-M, M] \times [0, \pi] \times [-M, M] \times [-M, M] \times [0, \pi]$, where M is a positive real number. Also we make the following assumption on the unknown parameters:

Assumption 2 *The true parameter vector $\boldsymbol{\theta}^0$ is an interior point of the parametric space $\boldsymbol{\Theta}$, and $A^{0^2} + B^{0^2} + C^{0^2} + D^{0^2} > 0$.*

Under these assumptions, we discuss two estimation procedures: the least squares estimation method and the sequential least squares estimation method. We then study the asymptotic properties of the estimators obtained using these methods.

2.1 Least Squares Estimators

The usual LSEs of the unknown parameters of model (5) can be obtained by minimising the error sum of squares:

$$Q(\boldsymbol{\theta}) = \sum_{t=1}^{n} \Big(y(t) - A\cos(\alpha t) - B\sin(\alpha t) - C\cos(\beta t^2) - D\sin(\beta t^2) \Big)^2,$$

with respect to A, B, α, C, D and β simultaneously. In matrix notation,

$$Q(\boldsymbol{\theta}) = (\boldsymbol{Y} - \boldsymbol{Z}(\alpha, \beta)\boldsymbol{\mu})^\top (\boldsymbol{Y} - \boldsymbol{Z}(\alpha, \beta)\boldsymbol{\mu}). \tag{8}$$

Here, $\boldsymbol{Y}_{n\times 1} = \begin{pmatrix} y(1) & \cdots & y(n) \end{pmatrix}^\top$, $\boldsymbol{\mu}_{4\times 1} = \begin{pmatrix} A & B & C & D \end{pmatrix}^\top$ and

$$\boldsymbol{Z}(\alpha,\beta)_{n\times 4} = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) & \cos(\beta) & \sin(\beta) \\ \vdots & \vdots & \vdots & \vdots \\ \cos(n\alpha) & \sin(n\alpha) & \cos(n^2\beta) & \sin(n^2\beta) \end{pmatrix}.$$

Since $\boldsymbol{\mu}$ is a vector of conditionally linear parameters, by separable linear regression technique of Richards [37], we have:

$$\widehat{\boldsymbol{\mu}}(\alpha,\beta) = [\boldsymbol{Z}(\alpha,\beta)^\top \boldsymbol{Z}(\alpha,\beta)]^{-1} \boldsymbol{Z}(\alpha,\beta)^\top \boldsymbol{Y}. \tag{9}$$

Using (9) in (8), we obtain:

$$\begin{aligned} R(\alpha,\beta) &= Q(\widehat{A}(\alpha,\beta), \widehat{B}(\alpha,\beta), \alpha, \widehat{C}(\alpha,\beta), \widehat{D}(\alpha,\beta), \beta) \\ &= \boldsymbol{Y}^\top (\boldsymbol{I} - \boldsymbol{Z}(\alpha,\beta)[\boldsymbol{Z}(\alpha,\beta)^\top \boldsymbol{Z}(\alpha,\beta)]^{-1} \boldsymbol{Z}(\alpha,\beta)^\top) \boldsymbol{Y}. \end{aligned}$$

To obtain $\widehat{\alpha}$ and $\widehat{\beta}$, the LSEs of α^0 and β^0 respectively, we minimise $R(\alpha,\beta)$ with respect to α and β simultaneously. Once we obtain $\widehat{\alpha}$ and $\widehat{\beta}$, by substituting them in (9), we obtain the LSEs of the linear parameters.

The following results provide the consistency and asymptotic normality properties of the LSEs.

Theorem 1 *Under Assumptions 1 and 2, $\widehat{\boldsymbol{\theta}}$ is a strongly consistent estimator of $\boldsymbol{\theta}^0$, that is,*

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

Proof See section B.1.

Theorem 2 *Under Assumptions 1 and 2,*

$$(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0)\boldsymbol{D}^{-1} \xrightarrow{d} \mathcal{N}(0, c\sigma^2 \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}^0)) \text{ as } n \to \infty,$$

where $\boldsymbol{D} = diag(\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}, \frac{1}{n\sqrt{n}}, \frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}, \frac{1}{n^2\sqrt{n}})$, $c = \sum_{j=-\infty}^{\infty} a(j)^2$ *and*

$$\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}^0) = \left(\begin{array}{c|c} {\boldsymbol{\Sigma^{(1)}}}^{-1}(\boldsymbol{\theta}^0) & \boldsymbol{0} \\ \hline \boldsymbol{0} & {\boldsymbol{\Sigma^{(2)}}}^{-1}(\boldsymbol{\theta}^0) \end{array}\right).$$

with

$${\boldsymbol{\Sigma^{(1)}}}^{-1}(\boldsymbol{\theta}^0) = \begin{pmatrix} \frac{2(A^{0^2}+4B^{0^2})}{A^{0^2}+B^{0^2}} & \frac{-6A^0B^0}{A^{0^2}+B^{0^2}} & \frac{-12B^0}{A^{0^2}+B^{0^2}} \\ \frac{-6A^0B^0}{A^{0^2}+B^{0^2}} & \frac{2(4A^{0^2}+B^{0^2})}{A^{0^2}+B^{0^2}} & \frac{12A^0}{A^{0^2}+B^{0^2}} \\ \frac{-12B^0}{A^{0^2}+B^{0^2}} & \frac{12A^0}{A^{0^2}+B^{0^2}} & \frac{24}{A^{0^2}+B^{0^2}} \end{pmatrix}$$

and

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

Proof See section B.1.

Note that to estimate the frequency and frequency rate parameters, we need to solve a 2D nonlinear optimization problem. Even for a special case of this model, when $C^0 = D^0 = 0$, it has been observed that the least squares surface is highly nonlinear and has several local minima near the true parameter value (for details, see Rice and Rosenblatt [36]). Therefore, it is evident that computation of the LSEs is a numerically challenging problem for the proposed model as well.

It is important to note that under stronger assumptions of *i.i.d.* Gaussian distribution on the error random variables $X(t)$, the asymptotic variances of the LSEs coincide with the corresponding CRLBs.

2.2 Sequential Least Squares Estimators

To overcome the computational difficulty of finding the LSEs without compromising on the efficiency of the estimates, we propose a sequential procedure to find the estimates of the unknown parameters of model (5). In this section, we present the algorithm to obtain the sequential estimators and study the asymptotic properties of these estimators.

Note that the matrix $\boldsymbol{Z}(\alpha,\beta)$ can be partitioned into two $n\times 2$ blocks as follows:

$$\boldsymbol{Z}(\alpha,\beta) = \left(\begin{array}{c|c} \boldsymbol{Z}^{(1)}(\alpha) & \boldsymbol{Z}^{(2)}(\beta) \end{array}\right).$$

Here

$$\boldsymbol{Z}^{(1)}(\alpha)_{n\times 2} = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ \vdots & \vdots \\ \cos(n\alpha) & \sin(n\alpha) \end{pmatrix} \text{ and } \boldsymbol{Z}^{(2)}(\beta)_{n\times 2} = \begin{pmatrix} \cos(\beta) & \sin(\beta) \\ \vdots & \vdots \\ \cos(n^2\beta) & \sin(n^2\beta) \end{pmatrix}.$$

Similarly, the linear parameter vector can be written as:

$$\boldsymbol{\mu} = \left(\begin{array}{c|c} \boldsymbol{\mu}^{(1)^\top} & \boldsymbol{\mu}^{(2)^\top} \end{array}\right)^\top,$$

where $\boldsymbol{\mu}^{(1)}_{2\times 1} = \begin{pmatrix} A, & B \end{pmatrix}^\top$ and $\boldsymbol{\mu}^{(2)}_{2\times 1} = \begin{pmatrix} C, & D \end{pmatrix}^\top$. Also, the parameter vector,

$$\boldsymbol{\theta} = \left(\begin{array}{c|c} \boldsymbol{\theta}^{(1)} & \boldsymbol{\theta}^{(2)} \end{array}\right),$$

with $\boldsymbol{\theta}^{(1)} = \begin{pmatrix} A, & B, & \alpha \end{pmatrix}$ and $\boldsymbol{\theta}^{(2)} = \begin{pmatrix} C, & D, & \beta \end{pmatrix}$. The parameter space can be written as $\boldsymbol{\Theta}^{(1)} \times \boldsymbol{\Theta}^{(2)}$ so that $\boldsymbol{\theta}^{(1)} \in \boldsymbol{\Theta}^{(1)}$ and $\boldsymbol{\theta}^{(2)} \in \boldsymbol{\Theta}^{(2)}$, with $\boldsymbol{\Theta}^{(1)} = \boldsymbol{\Theta}^{(2)} = [-M, M] \times [-M, M] \times [0, \pi]$.

Following is the algorithm to find the sequential estimators:

Step 1: First minimise the following error sum of squares:

$$Q_1(\boldsymbol{\theta}^{(1)}) = (\boldsymbol{Y} - \boldsymbol{Z}^{(1)}(\alpha)\boldsymbol{\mu}^{(1)})^\top(\boldsymbol{Y} - \boldsymbol{Z}^{(1)}(\alpha)\boldsymbol{\mu}^{(1)}) \tag{10}$$

with respect to A, B and α. Using separable linear regression technique, for fixed α, we have:

$$\widetilde{\boldsymbol{\mu}}^{(1)}(\alpha) = [\boldsymbol{Z}^{(1)}(\alpha)^\top \boldsymbol{Z}^{(1)}(\alpha)]^{-1}\boldsymbol{Z}^{(1)}(\alpha)^\top \boldsymbol{Y}. \tag{11}$$

Now replacing $\boldsymbol{\mu}^{(1)}$ by $\widetilde{\boldsymbol{\mu}}^{(1)}(\alpha)$ in (10), we have:

$$R_1(\alpha) = \boldsymbol{Y}^\top(\boldsymbol{I} - \boldsymbol{Z}^{(1)}(\alpha)[\boldsymbol{Z}^{(1)}(\alpha)^\top \boldsymbol{Z}^{(1)}(\alpha)]^{-1}\boldsymbol{Z}^{(1)}(\alpha)^\top)\boldsymbol{Y}.$$

Minimising $R_1(\alpha)$, we obtain $\widetilde{\alpha}$ and replacing α by $\widetilde{\alpha}$ in (11), we get the linear parameter estimates $\widetilde{A}$ and $\widetilde{B}$.

Step 2: At this step, we eliminate the effect of the sinusoidal component from the original data, and obtain a new data vector:

$$\boldsymbol{Y}_1 = \boldsymbol{Y} - \boldsymbol{Z}^{(1)}(\widetilde{\alpha})\widetilde{\boldsymbol{\mu}}^{(1)}.$$

Now we minimise the error sum of squares:

$$Q_2(\boldsymbol{\theta}^{(2)}) = (\boldsymbol{Y}_1 - \boldsymbol{Z}^{(2)}(\beta)\boldsymbol{\mu}^{(2)})^\top(\boldsymbol{Y}_1 - \boldsymbol{Z}^{(2)}(\beta)\boldsymbol{\mu}^{(2)}), \tag{12}$$

with respect to C, D and β. Again by separable linear regression technique, we have:

$$\widetilde{\boldsymbol{\mu}}^{(2)}(\beta) = [\boldsymbol{Z}^{(2)}(\beta)^\top \boldsymbol{Z}^{(2)}(\beta)]^{-1}\boldsymbol{Z}^{(2)}(\beta)^\top \boldsymbol{Y}_1 \tag{13}$$

for a fixed β. Now replacing $\boldsymbol{\mu}^{(2)}$ by $\widetilde{\boldsymbol{\mu}}^{(2)}$ in (12), we obtain:

$$R_2(\beta) = \boldsymbol{Y}_1^\top(\boldsymbol{I} - \boldsymbol{Z}^{(2)}(\beta)[\boldsymbol{Z}^{(2)}(\beta)^\top \boldsymbol{Z}^{(2)}(\beta)]^{-1}\boldsymbol{Z}^{(2)}(\beta)^\top)\boldsymbol{Y}_1.$$

Minimizing $R_2(\beta)$, with respect to β, we obtain $\widetilde{\beta}$, and using $\widetilde{\beta}$ in (13), we obtain $\widetilde{C}$ and $\widetilde{D}$, the linear parameter estimates.

We use the following notations: ${\boldsymbol{\theta}^{(1)}}^0 = (A^0,\ B^0,\ \alpha^0)$ and ${\boldsymbol{\theta}^{(2)}}^0 = (C^0,\ D^0,\ \beta^0)$ are the true parameter vectors, $\widetilde{\boldsymbol{\theta}}^{(1)} = (\widetilde{A},\ \widetilde{B},\ \widetilde{\alpha})$ is the sequential LSE of ${\boldsymbol{\theta}^{(1)}}^0$ and $\widetilde{\boldsymbol{\theta}}^{(2)} = (\widetilde{C},\ \widetilde{D},\ \widetilde{\beta})$ that of ${\boldsymbol{\theta}^{(2)}}^0$.

In the following theorems, we prove that the proposed sequential LSEs are strongly consistent as the usual LSEs. Moreover, if Conjecture 1 (see section A) holds true, the sequential LSEs have the same asymptotic distribution as the corresponding LSEs.

Theorem 3 *Under Assumptions 1 and 2, $\widetilde{\boldsymbol{\theta}}^{(1)}$ and $\widetilde{\boldsymbol{\theta}}^{(2)}$ are strongly consistent estimators of ${\boldsymbol{\theta}^{(1)}}^0$ and ${\boldsymbol{\theta}^{(2)}}^0$, respectively, that is,*

(a)

$$\widetilde{\boldsymbol{\theta}}^{(1)} \xrightarrow{a.s.} {\boldsymbol{\theta}^{(1)}}^0 \text{ as } n \to \infty,$$

(b)

$$\widetilde{\boldsymbol{\theta}}^{(2)} \xrightarrow{a.s.} {\boldsymbol{\theta}^{(2)}}^0 \text{ as } n \to \infty.$$

Proof See section B.2.

Theorem 4 *Under Assumptions 1 and 2 and presuming Conjecture 1 (see section A) holds true,*

(a)

$$(\widetilde{\boldsymbol{\theta}}^{(1)} - {\boldsymbol{\theta}^{(1)}}^0)\boldsymbol{D}_1^{-1} \xrightarrow{d} \mathcal{N}_3(0, c\sigma^2 {\boldsymbol{\Sigma}^{(1)}}^{-1}) \text{ as } n \to \infty,$$

(b)

$$(\widetilde{\boldsymbol{\theta}}^{(2)} - {\boldsymbol{\theta}^{(2)}}^0)\boldsymbol{D}_2^{-1} \xrightarrow{d} \mathcal{N}_3(0, c\sigma^2 {\boldsymbol{\Sigma}^{(2)}}^{-1}) \text{ as } n \to \infty,$$

where $\boldsymbol{D}_1$ and $\boldsymbol{D}_2$, are sub-matrices of order 3×3, of the diagonal matrix $\boldsymbol{D}$ such that $\boldsymbol{D} = \left(\begin{array}{c|c} \boldsymbol{D}_1 & \boldsymbol{0} \\ \hline \boldsymbol{0} & \boldsymbol{D}_2 \end{array}\right)$. Note that, $\boldsymbol{D}$, c and $\boldsymbol{\Sigma}_1^{-1}(\boldsymbol{\theta}^0)$ and $\boldsymbol{\Sigma}_2^{-1}(\boldsymbol{\theta}^0)$ are as defined in Theorem 2.

Proof See section B.2.

3 Multiple Component Chirp-like Model

To model real life data, we require a more adaptable model. In this section, we consider a multiple-component chirp-like model (3), a natural generalization of the one-component model. Under certain assumptions in addition to Assumption 1 on the noise component, that we state below, we study the asymptotic properties of the LSEs and provide the results in the following subsection.

Let us denote by $\boldsymbol{\vartheta}$ the parameter vector for model (3),

$$\boldsymbol{\vartheta} = (A_1, B_1, \alpha_1, \cdots, A_p, B_p, \alpha_p, C_1, D_1, \beta_1, \cdots, C_q, D_q, \beta_q).$$

Also, let $\boldsymbol{\vartheta}^0$ denote the true parameter vector and $\widehat{\boldsymbol{\vartheta}}$, the LSE of $\boldsymbol{\vartheta}^0$.

Assumption 3 *$\boldsymbol{\vartheta}^0$ is an interior point of $\boldsymbol{\mathcal{V}} = {\boldsymbol{\Theta}_1}^{(p+q)}$, the parameter space and the frequencies $\alpha_j^0 s$ are distinct for $j = 1, \cdots p$ and so are the frequency rates $\beta_k^0 s$ for $k = 1, \cdots q$. Note that $\boldsymbol{\Theta}_1 = [-M, M] \times [-M, M] \times [0, \pi]$.*

Assumption 4 *The amplitudes, $A_j^0 s$ and $B_j^0 s$ satisfy the following relationship:*

$$2M^2 > A_1^{0\,2} + B_1^{0\,2} > A_2^{0\,2} + B_2^{0\,2} > \cdots > A_p^{0\,2} + B_p^{0\,2} > 0.$$

Similarly, $C_k^0 s$ and $D_k^0 s$ satisfy the following relationship:

$$2M^2 > C_1^{0\,2} + D_1^{0\,2} > C_2^{0\,2} + D_2^{0\,2} > \cdots > C_q^{0\,2} + D_q^{0\,2} > 0.$$

3.1 Least Squares Estimators

The LSEs of the unknown parameters of the proposed model, see (3), can be obtained by minimising the error sum of squares:

$$Q(\boldsymbol{\vartheta}) = \sum_{t=1}^{n} \left(y(t) - \sum_{j=1}^{p} \{A_j cos(\alpha_j t) + B_j sin(\alpha_j t)\} - \sum_{k=1}^{q} \{C_k cos(\beta_k t^2) + D_k sin(\beta_k t^2)\} \right)^2 \tag{14}$$

with respect to A_1, B_1, α_1, $\cdots$, A_p, B_p α_p, C_1, D_1, β_1, $\cdots$, C_q, D_q and β_q simultaneously. Similar to the one-component model, $Q(\boldsymbol{\vartheta})$ can be expressed in matrix notation and then the LSE, $\widehat{\boldsymbol{\vartheta}}$ of $\boldsymbol{\vartheta}^0$, can be obtained along the similar lines.

Next, we examine the consistency property of the LSE $\widehat{\boldsymbol{\vartheta}}$ along with its asymptotic distribution.

Theorem 5 *If Assumptions, 1, 3 and 4, hold true, then:*

$$\widehat{\boldsymbol{\vartheta}} \xrightarrow{a.s.} \boldsymbol{\vartheta}^0 \text{ as } n \to \infty.$$

Proof The consistency of the LSE $\widehat{\boldsymbol{\vartheta}}$ can be proved along the similar lines as the consistency of the LSE $\widehat{\boldsymbol{\theta}}$, for the one-component model.

Theorem 6 *If the above Assumptions, 1, 3 and 4, then:*

$$(\widehat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}^0)\mathfrak{D}^{-1} \xrightarrow{d} \mathcal{N}_{3(p+q)}(0, c\sigma^2 \mathcal{E}^{-1}(\boldsymbol{\vartheta}^0)) \text{ as } n \to \infty.$$

Here, $\mathfrak{D} = diag(\underbrace{\boldsymbol{D}_1, \cdots \boldsymbol{D}_1}_{p \text{ times}}, \underbrace{\boldsymbol{D}_2, \cdots, \boldsymbol{D}_2}_{q \text{ times}})$, *where* $\boldsymbol{D}_1 = diag(\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}, \frac{1}{n\sqrt{n}})$ *and* $\boldsymbol{D}_2 = diag(\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}, \frac{1}{n^2\sqrt{n}})$.

$$\mathcal{E}(\boldsymbol{\vartheta}^0) = \begin{pmatrix} \boldsymbol{\Sigma}_1^{(1)} & 0 & \cdots & & \cdots & 0 \\ 0 & \ddots & 0 & & \cdots & 0 \\ \vdots & \vdots & \boldsymbol{\Sigma}_p^{(1)} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \boldsymbol{\Sigma}_1^{(2)} & 0 & 0 \\ 0 & \cdots & \cdots & 0 & \ddots & 0 \\ 0 & \cdots & & \cdots & 0 & \boldsymbol{\Sigma}_q^{(2)} \end{pmatrix},$$

with $\boldsymbol{\Sigma}_j^{(1)} = \begin{pmatrix} \frac{1}{2} & 0 & \frac{B_j^0}{4} \\ 0 & \frac{1}{2} & \frac{-A_j^0}{4} \\ \frac{B_j^0}{4} & \frac{-A_j^0}{4} & \frac{A_j^{0\,2}+B_j^{0\,2}}{6} \end{pmatrix}$, $j = 1, \cdots, p$ *and* $\boldsymbol{\Sigma}_k^{(2)} = \begin{pmatrix} \frac{1}{2} & 0 & \frac{D_k^0}{6} \\ 0 & \frac{1}{2} & \frac{-C_k^0}{6} \\ \frac{D_k^0}{6} & \frac{-C_k^0}{6} & \frac{C_k^{0\,2}+D_k^{0\,2}}{10} \end{pmatrix}$, $k = 1, \cdots, q$.

Proof See section C.1.

3.2 Sequential Least Squares Estimators

For the multiple-component chirp-like model, if the number of components, p and q are very large, finding the LSEs becomes computationally challenging. To resolve this issue, we propose a sequential procedure to estimate the unknown parameters similar to the one-component model. Using the sequential procedure, the $(p+q)$-dimensional optimization problem can be reduced to $p+q$, 1D optimization problems. The algorithm for the sequential estimation is as follows:

Step 1: Perform **Step 1** of the sequential algorithm for the one-component chirp-like model as explained in Section 2.2 and obtain the estimate, $\widetilde{\boldsymbol{\theta}}_1^{(1)} = (\widetilde{A}_1, \widetilde{B}_1, \widetilde{\alpha}_1)$.

Step 2: Eliminate the effect of the estimated sinusoidal component and obtain new data vector:

$$y_1(t) = y(t) - \widetilde{A}_1 \cos(\widetilde{\alpha}_1 t) - \widetilde{B}_1 \sin(\widetilde{\alpha}_1 t).$$

Step 3: Minimize the following error sum of squares to obtain the estimates of the next sinusoid, $\widetilde{\boldsymbol{\theta}}_2^{(1)} = (\widetilde{A}_2, \widetilde{B}_2, \widetilde{\alpha}_2)$:

$$Q_2(A, B, \alpha) = \sum_{t=1}^{n} \Big(y_1(t) - A\cos(\alpha t) - B\sin(\alpha t) \Big)^2.$$

Repeat these steps until all the p sinusoids are estimated.

Step p + 1: At $(p+1)$-*th* step, we obtain the data:

$$y_p(t) = y_{p-1}(t) - \widetilde{A}_p \cos(\widetilde{\alpha}_p t) - \widetilde{B}_p \sin(\widetilde{\alpha}_p t).$$

Step p + 2: Using this data we estimate the first chirp component parameters, and obtain $\widetilde{\boldsymbol{\theta}}_1^{(2)} = (\widetilde{C}_1, \widetilde{D}_1, \widetilde{\beta}_1)$: by minimizing:

$$Q_{p+1}(C, D, \beta) = \sum_{t=1}^{n} \Big(y_p(t) - C\cos(\beta t^2) - D\sin(\beta t^2) \Big)^2.$$

Step p + 3: Now eliminate the effect of this estimated chirp component and obtain: $y_{p+1}(t) = y_p(t) - \widetilde{C}_1 \cos(\widetilde{\beta}_1 t^2) - \widetilde{D}_1 \sin(\widetilde{\beta}_1 t^2)$ and minimize $Q_{p+2}(C, D, \beta)$ to obtain $\widetilde{\boldsymbol{\theta}}_2^{(2)} = (\widetilde{C}_2, \widetilde{D}_2, \widetilde{\beta}_2)$.

Continue to do so and estimate all the q chirp components.

We now investigate the consistency property of the proposed sequential estimators, when p and q are unknown. Thus, we consider the following two cases: (a) when the number of components of the fitted model is less than the actual number of components, and (b) when the number of components of the fitted model is more than the actual number of components.

Theorem 7 *If Assumptions 1, 3 and 4 are satisfied, then the following are true:*

(a)

$$\widetilde{\boldsymbol{\theta}}_1^{(1)} \xrightarrow{a.s.} {\boldsymbol{\theta}_1^{(1)}}^0 \text{ as } n \to \infty,$$

(b)

$$\widetilde{\boldsymbol{\theta}}_1^{(2)} \xrightarrow{a.s.} {\boldsymbol{\theta}_1^{(2)}}^0 \text{ as } n \to \infty.$$

Proof See section C.2.

Theorem 8 *If Assumptions 1, 3 and 4 are satisfied, the following are true:*

(a)

$$\widetilde{\boldsymbol{\theta}}_j^{(1)} \xrightarrow{a.s.} {\boldsymbol{\theta}_j^{(1)}}^0 \text{ as } n \to \infty, \text{ for all } j = 2, \cdots, p,$$

(b)

$$\widetilde{\boldsymbol{\theta}}_k^{(2)} \xrightarrow{a.s.} {\boldsymbol{\theta}_k^{(2)}}^0 \text{ as } n \to \infty, \text{ for all } k = 2, \cdots, q.$$

Proof See section C.2.

Theorem 9 *If Assumptions 1, 3 and 4 are true, then the following are true:*

(a)

$$\widetilde{A}_{p+k} \xrightarrow{a.s.} 0, \quad \widetilde{B}_{p+k} \xrightarrow{a.s.} 0 \text{ for } k = 1, 2, \ldots, \text{ as } n \to \infty,$$

(b)

$$\widetilde{C}_{q+k} \xrightarrow{a.s.} 0, \quad \widetilde{D}_{q+k} \xrightarrow{a.s.} 0 \text{ for } k = 1, 2, \ldots, \text{ as } n \to \infty.$$

Proof See section C.2.

Next we determine the asymptotic distribution of the proposed estimators at each step through the following theorems:

Theorem 10 *If Assumptions 1, 3 and 4 are satisfied and presuming Conjecture 1 (see section A) hold true, then:*

(a)

$$(\widetilde{\boldsymbol{\theta}}_1^{(1)} - \boldsymbol{\theta}_1^{0(1)}) \boldsymbol{D}_1^{-1} \xrightarrow{d} \mathcal{N}_3(0, c\sigma^2 {\boldsymbol{\Sigma}_1^{(1)}}^{-1}) \text{ as } n \to \infty,$$

(b)

$$(\widetilde{\boldsymbol{\theta}}_1^{(2)} - \boldsymbol{\theta}_1^{0(2)})\boldsymbol{D}_2^{-1} \xrightarrow{d} \mathcal{N}_3(0, c\sigma^2 {\boldsymbol{\Sigma}_1^{(2)}}^{-1}) \text{ as } n \to \infty.$$

Here, c, the diagonal matrices $\boldsymbol{D}_1$ *and* $\boldsymbol{D}_2$ *and the matrices* ${\boldsymbol{\Sigma}_1^{(2)}}^{-1}$ *and* ${\boldsymbol{\Sigma}_1^{(2)}}^{-1}$ *are as defined in Theorem 6.*

Proof See section C.2.

This result can be extended for $2 \leqslant j \leqslant p$ and $2 \leqslant k \leqslant q$ as follows:

Theorem 11 *If Assumptions 1, 3 and 4 are satisfied and presuming Conjecture 1 (see section A) hold true,, then for all* $j = 2, \ldots, p$ *and* $k = 2, \ldots, q$*:*

(a)

$$(\widetilde{\boldsymbol{\theta}}_j^{(1)} - \boldsymbol{\theta}_j^{0(1)})\boldsymbol{D}_1^{-1} \xrightarrow{d} \mathcal{N}_3(0, c\sigma^2 {\boldsymbol{\Sigma}_j^{(1)}}^{-1}) \text{ as } n \to \infty,$$

(b)

$$(\widetilde{\boldsymbol{\theta}}_k^{(2)} - \boldsymbol{\theta}_k^{0(2)})\boldsymbol{D}_2^{-1} \xrightarrow{d} \mathcal{N}_3(0, c\sigma^2 {\boldsymbol{\Sigma}_k^{(2)}}^{-1}) \text{ as } n \to \infty.$$

$\boldsymbol{\Sigma}_j^{(1)}$ *and* $\boldsymbol{\Sigma}_k^{(2)}$ *are as defined in Theorem 6.*

Proof This can be obtained along the same lines as the proof of Theorem 10.

From the above results, it is evident that the sequential LSEs are strongly consistent and have the same asymptotic distribution as the LSEs and at the same time can be computed more efficiently. Thus for the simulation studies and to analyze the real data sets as well, we compute the sequential LSEs instead of the LSEs.

4 Simulation Studies

In this section, we present the results obtained from some numerical experiments, performed both for a one-component and a multiple-component model. These results demonstrate the applicability of our model and the performance of the LSEs and the sequential LSEs. Since we are primarily interested in the estimation of the nonlinear parameters, here we report only these estimates. The linear parameter estimates can be obtained by simple linear regression.

4.1 Results for a one-component chirp-like model

In the first set of experiments, we consider a one-component chirp-like model (5) with the following true parameter values:

$$A^0 = 10, B^0 = 10, \alpha^0 = 1.5, C^0 = 10, D^0 = 10 \text{ and } \beta^0 = 0.1$$

The error structure used to generate the data is as follows:

$$X(t) = \epsilon(t) + 0.5\epsilon(t-1).$$

Here $\epsilon(t)$s are *i.i.d.* normal random variables with mean zero and variance σ^2. We consider different sample sizes: $n = 100, 200, 300, 400$ and 500 and different error variances: σ^2: 0.1, 0.25, 0.5, 0.75 and 1. For each n and σ^2, we generate the data and obtain the LSEs. Based on 1000 iterations,we compute the biases and MSEs of the LSEs. We also compute the theoretical asymptotic variances of the proposed estimators to compare with the corresponding computed MSEs. Figures 2 and 3 represent the biases and MSEs of the LSEs of α and β compared to the asymptotic variances versus different sample sizes. Similarly in figures 4 and 5 the biases and MSEs versus signal-to-noise ratio (SNR) are shown.

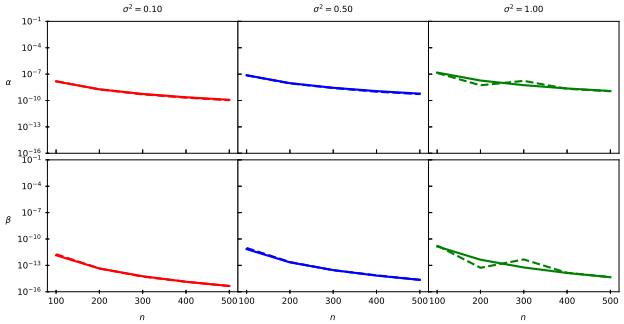

Fig. 3 In each sub-plot the dashed line represents the MSEs of the estimates and the solid line represents the corresponding theoretical asymptotic variances of the estimators of parameters of the underlying simulated one-component model versus the sample size.

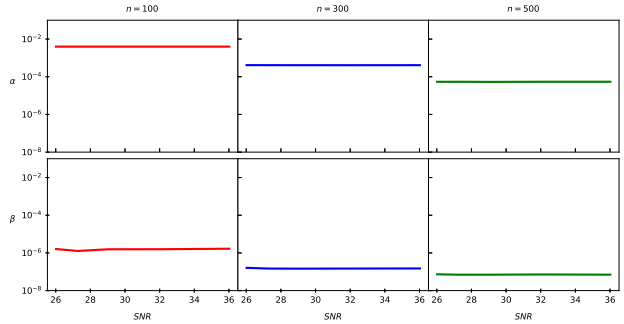

Fig. 4 In each sub-plot, the solid line represents the absolute value of the biases of the estimators of parameters of the underlying simulated one-component model versus SNR.

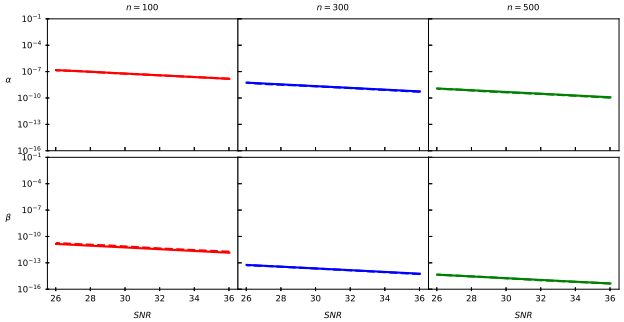

Fig. 5 In each sub-plot the dashed line represents the MSEs of the estimates and the solid line represents the corresponding theoretical asymptotic variances of the estimators of parameters of the underlying simulated one-component model versus SNR.

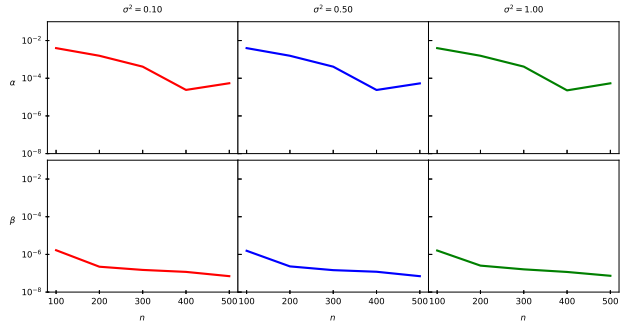

Fig. 2 In each sub-plot, the solid line represents the absolute value of the biases of the estimators of parameters of the underlying simulated one-component model versus the sample size.

Figures 2 and 4 show that the biases of the estimates are quite small and therefore the average estimates are close to the true values. Figure 3 depicts the consistent behaviour of the LSEs. It can be seen that as n increases,

the MSEs decrease. Similarly, Figure 5 represents the performance of the LSEs for different SNRs compared with the asymptotic variances. The figure shows that MSEs decrease as the SNR increases and they match the corresponding asymptotic variances quite well.

4.2 Results for a multiple-component chirp-like model

Here we present the simulation results for the multiple-component chirp-like model (3) with $p = q = 2$. Following are the true parameter values used for data generation:

$$A_1^0 = 10, B_1^0 = 10, \alpha_1^0 = 1.5, C_1^0 = 10, D_1^0 = 10 \text{ and } \beta_1^0 = 0.1,$$

$$A_2^0 = 8, B_2^0 = 8, \alpha_2^0 = 2.5, C_2^0 = 8, D_2^0 = 8 \text{ and } \beta_2^0 = 0.2.$$

The error structure used for data generation is again a moving average process, the same as used for the one-component model simulations. We compute the sequential LSEs of the parameters and report their biases, MSEs, and asymptotic variances. Again, the different sample sizes and error variances used for the simulations are the same as those for the one-component model. Figures 6 and 7 display the biases and the MSEs of the obtained estimates versus the varying sample size. It is observed that the estimates obtained have significantly small biases and thereby are close to the true values. We also observe that as n increases the biases and the MSEs decrease, thus depicting the desired consistency of the estimators. Moreover, the MSEs are on an equal footing with the corresponding asymptotic variances. The observations, therefore, validate the derived theoretical properties of the sequential estimators.

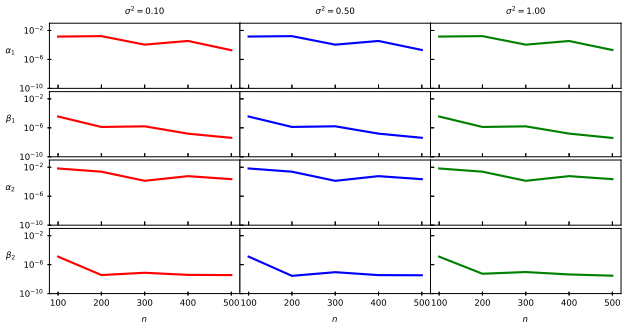

Fig. 6 In each sub-plot, the solid line represents the absolute value of the biases of the estimators of parameters of the underlying simulated two component model versus sample size.

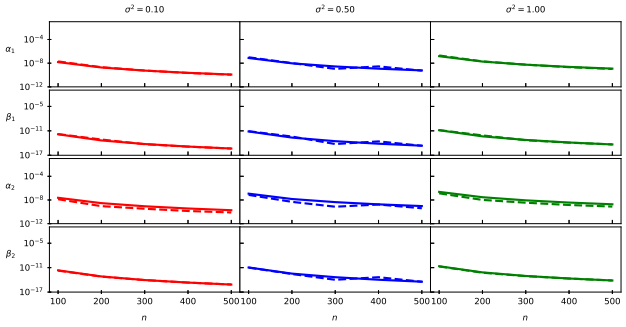

Fig. 7 In each sub-plot, the dashed line represents the MSEs of the estimates and the solid line represents the corresponding theoretical asymptotic variances of the estimators of parameters of the underlying simulated two component model versus sample size.

In figures 8 and 9 the biases and the MSEs of the sequential estimates of the first component parameters versus the SNR are shown and in figures 10 and 11 the biases and the MSEs of the sequential estimates of the second component parameters versus SNR are displayed. The lines representing the MSEs and the asymptotic variances in figures 9 and 11 are visually indistinguishable, indicating high accuracy of the estimators.

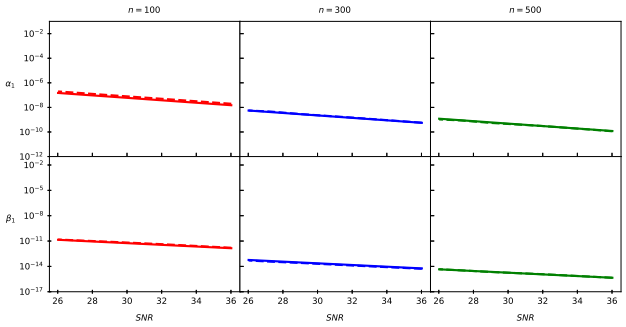

Fig. 9 In each sub-plot, the dashed line represents the MSEs of the estimates and the solid line represents the corresponding theoretical asymptotic variances of the estimators of parameters of the first component of the simulated two component model versus SNR.

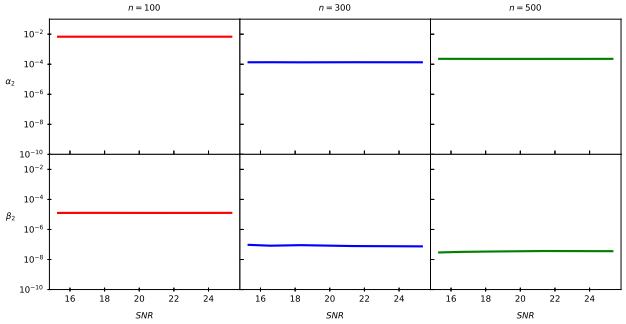

Fig. 10 In each sub-plot, the solid line represents the absolute value of the biases of the estimators of parameters of the second component of the simulated two component model versus SNR.

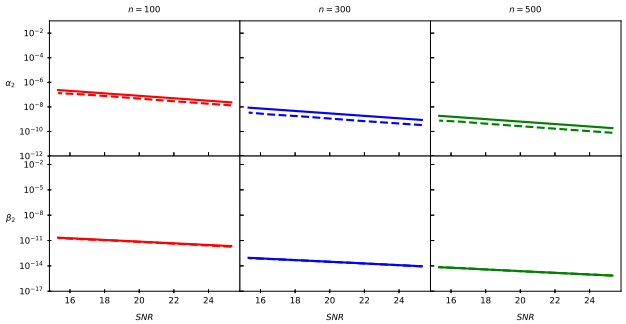

Fig. 11 In each sub-plot, the dashed line represents the MSEs of the estimates and the solid line represents the corresponding theoretical asymptotic variances of the estimators of parameters of the second component of the simulated two component model versus SNR.

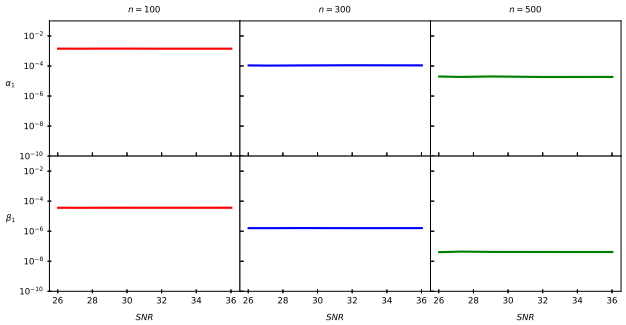

Fig. 8 In each sub-plot, the solid line represents the absolute value of the biases of the estimators of parameters of the first component of the simulated two component model versus SNR.

4.3 Fitting a chirp model versus a chirp-like model to a given data

In this section, we make a comparison of computational complexities involved in fitting a chirp-like model to a data set and that involved in modeling a data set using a chirp model. The method of estimation that we will use to fit either of these models is sequential LSEs as it significantly reduces the computational burden involved in finding the traditional LSEs. This is discussed in more detail in the next section.

It must be noted that for fitting a non-linear model finding the initial values is of prime importance. Once we have found the initial values, we can employ any iterative algorithm to find the sequential LSEs of the parameters. To find precise initial values, we have to resort to a fine grid search throughout the parameter space, but performing a grid search entails high computational load. We demonstrate in the following figures how replacing a chirp model with a chirp-like model reduces this computational load significantly. In Figure 12, we plot the size of the grid required to find accurate initial values of the parameters of one component of a chirp model and a chirp-like model. It is visually evident that the difference in the computational complexity involved in fitting a chirp model and a chirp-like model is huge. In Figure 13, time taken to fit a component of a chirp model and a chirp-like model is shown and the picture gives more insight into the computational difference.

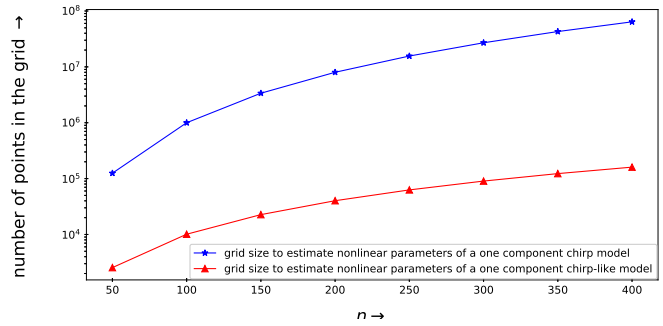

Fig. 12 Comparison of computational complexity involved in fitting a component of chirp model and of a chirp-like model.

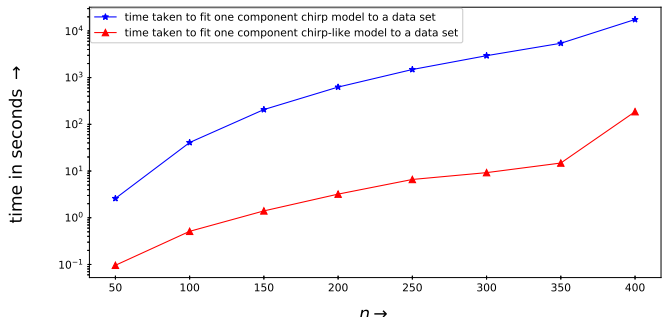

Fig. 13 Comparison of time consumption of fitting a component of chirp model and of a chirp-like model.

4.4 Fitting a chirp-like model using LSEs versus sequential LSEs.

In this section, we see how it is more expensive from a computational point of view to find the LSEs. As discussed before, finding the initial values accounts for most of the time consumption of finding the estimators of the non-linear parameters. Therefore the computational complexity of finding the LSEs and the sequential LSEs heavily depends on the grid search for initial values. In Figure 14, we bring out this comparison between the LSEs and the sequential LSEs. The number of grid points in the parameter space required for precise initial values of the parameter estimates of a chirp-like model with two sinusoids and two chirp components is reported. The figure reveals that the sequential method reduces the computational burden involved in finding the LSEs significantly.

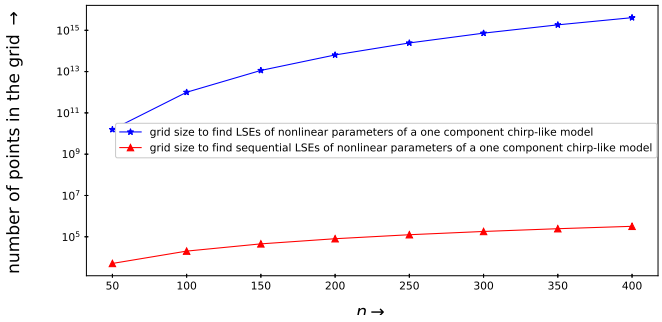

Fig. 14 Comparison of computational complexity involved in finding the LSEs and sequential LSEs of a chirp-like model with two sinusoids and two chirp components.

5 Data Analysis

5.1 Real Data Analysis

In this section, we analyze four different speech signal data sets: "AAA", "AHH", "UUU" and "EEE" using the chirp model as well as the proposed chirp-like model. These data sets have been obtained from a sound instrument at the Speech Signal Processing laboratory of the Indian Institute of Technology Kanpur. The data set "AAA", has 477 data points, the set "AHH" has 469 data points and the rest of them have 512 points each.

We fit the chirp-like model to these data sets using the sequential LSEs following the algorithm described in Section 3.2. As is evident from the description, we need to solve a 1D optimization problem to find these estimators and since the problem is nonlinear, we need to employ some iterative method to do so. Here we use Brent's method [3] to solve 1D optimization problems, using an inbuilt function in R, known as 'optim'. For this method to work, we require very good initial values in the sense that they need to be close to the true values. Now one of the well-received methods for finding initial values for the frequencies of the sinusoidal model is to maximize the periodogram function:

$$I_1(\alpha) = \frac{1}{n}\left|\sum_{t=1}^{n} y(t)e^{-i\alpha t}\right|^2$$

at the points: $\frac{\pi j}{n}$; $j = 1, \cdots, n-1$, called the Fourier frequencies. The estimators obtained by this method, are called the Periodogram Estimators. After all the p sinusoidal components are fitted, we need to fit the q chirp components. Again we need to solve 1D optimization problem at each stage and for that we need good initial values. Analogous to the periodogram function $I_1(\alpha)$, we define a periodogram-type function as follows:

$$I_2(\beta) = \frac{1}{n}\left|\sum_{t=1}^{n} y(t)e^{-i\beta t^2}\right|^2.$$

To obtain the starting points for the frequency rate parameter β, we maximise this function at the points: $\frac{\pi k}{n^2}$; $k = 1, \cdots, n^2 - 1$, similar to the Fourier frequencies.

Since in practice, the number of components of a model are unknown, we need to estimate them. We use the following Bayesian Information Criterion (BIC) criterion, as a tool to estimate p and q:

$$\text{BIC}(j,k) = n\ ln(\text{SS}_{res}(j,k)) + 2\ (3j + 3k)\ ln(n); j = 1, \ldots, J;\ k = 1, \ldots, K, \tag{15}$$

for the present analysis of the data sets. For reference on the form of this criterion function, one may refer to the monograph by Kundu and Nandi [19]. Here, SS_{res} is the residual sum of squares when j sinusoidal components and k chirp components are fitted to the data. This is based on the assumption that the maximum number of sinusoidal components is J and chirp components is K and in practice, we choose a large values of J and K. We select the pair (j,k) as an estimate of the pair (p,q) corresponding to the minimum BIC.

For comparison of the chirp-like model with the chirp model, we re-analyze these data sets by fitting a chirp model to each of them (for methodology, see Lahiri, Kundu and Mitra [22]). In the following table, we report the number of components required to fit the chirp model and the chirp-like model to each of the data sets and in the subsequent figures, we plot the original data along with the estimated signals obtained by fitting a

Table 1 Number of components used to fit chirp and chirp-like model to the speech data sets.

Number of components				
Data Set	Chirp Model		Chirp-like model	
	p	Number of parameters	(p, q)	Number of parameters
AAA	9	36	(10, 1)	33
AHH	8	32	(7, 1)	24
UUU	9	36	(8, 1)	27
EEE	11	44	(14, 1)	45

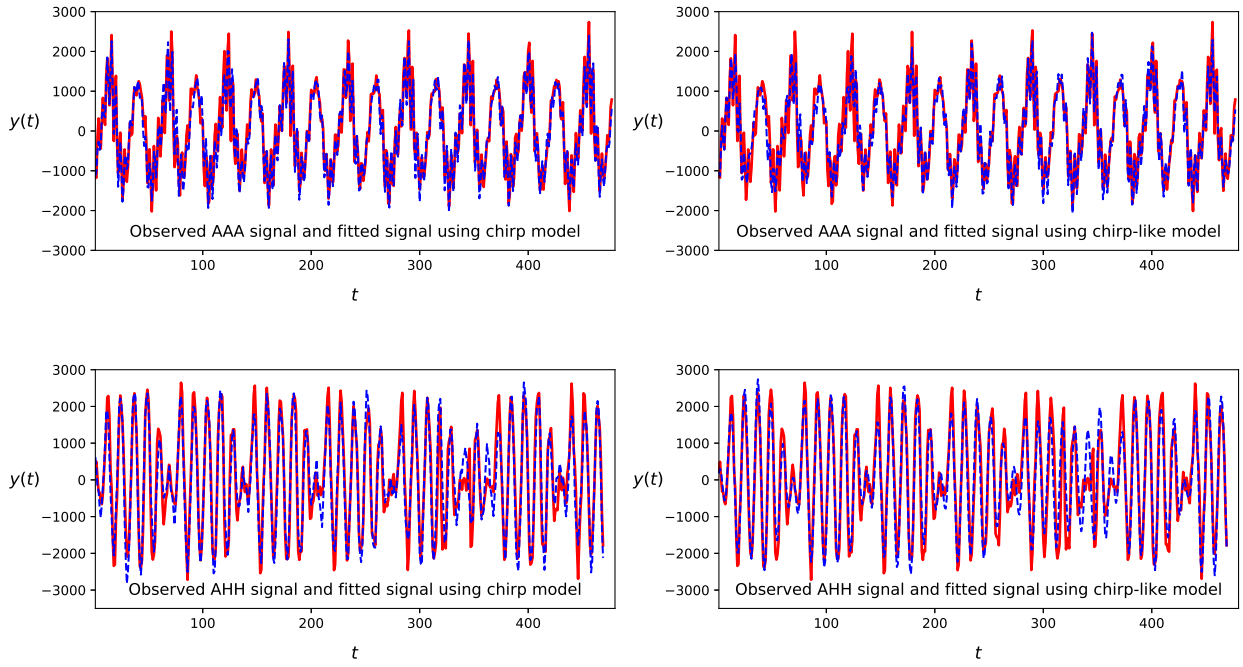

Fig. 15 Speech Signal data sets: "AAA" and "AHH"; Observed data (red solid line) and fitted signal (blue dashed line). The sub-plots on the left represent chirp model fitting and those on the right represent chirp-like model fitting.

chirp model and a chirp-like model to these data. In both scenarios, the model is fitted using the sequential LSEs.

To validate the error assumption of stationarity, we test the residuals, for all the cases, using the augmented Dickey-Fuller test (for more details see Fuller [7]). This tests the following null hypothesis:
H_0: There is a unit root present in the series,
against the following alternative:
H_1 : No unit root present in the series, that is, the series is stationary.
We use an inbuilt function 'adftest' in MATLAB for this purpose. The test statistic values result in rejection of the null hypothesis of a presence of a unit root indicating that residuals, in all the cases, are stationary.

It is evident from the figures above, that visually both the models provide a good fit for all the speech data sets. However, to fit a chirp-like model using the sequential LSEs, we solve a 1D optimization problem at each step while for the fitting of a chirp model, at each step we need to deal with a 2D optimization problem. Moreover, to find the initial values, in both cases, a grid search is performed and for the chirp-like model, this means evaluation of the periodogram functions $I_1(\alpha)$ and $I_2(\beta)$ at n and n^2 grid points respectively as opposed to the n^3 grid points for the chirp model. Note that this is done at each step for the sequential estimators and hence becomes more complex as the number of components increases. Thus fitting a chirp-like model is computationally much more efficient than fitting a chirp model.

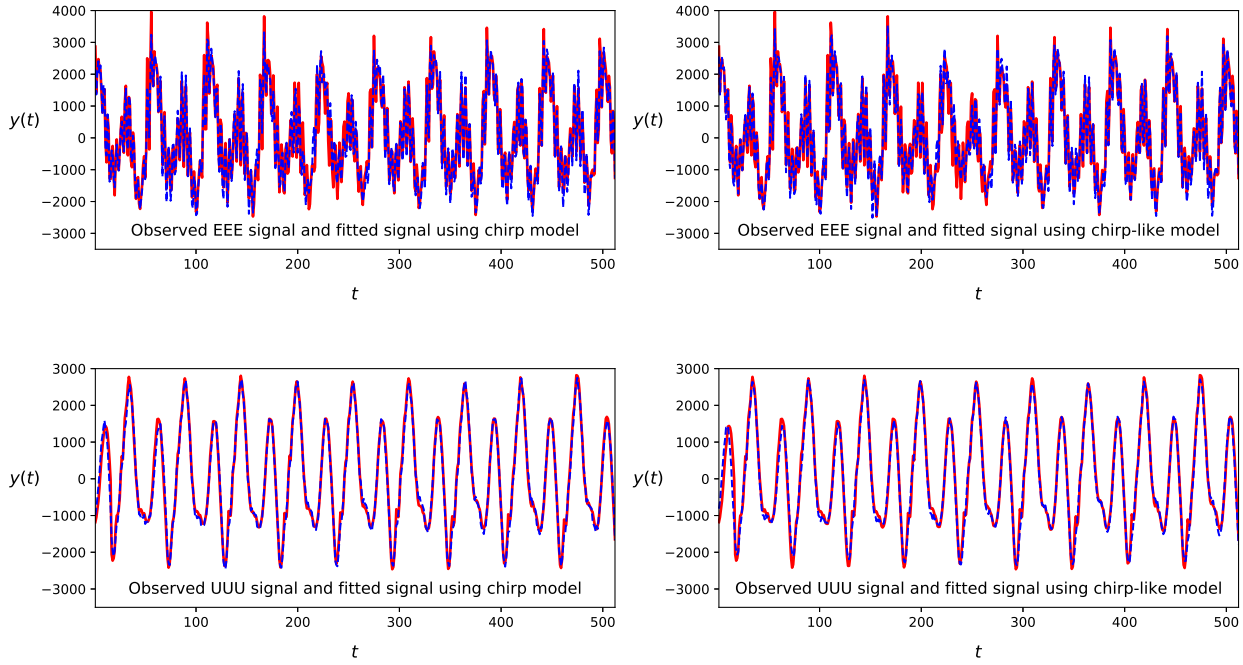

Fig. 16 Speech Signal data sets: "EEE" and "UUU"; Observed data (red solid line) and fitted signal (blue dashed line). The sub-plots on the left represent chirp model fitting and those on the right represent chirp-like model fitting.

Table 2 True parameters values of the synthetic data.

A_k^0	951.12877	410.28077	-205.56061	-205.77076	160.16585
B_k^0	942.76161	-325.20751	-412.32128	-74.00592	-67.47605
α_k^0	2.6838036	2.7794191	0.5512842	0.2698761	2.9871480
β_k^0	3.141577	3.141573	1.531575e-05	1.814701e-05	3.140960

5.2 Simulated Data Analysis

We generate the data from a multiple-component chirp model. The number of components is set to 5 and the parameters, amplitudes, frequencies, and frequency rates are assigned prefixed values provided in Table 2.

The data here is generated with the following error structure:

$$X(t) = 0.8897X(t-1) - 0.4858X(t-2) + e(t) - 0.2279e(t-1) + 0.2488e(t-2)$$

Here $e(t)$ are i.i.d. Gaussian random variables with mean 0 and variance 2. The simulated signal consists of 512 sample points and is shown in Figure 17.

The objective is to evaluate and compare the performance of the chirp model and the chirp-like model to fit the simulated data. First, we fit a chirp model to the data using the sequential least squares estimation method. For estimating the number of chirp components, we use the following form of BIC:

$$\text{BIC}(j) = n\ ln(\text{SS}_{res}(j)) + 2\ *4j\ *ln(n); j = 1, \ldots, J. \tag{16}$$

Corresponding to the minimum value of BIC, $\widetilde{p} = 5$. Using these 5 chirp components and sequential LSEs of the parameters we compute the estimated signal. The fitted chirp signal overlapping the synthesized signal is shown in Figure 18.

Next, to illustrate the effectiveness of a chirp-like model to clone a chirp signal, we fit a chirp-like model to the above synthesized data. We fit this model using the proposed sequential LSEs. Since the number of sinusoids and chirp components needed to fit this model to the simulated data is unknown, we again use BIC for the model selection as defined in (15). Corresponding to the minimum value of BIC, we choose $\widetilde{p} = 9$ and $\widetilde{q} = 1$, the estimates of the model order. In Figure 19, the model fitting corresponding to the selected model is shown

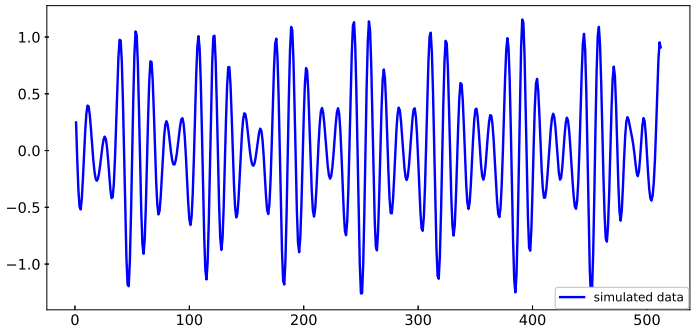

Fig. 17 Simulated data.

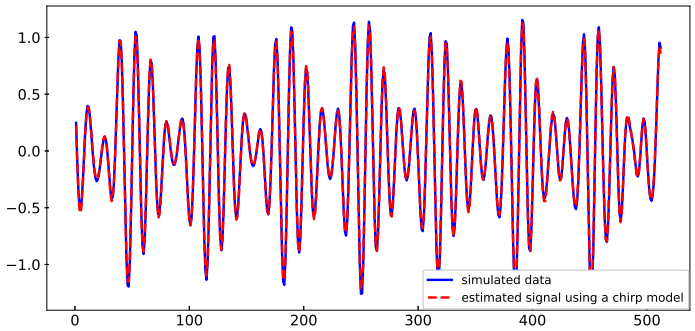

Fig. 18 Simulated data signal along with estimated signal using chirp model.

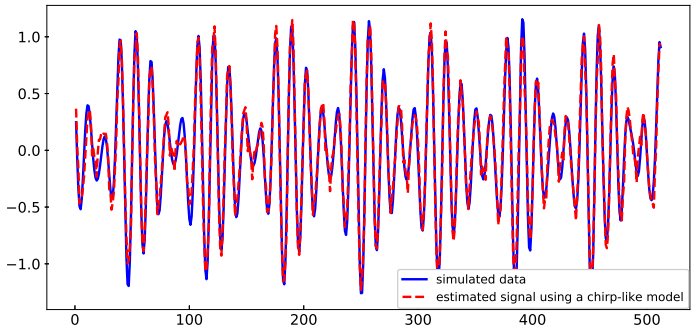

Fig. 19 Simulated data signal along with estimated signal using chirp-like model.

along with the simulated data.

It can be seen that the estimated signals using the chirp model as well as the chirp-like model envelop the simulated chirp data quite accurately. We also measure the accuracy of these fittings by calculating the residual root mean square errors (RMSEs) for the two model fittings. The residual RMSE for the chirp model fitting is 0.8750 while for the chirp-like model fitting it is 4.0727. The difference in the RMSEs is also reflected in Figure 19 as a small gap can be observed at some time points. However, this gap can be reduced with an increase in the number of components fitted to the model. It is important to note that fitting 5 components of a chirp model to a data set of size 512, needs $512^3 * 5 = 67,10,88,640$ function evaluations. On the other hand, fitting 9 sinusoid components and 1 chirp component of a chirp-like model to this data set requires $(512 * 9) + (512 * 512 * 1) = 2,66,752$ function evaluations. Therefore, a trade-off must be made between the computational complexity and accuracy of the estimated fitting.

Another important point is that by increasing the number of components, we can improve the performance of the chirp-like model and reduce the residual RMSE to get at par with that in the case of the chirp model fitting. For example, if we use 48 sinusoids and 41 chirp components of a chirp-like model to explain this simulated data,

the residual RMSE of the new fitting turns out to be 0.8720, which is less than that obtained by chirp model fitting. Moreover the computational expense $((512 * 48) + (512 * 512 * 41) = 1,07,47,904$ function evaluations which is approximately 63 times less expensive than the chirp model fitting) will still be lower than that involved in fitting a 5 component chirp model. Therefore a chirp-like model can provide a better fit at a lower expense. From here, we can also conclude that the BIC method for model selection can under-estimate the model order and may not give the best estimation performance. We believe that there is a need to develop more efficient methods of model selection for better results. However, this is not explored here and is an open problem.

6 Conclusion

Chirp signals are ubiquitous in many areas of science and engineering and hence their parameter estimation is of great significance in signal processing. But it has been observed that parameter estimation of this model, particularly using the method of least squares is computationally complex. In this paper, we put forward an alternate model, named the chirp-like model. We observe that the data that have been analyzed using chirp models can also be analyzed using the chirp-like model and estimating its parameters using sequential LSEs is simpler than that for the chirp model. We show that the LSEs and the sequential LSEs of the parameters of this model are strongly consistent and asymptotically normally distributed. The rates of convergence of the parameter estimates of this model are the same as those for the chirp model. We analyze four speech data sets, and it is observed that the proposed model can be used quite effectively to analyze these data sets.

7 Acknowledgements

The authors would like to thank the unknown reviewers for their constructive comments which have helped to improve the manuscript significantly. Part of the work of the second author has been supported by a research grant from the Science and Engineering Research Board, Government of India.

A Some Preliminary Results

To provide the proofs of the asymptotic properties established in this manuscript, we will require following results:

Lemma 1 *If $\phi \in (0, \pi)$, then the following hold true:*

(a) $\lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \cos(\phi t) = \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \sin(\phi t) = 0.$

(b) $\lim_{n\to\infty} \frac{1}{n^{k+1}} \sum_{t=1}^{n} t^k \cos^2(\phi t) = \lim_{n\to\infty} \frac{1}{n^{k+1}} \sum_{t=1}^{n} t^k \sin^2(\phi t) = \frac{1}{2(k+1)};\ k = 0, 1, 2, \cdots.$

(c) $\lim_{n\to\infty} \frac{1}{n^{k+1}} \sum_{t=1}^{n} t^k \sin(\phi t)\cos(\phi t) = 0;\ k = 0, 1, 2, \cdots.$

Proof Refer to Kundu and Nandi [19].

Lemma 2 *If $\phi \in (0, \pi)$, then except for a countable number of points, the following hold true:*

(a) $\lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \cos(\phi t^2) = \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \sin(\phi t^2) = 0.$

(b) $\lim_{n\to\infty} \frac{1}{n^{k+1}} \sum_{t=1}^{n} t^k \cos^2(\phi t^2) = \lim_{n\to\infty} \frac{1}{n^{k+1}} \sum_{t=1}^{n} t^k \sin^2(\phi t^2) = \frac{1}{2(k+1)};\ k = 0, 1, 2, \cdots.$

(c) $\lim_{n\to\infty} \frac{1}{n^{k+1}} \sum_{t=1}^{n} t^k \sin(\phi t^2)\cos(\phi t^2) = 0;\ k = 0, 1, 2, \cdots.$

Proof Refer to Lahiri [20].

Lemma 3 *If $(\phi_1, \phi_2) \in (0, \pi) \times (0, \pi)$, then except for a countable number of points, the following hold true:*

(a) $\lim_{n\to\infty} \frac{1}{n^{k+1}} \sum_{t=1}^{n} t^k \cos(\phi_1 t)\cos(\phi_2 t^2) = 0$

(b) $\lim_{n\to\infty} \frac{1}{n^{k+1}} \sum_{t=1}^{n} t^k \cos(\phi_1 t)\sin(\phi_2 t^2) = 0$

(c) $\lim_{n\to\infty} \frac{1}{n^{k+1}} \sum_{t=1}^{n} t^k \sin(\phi_1 t)\cos(\phi_2 t^2) = 0$

(d) $\lim_{n\to\infty} \frac{1}{n^{k+1}} \sum_{t=1}^{n} t^k \sin(\phi_1 t)\sin(\phi_2 t^2) = 0$

$k = 0, 1, 2, \cdots$

Proof This proof follows from the number theoretic result proved by Lahiri [20] (see Lemma 2.2.1 of the reference).

Lemma 4 *If X(t) satisfies Assumptions 1, 3 and 4, then for $k \geqslant 0$:*

(a) $\sup_{\phi} \left| \frac{1}{n^{k+1}} \sum_{t=1}^{n} t^k X(t) e^{i(\phi t)} \right| \xrightarrow{a.s.} 0$ *(b)* $\sup_{\phi} \left| \frac{1}{n^{k+1}} \sum_{t=1}^{n} t^k X(t) e^{i(\phi t^2)} \right| \xrightarrow{a.s.} 0$

Here $i = \sqrt{-i}$.

Proof These can be obtained as particular cases of Lemma 2.2.2 of Lahiri [20].

Following is the famous number theoretic conjecture of Montgomery [29].

Conjecture 1 If θ_1, θ_2, θ_1', $\theta_2' \in (0, \pi)$, then except for a countable number of points:

(a)

$$\lim_{n\to\infty} \frac{1}{n^k\sqrt{n}} \sum_{t=1}^{n} t^k \cos(\theta_1 t + \theta_2 t^2) \sin(\theta_1' t + \theta_2' t^2) = 0; \ k = 0, 1, 2, \ldots,$$

(b)

$$\lim_{n\to\infty} \frac{1}{n^k\sqrt{n}} \sum_{t=1}^{n} t^k \cos(\theta_1 t + \theta_2 t^2) \cos(\theta_1' t + \theta_2' t^2) = 0; \ k = 0, 1, 2, \ldots; \text{ if } \theta_2 \neq \theta_2',$$

$$\lim_{n\to\infty} \frac{1}{n^k\sqrt{n}} \sum_{t=1}^{n} t^k \sin(\theta_1 t + \theta_2 t^2) \sin(\theta_1' t + \theta_2' t^2) = 0; \ k = 0, 1, 2, \ldots; \text{ if } \theta_2 \neq \theta_2'.$$

The following conjecture follows from Montgomery's conjecture:

Conjecture 2 If $(\phi_1, \phi_2) \in (0, \pi) \times (0, \pi)$, then except for a countable number of points, the following hold true:

(a) $\lim_{n\to\infty} \frac{1}{n^k\sqrt{n}} \sum_{t=1}^{n} t^k \cos(\phi_1 t^2) = 0$

(b) $\lim_{n\to\infty} \frac{1}{n^k\sqrt{n}} \sum_{t=1}^{n} t^k \sin(\phi_1 t^2) = 0$

(c) $\lim_{n\to\infty} \frac{1}{n^k\sqrt{n}} \sum_{t=1}^{n} t^k \cos(\phi_1 t) \cos(\phi_2 t) = 0$

(d) $\lim_{n\to\infty} \frac{1}{n^k\sqrt{n}} \sum_{t=1}^{n} t^k \cos(\phi_1 t) \sin(\phi_2 t) = 0$

(e) $\lim_{n\to\infty} \frac{1}{n^k\sqrt{n}} \sum_{t=1}^{n} t^k \sin(\phi_1 t) \sin(\phi_2 t) = 0$

(f) $\lim_{n\to\infty} \frac{1}{n^k\sqrt{n}} \sum_{t=1}^{n} t^k \cos(\phi_1 t) \cos(\phi_2 t^2) = 0$

(g) $\lim_{n\to\infty} \frac{1}{n^k\sqrt{n}} \sum_{t=1}^{n} t^k \cos(\phi_1 t) \sin(\phi_2 t^2) = 0$

(h) $\lim_{n\to\infty} \frac{1}{n^k\sqrt{n}} \sum_{t=1}^{n} t^k \sin(\phi_1 t) \cos(\phi_2 t^2) = 0$

(i) $\lim_{n\to\infty} \frac{1}{n^k\sqrt{n}} \sum_{t=1}^{n} t^k \sin(\phi_1 t) \sin(\phi_2 t^2) = 0$

$k = 0, 1, 2, \cdots$.

In the subsequent appendices, we show that if the above conjecture holds, then the asymptotic distribution of the sequential LSEs coincides with that of the usual LSEs.

B One Component Chirp-like Model

B.1 Proofs of the asymptotic properties of the LSEs

We need the following lemmas to prove the consistency of the LSEs:

Lemma 5 *Consider the set* $S_c = \{\boldsymbol{\theta} : |\boldsymbol{\theta} - \boldsymbol{\theta}^0| > c; \boldsymbol{\theta} \in \boldsymbol{\Theta}\}$. *If the following holds true:*

$$\liminf \inf_{S_c} \frac{1}{n}(Q(\boldsymbol{\theta}) - Q(\boldsymbol{\theta}^0)) > 0 \ a.s., \tag{17}$$

then $\hat{\boldsymbol{\theta}} \xrightarrow{a.s.} \boldsymbol{\theta}^0$ *as* $n \to \infty$

Proof Let us denote $\hat{\boldsymbol{\theta}}$ by $\hat{\boldsymbol{\theta}}_n$, to highlight the fact that the estimates depend on the sample size n. Now suppose, $\hat{\boldsymbol{\theta}}_n \nrightarrow \boldsymbol{\theta}^0$, then there exists a subsequence $\{n_k\}$ of $\{n\}$, such that $\hat{\boldsymbol{\theta}}_{n_k} \nrightarrow \boldsymbol{\theta}^0$. In such a situation, one of two cases may arise:

1. $|\hat{A}_{n_k}|+|\hat{B}_{n_k}|+|\hat{C}_{n_k}|+|\hat{D}_{n_k}|$ is not bounded, that is, at least one of the $|\hat{A}_{n_k}|$ or $|\hat{B}_{n_k}|$ or $|\hat{C}_{n_k}|$ or $|\hat{D}_{n_k}| \to \infty \Rightarrow \frac{1}{n_k} Q_{n_k}(\hat{\boldsymbol{\theta}}_{n_k}) \to \infty$
   But, $\lim_{n_k\to\infty} \frac{1}{n_k} Q_{n_k}(\boldsymbol{\theta}^0) < \infty$ which implies, $\frac{1}{n_k}(Q_{n_k}(\hat{\boldsymbol{\theta}}_{n_k}) - Q_{n_k}(\boldsymbol{\theta}^0)) \to \infty$. This contradicts the fact that:

$$Q_{n_k}(\hat{\boldsymbol{\theta}}_{n_k}) \leqslant Q_{n_k}(\boldsymbol{\theta}^0), \tag{18}$$

   which holds true as $\hat{\boldsymbol{\theta}}_{n_k}$ is the LSE of $\boldsymbol{\theta}^0$.
2. $|\hat{A}_{n_k}| + |\hat{B}_{n_k}| + |\hat{C}_{n_k}| + |\hat{D}_{n_k}|$ is bounded, then there exists a $c > 0$ such that $\hat{\boldsymbol{\theta}}_{n_k} \in S_c$, for all $k = 1, 2, \cdots$. Now, since (17) is true, this contradicts (18).

Hence, the result.

Proof of Theorem 1: Consider the difference:

$$
\begin{aligned}
&\frac{1}{n}\Big(Q(\boldsymbol{\theta})-Q(\boldsymbol{\theta}^0)\Big)\\
&=\frac{1}{n}\sum_{t=1}^{n}\Big(y(t)-A\cos(\alpha t)-B\sin(\alpha t)-C\cos(\beta t^2)-D\sin(\beta t^2)\Big)^2\\
&-\frac{1}{n}\sum_{t=1}^{n}\Big(y(t)-A^0\cos(\alpha^0 t)-B^0\sin(\alpha^0 t)-C^0\cos(\beta^0 t^2)-D^0\sin(\beta^0 t^2)\Big)^2\\
&=\frac{1}{n}\sum_{t=1}^{n}\Big(A^0\cos(\alpha^0 t)-A\cos(\alpha t)+B^0\sin(\alpha^0 t)-B\sin(\alpha t)+C^0\cos(\beta^0 t^2)-C\cos(\beta t^2)+D^0\sin(\beta^0 t^2)-D\sin(\beta t^2)\Big)^2\\
&+\frac{1}{n}\sum_{t=1}^{n}X(t)\Big(A^0\cos(\alpha^0 t)-A\cos(\alpha t)+B^0\sin(\alpha^0 t)-B\sin(\alpha t)+C^0\cos(\beta^0 t^2)-C\cos(\beta t^2)+D^0\sin(\beta^0 t^2)-D\sin(\beta t^2)\Big)\\
&=f(\boldsymbol{\theta})+g(\boldsymbol{\theta}).
\end{aligned}
$$

Now using Lemma 4, it can be easily seen that:

$$\lim_{n\to\infty}\sup_{\boldsymbol{\theta}\in S_c} g(\boldsymbol{\theta})=0 \text{ a.s.} \tag{19}$$

Thus, we have:

$$\liminf_{\boldsymbol{\theta}\in S_c}\inf\frac{1}{n}\Big(Q(\boldsymbol{\theta})-Q(\boldsymbol{\theta}^0)\Big)=\liminf_{\boldsymbol{\theta}\in S_c}\inf f(\boldsymbol{\theta}).$$

Note that the proof will follow if we show that $\liminf\inf_{\boldsymbol{\theta}\in S_c} f(\boldsymbol{\theta})>0$. Consider the set $S_c=\{\boldsymbol{\theta}:|\boldsymbol{\theta}-\boldsymbol{\theta}^0|\geqslant 6c;\boldsymbol{\theta}\in\boldsymbol{\Theta}\}\subset S_c^{(1)}\cup S_c^{(2)}\cup S_c^{(3)}\cup S_c^{(4)}\cup S_c^{(5)}\cup S_c^{(6)}$, where

$$
\begin{aligned}
&S_c^{(1)}=\{\boldsymbol{\theta}:|A-A^0|\geqslant c;\boldsymbol{\theta}\in\boldsymbol{\Theta}\} \qquad S_c^{(2)}=\{\boldsymbol{\theta}:|B-B^0|\geqslant c;\boldsymbol{\theta}\in\boldsymbol{\Theta}\}\\
&S_c^{(3)}=\{\boldsymbol{\theta}:|\alpha-\alpha^0|\geqslant c;\boldsymbol{\theta}\in\boldsymbol{\Theta}\} \qquad S_c^{(4)}=\{\boldsymbol{\theta}:|C-C^0|\geqslant c;\boldsymbol{\theta}\in\boldsymbol{\Theta}\}\\
&S_c^{(5)}=\{\boldsymbol{\theta}:|D-D^0|\geqslant c;\boldsymbol{\theta}\in\boldsymbol{\Theta}\} \qquad S_c^{(6)}=\{\boldsymbol{\theta}:|\beta-\beta^0|\geqslant c;\boldsymbol{\theta}\in\boldsymbol{\Theta}\}
\end{aligned}
$$

Now, we split the set $S_c^{(1)}$ as follows:

$$
\begin{aligned}
&S_c^{(1)}=\{\boldsymbol{\theta}:|A-A^0|\geqslant c;\boldsymbol{\theta}\in\boldsymbol{\Theta}\}\\
&\subset\{\boldsymbol{\theta}:|A-A^0|\geqslant c;\boldsymbol{\theta}\in\boldsymbol{\Theta};\alpha=\alpha^0;\beta=\beta^0\}\cup\{\boldsymbol{\theta}:|A-A^0|\geqslant c;\boldsymbol{\theta}\in\boldsymbol{\Theta};\alpha\neq\alpha^0;\beta=\beta^0\}\\
&\cup\{\boldsymbol{\theta}:|A-A^0|\geqslant c;\boldsymbol{\theta}\in\boldsymbol{\Theta};\alpha=\alpha^0;\beta\neq\beta^0\}\cup\{\boldsymbol{\theta}:|A-A^0|\geqslant c;\boldsymbol{\theta}\in\boldsymbol{\Theta};\alpha\neq\alpha^0;\beta\neq\beta^0\}\\
&=S_c^{(1)_1}\cup S_c^{(1)_2}\cup S_c^{(1)_3}\cup S_c^{(1)_4}
\end{aligned}
$$

Now let us consider:

$$
\begin{aligned}
&\liminf\inf_{\boldsymbol{\theta}\in S_c^{(1)_1}} f(\boldsymbol{\theta})\\
&=\liminf\inf_{\boldsymbol{\theta}\in S_c^{(1)_1}}\frac{1}{n}\sum_{t=1}^{n}\Big(A^0\cos(\alpha^0 t)-A\cos(\alpha t)+B^0\sin(\alpha^0 t)-B\sin(\alpha t)+C^0\cos(\beta^0 t^2)-C\cos(\beta t^2)\\
&+D^0\sin(\beta^0 t^2)-D\sin(\beta t^2)\Big)^2\\
&=\liminf\inf_{\boldsymbol{\theta}\in S_c^{(1)_1}}\frac{1}{n}\sum_{t=1}^{n}\Big((A^0-A)\cos(\alpha^0 t)+(B^0-B)\sin(\alpha^0 t)+(C^0-C)\cos(\beta^0 t^2)+(D^0-D)\sin(\beta^0 t^2)\Big)^2\\
&=\frac{(A^0-A)^2}{2}+\frac{(B^0-B)^2}{2}+\frac{(C^0-C)^2}{2}+\frac{(D^0-D)^2}{2}>0
\end{aligned}
$$

$$
\begin{aligned}
&\liminf\inf_{\boldsymbol{\theta}\in S_c^{(1)_2}} f(\boldsymbol{\theta})\\
&=\liminf\inf_{\boldsymbol{\theta}\in S_c^{(1)_1}}\frac{1}{n}\sum_{t=1}^{n}\Big(A^0\cos(\alpha^0 t)-A\cos(\alpha t)+B^0\sin(\alpha^0 t)-B\sin(\alpha t)+(C^0-C)\cos(\beta^0 t^2)\\
&\quad+(D^0-D)\sin(\beta^0 t^2)\Big)^2=\frac{{A^0}^2}{2}+\frac{A^2}{2}+\frac{{B^0}^2}{2}+\frac{B^2}{2}+\frac{(C^0-C)^2}{2}+\frac{(D^0-D)^2}{2}>0
\end{aligned}
$$

$$
\begin{aligned}
&\liminf\inf_{\boldsymbol{\theta}\in S_c^{(1)_3}} f(\boldsymbol{\theta})\\
&=\liminf\inf_{\boldsymbol{\theta}\in S_c^{(1)_3}}\frac{1}{n}\sum_{t=1}^{n}\Big((A^0-A)\cos(\alpha^0 t)+(B^0-B)\sin(\alpha^0 t)+C^0\cos(\beta^0 t^2)-C\cos(\beta t^2)\\
&\quad+D^0\sin(\beta^0 t^2)-D\sin(\beta t^2)\Big)^2=\frac{(A^0-A)^2}{2}+\frac{(B^0-B)^2}{2}+\frac{{C^0}^2}{2}+\frac{C^2}{2}+\frac{{D^0}^2}{2}+\frac{D^2}{2}>0
\end{aligned}
$$

Finally, $\liminf \inf_{\boldsymbol{\theta}\in S_c^{(1)_4}} f(\boldsymbol{\theta})$

$$= \liminf \inf_{\boldsymbol{\theta}\in S_c^{(1)_1}} \frac{1}{n}\sum_{t=1}^{n}\Big(A^0\cos(\alpha^0 t) - A\cos(\alpha t) + B^0\sin(\alpha^0 t) - B\sin(\alpha t) + C^0\cos(\beta^0 t^2) - C\cos(\beta t^2)$$

$$+ D^0\sin(\beta^0 t^2) - D\sin(\beta t^2)\Big)^2 = \frac{A^{0\,2}}{2} + \frac{A^2}{2} + \frac{B^{0\,2}}{2} + \frac{B^2}{2} + \frac{C^{0\,2}}{2} + \frac{C^2}{2} + \frac{D^{0\,2}}{2} + \frac{D^2}{2} > 0$$

Note that we used lemmas 1 and 2 in all the above computations of the limits. On combining all the above, we have $\liminf \inf_{\boldsymbol{\theta}\in S_c^{(1)}} f(\boldsymbol{\theta}) > 0$. Similarly, it can be shown that the result holds for the rest of the sets. Therefore, by Lemma 5, $\hat{\boldsymbol{\theta}}$ is a strongly consistent estimator of $\boldsymbol{\theta}^0$.

□

Proof of Theorem 2: To obtain the asymptotic distribution of the LSEs, we express $\mathbf{Q}'(\hat{\boldsymbol{\theta}})$ using multivariate Taylor series expansion arount the point $\boldsymbol{\theta}^0$, as follows:

$$\mathbf{Q}'(\hat{\boldsymbol{\theta}}) - \mathbf{Q}'(\boldsymbol{\theta}^0) = (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0)\mathbf{Q}''(\bar{\boldsymbol{\theta}}). \tag{20}$$

Here, $\bar{\boldsymbol{\theta}}$ is a point between $\hat{\boldsymbol{\theta}}$ and $\boldsymbol{\theta}^0$. Since, $\hat{\boldsymbol{\theta}}$ is the LSE of $\boldsymbol{\theta}^0$, $\mathbf{Q}'(\hat{\boldsymbol{\theta}}) = 0$. Thus, we have:

$$(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0) = -\mathbf{Q}'(\boldsymbol{\theta}^0)[\mathbf{Q}''(\bar{\boldsymbol{\theta}})]^{-1}. \tag{21}$$

Multiplying both sides of (21)) by the 6×6 diagonal matrix $\mathbf{D} = diag(\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}, \frac{1}{n\sqrt{n}}, \frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}, \frac{1}{n^2\sqrt{n}})$, we get:

$$(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0)\mathbf{D}^{-1} = -\mathbf{Q}'(\boldsymbol{\theta}^0)\mathbf{D}[\mathbf{D}\mathbf{Q}''(\bar{\boldsymbol{\theta}})\mathbf{D}]^{-1}. \tag{22}$$

First, we will show that:

$$\lim_{n\to\infty} \mathbf{Q}'(\boldsymbol{\theta}^0)\mathbf{D} \xrightarrow{d} N(0, 4c\sigma^2\boldsymbol{\Sigma}). \tag{23}$$

Here,

$$\boldsymbol{\Sigma} = \begin{pmatrix} \frac{1}{2} & 0 & \frac{B^0}{4} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{-A^0}{4} & 0 & 0 & 0 \\ \frac{B^0}{4} & \frac{-A^0}{4} & \frac{A^{0\,2}+B^{0\,2}}{6} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{D^0}{6} \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{-C^0}{6} \\ 0 & 0 & 0 & \frac{D^0}{6} & \frac{-C^0}{6} & \frac{C^{0\,2}+D^{0\,2}}{10} \end{pmatrix} \tag{24}$$

To prove (23)), we compute the elements of the 6×1 vector $\mathbf{Q}'(\boldsymbol{\theta}^0)\mathbf{D} = \Big(\frac{1}{\sqrt{n}}\frac{\partial Q(\boldsymbol{\theta})}{\partial A} \quad \frac{1}{\sqrt{n}}\frac{\partial Q(\boldsymbol{\theta})}{\partial B} \quad \frac{1}{n\sqrt{n}}\frac{\partial Q(\boldsymbol{\theta})}{\partial \alpha} \quad \frac{1}{\sqrt{n}}\frac{\partial Q(\boldsymbol{\theta})}{\partial C} \quad \frac{1}{\sqrt{n}}\frac{\partial Q(\boldsymbol{\theta})}{\partial D} \quad \frac{1}{n^2\sqrt{n}}\frac{\partial Q(\boldsymbol{\theta})}{\partial \beta}\Big)$ as follows:

$$\frac{1}{\sqrt{n}}\frac{\partial Q(\boldsymbol{\theta})}{\partial A} = \frac{-2}{\sqrt{n}}\sum_{t=1}^{n}\Big(y(t) - A\cos(\alpha t) - B\sin(\alpha t) - C\cos(\beta t^2) - D\sin(\beta t^2)\Big)\cos(\alpha t)$$

$$\Rightarrow \frac{1}{\sqrt{n}}\frac{\partial Q(\boldsymbol{\theta}^0)}{\partial A} = \frac{-2}{\sqrt{n}}\sum_{t=1}^{n} X(t)\cos(\alpha^0 t).$$

Similarly, the rest of the elements can be computed and we get:

$$\mathbf{Q}'(\boldsymbol{\theta}^0)\mathbf{D} = \begin{pmatrix} \frac{-2}{\sqrt{n}}\sum_{t=1}^{n} X(t)\cos(\alpha^0 t) \\ \frac{-2}{\sqrt{n}}\sum_{t=1}^{n} X(t)\sin(\alpha^0 t) \\ \frac{-2}{n\sqrt{n}}\sum_{t=1}^{n} tX(t)\big(-A^0\sin(\alpha^0 t) + B^0\cos(\alpha^0 t)\big) \\ \frac{-2}{\sqrt{n}}\sum_{t=1}^{n} X(t)\cos(\beta^0 t^2) \\ \frac{-2}{\sqrt{n}}\sum_{t=1}^{n} X(t)\sin(\beta^0 t^2) \\ \frac{-2}{n^2\sqrt{n}}\sum_{t=1}^{n} t^2X(t)\big(-C^0\sin(\beta^0 t^2) + D^0\cos(\beta^0 t^2)\big) \end{pmatrix}.$$

Now using the Central Limit Theorem (CLT) of stochastic processes (see Fuller [7]), the above vector tends to a 6-variate Gaussian distribution with mean 0 and variance $4c\sigma^2\boldsymbol{\Sigma}$ and hence (23) holds true. Now we consider the second derivative matrix $\mathbf{D}\mathbf{Q}''(\bar{\boldsymbol{\theta}})\mathbf{D}$. Note that, since $\hat{\boldsymbol{\theta}} \xrightarrow{a.s.} \boldsymbol{\theta}^0$ as $n\to\infty$ and $\bar{\boldsymbol{\theta}}$ is a point between $\hat{\boldsymbol{\theta}}$ and $\boldsymbol{\theta}^0$,

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

Using lemmas 1, 2, 3 and 4 and after some calculations, it can be shown that:

$$\mathbf{D}\mathbf{Q}''(\boldsymbol{\theta}^0)\mathbf{D} = 2\boldsymbol{\Sigma}, \tag{25}$$

where $\boldsymbol{\Sigma}$ is as defined in (24). On combining, (22),(23) and (25), the desired result follows.

□

B.2 Proofs of the asymptotic properties of the sequential LSEs

Following lemmas are required to prove the consistency of the sequential LSEs:

Lemma 6 *Let us define the set* $M_c = \{\boldsymbol{\theta}^{(1)} : |\boldsymbol{\theta}^{(1)} - \boldsymbol{\theta}^{0(1)}| \geqslant 3c; \boldsymbol{\theta}^{(1)} \in \boldsymbol{\Theta}^{(1)}\}$. *If the following holds true:*

$$\liminf \inf_{M_c} \frac{1}{n}(Q_1(\boldsymbol{\theta}^{(1)}) - Q_1(\boldsymbol{\theta}^{0(1)})) > 0 \ a.s. \tag{26}$$

then $\tilde{\boldsymbol{\theta}}^{(1)} \xrightarrow{a.s.} \boldsymbol{\theta}^{0(1)}$ *as* $n \to \infty$

Proof This can be proved by contradiction along the same lines as Lemma 5.

Lemma 7 *Let us define the set* $N_c = \{\boldsymbol{\theta}^{(2)} : \boldsymbol{\theta}^{(2)} \in \boldsymbol{\Theta}^{(2)};\ |\boldsymbol{\theta}^{(2)} - \boldsymbol{\theta}^{0(2)}| \geqslant 3c\}$. *If for any* $c > 0$,

$$\liminf \inf_{\boldsymbol{\theta}^{(2)} \in N_c} \frac{1}{n}(Q_2(\boldsymbol{\theta}^{(2)}) - Q_2(\boldsymbol{\theta}^{0(2)})) > 0 \ a.s. \tag{27}$$

then $\tilde{\boldsymbol{\theta}}^{(2)} \xrightarrow{a.s.} \boldsymbol{\theta}^{0(2)}$ *as* $n \to \infty$.

Proof This can be proved by contradiction along the same lines as Lemma 5.

Proof of Theorem 3: First we prove the consistency of the parameter estimates of the sinusoidal component, $\tilde{\boldsymbol{\theta}}^{(1)}$. For this, consider the difference:

$$\begin{aligned}
&\frac{1}{n}(Q_1(\boldsymbol{\theta}^{(1)}) - Q_1(\boldsymbol{\theta}^{0(1)})) \\
&= \frac{1}{n}\left[\sum_{t=1}^{n}\Big(y(t) - A\cos(\alpha t) - B\sin(\alpha t)\Big)^2 - \Big(y(t) - A^0\cos(\alpha^0 t) - B^0\sin(\alpha^0 t)\Big)^2\right] \\
&= \frac{1}{n}\sum_{t=1}^{n}\Big(A^0\cos(\alpha^0 t) - A\cos(\alpha t) + B^0\sin(\alpha^0 t) - B\sin(\alpha t) + C^0\cos(\beta^0 t^2) + D^0\sin(\beta^0 t^2) + X(t)\Big)^2 \\
&\quad - \frac{1}{n}\sum_{t=1}^{n}\Big(C^0\cos(\beta^0 t^2) + D^0\sin(\beta^0 t^2) + X(t)\Big)^2 \\
&= \frac{1}{n}\sum_{t=1}^{n}\Big(A^0\cos(\alpha^0 t) + B^0\sin(\alpha^0 t) - A\cos(\alpha t) - B\sin(\alpha t)\Big)^2 \\
&\quad + \frac{2}{n}\sum_{t=1}^{n}(C^0\cos(\beta^0 t^2) + D^0\sin(\beta^0 t^2) + X(t))\Big)\Big(A^0\cos(\alpha^0 t) + B^0\sin(\alpha^0 t) - A\cos(\alpha t) - B\sin(\alpha t)\Big) \\
&= f_1(\boldsymbol{\theta}^{(1)}) + g_1(\boldsymbol{\theta}^{(1)}).
\end{aligned}$$

Here,

$$f_1(\boldsymbol{\theta}^{(1)}) = \frac{1}{n}\sum_{t=1}^{n}\Big(A^0\cos(\alpha^0 t) + B^0\sin(\alpha^0 t) - A\cos(\alpha t) - B\sin(\alpha t)\Big)^2 \text{ and,}$$

$$g_1(\boldsymbol{\theta}^{(1)}) = \frac{2}{n}\sum_{t=1}^{n}\Big(C^0\cos(\beta^0 t^2) + D^0\sin(\beta^0 t^2) + X(t)\Big)\Big(A^0\cos(\alpha^0 t) + B^0\sin(\alpha^0 t) - A\cos(\alpha t) - B\sin(\alpha t)\Big)$$

Now using lemmas 3 and 4, it is easy to see that:

$$\sup_{\boldsymbol{\theta} \in M_c} |g_1(\boldsymbol{\theta}^{(1)})| \xrightarrow{a.s.} 0.$$

Thus if we prove that $\liminf \inf_{M_c} f_1(\boldsymbol{\theta}^{(1)}) > 0$ a.s., it will follow that $\liminf \inf_{M_c} \frac{1}{n}(Q_1(\boldsymbol{\theta}^{(1)}) - Q_1(\boldsymbol{\theta}^{0(1)})) > 0$ a.s. . First consider the set $M_c = \{\boldsymbol{\theta}^{(1)} : |\boldsymbol{\theta}^{(1)} - \boldsymbol{\theta}^{0(1)}| \geqslant 3c; \boldsymbol{\theta}^{(1)} \in \boldsymbol{\Theta}^{(1)}\}$. It is evident that:

$$M_c \subset M_c^{(1)} \cup M_c^{(2)} \cup M_c^{(3)},$$

where $M_c^{(1)} = \{\boldsymbol{\theta}^{(1)} : |A - A^0| \geqslant c; \boldsymbol{\theta}^{(1)} \in \boldsymbol{\Theta}^{(1)}\}$, $M_c^{(2)} = \{\boldsymbol{\theta}^{(1)} : |B - B^0| \geqslant c; \boldsymbol{\theta}^{(1)} \in \boldsymbol{\Theta}^{(1)}\}$ and $M_c^{(3)} = \{\boldsymbol{\theta}^{(1)} : |\alpha - \alpha^0| \geqslant c; \boldsymbol{\theta}^{(1)} \in \boldsymbol{\Theta}^{(1)}\}$. Now we further split the set $M_c^{(1)}$ which can be written as: $M_c^{(1)_1} \cup M_c^{(1)_2}$, where

$M_c^{(1)_1} = \{\boldsymbol{\theta}^{(1)} : |A - A^0| \geqslant c; \boldsymbol{\theta}^{(1)} \in \boldsymbol{\Theta}^{(1)}; \alpha = \alpha^0\}$ and $M_c^{(1)_2} = \{\boldsymbol{\theta}^{(1)} : |A - A^0| \geqslant c; \boldsymbol{\theta}^{(1)} \in \boldsymbol{\Theta}^{(1)}; \alpha \neq \alpha^0\}$

Consider, $\liminf \inf_{M_c^{(1)_1}} f_1(\boldsymbol{\theta}^{(1)}) = \liminf \inf_{M_c^{(1)_1}} \frac{1}{n}\sum_{t=1}^{n}\Big(A^0\cos(\alpha^0 t) + B^0\sin(\alpha^0 t) - A\cos(\alpha t) - B\sin(\alpha t)\Big)^2$

$= \frac{(A^0 - A)^2}{2} + \frac{(B^0 - B)^2}{2} > 0$ a.s. (using Lemma 1).

Again, using Lemma 1, $\liminf \inf_{M_c^{(1)_2}} \frac{1}{n}\sum_{t=1}^{n}\Big(A^0\cos(\alpha^0 t) + B^0\sin(\alpha^0 t) - A\cos(\alpha t) - B\sin(\alpha t)\Big)^2$

$$= \frac{A^{0\,2}}{2} + \frac{B^{0\,2}}{2} + \frac{A^2}{2} + \frac{B^2}{2} > 0 \text{ a.s..}$$

Similarly, it can be shown that $\liminf \inf_{M_c^{(2)}} f_1(\boldsymbol{\theta}^{(1)}) > 0$ a.s. and $\liminf \inf_{M_c^{(3)}} f_1(\boldsymbol{\theta}^{(1)}) > 0$ a.s.. Now using Lemma 6, $\tilde{A}$, $\tilde{B}$ and $\tilde{\alpha}$ are strongly consistent estimators of A^0, B^0 and α^0 respectively. To prove the consistency of the chirp parameter sequential estimates, $\tilde{C}$, $\tilde{D}$ and $\tilde{\beta}$, we need the following lemma:

Lemma 8 *If Assumptions 1,2 and 3 are satisfied, then:*

$$(\tilde{\boldsymbol{\theta}}^{(1)} - \boldsymbol{\theta}^{0(1)})(\sqrt{n}\boldsymbol{D}_1)^{-1} \xrightarrow{a.s.} 0.$$

Here, $\boldsymbol{D}_1 = diag(\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}, \frac{1}{n\sqrt{n}})$.

Proof Consider the error sum of squares: $Q_1(\boldsymbol{\theta}) = \frac{1}{n}\sum_{t=1}^{n}(y(t) - A\cos(\alpha t) - B\sin(\alpha t))^2$.

By Taylor series expansion of $\mathbf{Q}_1'(\tilde{\boldsymbol{\theta}}^{(1)})$ around the point $\boldsymbol{\theta}^{0(1)}$, we get:

$$\mathbf{Q}_1'(\tilde{\boldsymbol{\theta}}^{(1)}) - \mathbf{Q}_1'(\boldsymbol{\theta}^{0(1)}) = (\tilde{\boldsymbol{\theta}}^{(1)} - \boldsymbol{\theta}^{0(1)})\mathbf{Q}_1''(\bar{\boldsymbol{\theta}}^{(1)}) \tag{28}$$

where, $\bar{\boldsymbol{\theta}}^{(1)}$ is a point lying between $\tilde{\boldsymbol{\theta}}^{(1)}$ and $\boldsymbol{\theta}^{0(1)}$. Since, $\tilde{\boldsymbol{\theta}}^{(1)}$ minimises $Q_1(\boldsymbol{\theta})$, it implies that $\mathbf{Q}_1'(\tilde{\boldsymbol{\theta}}^{(1)}) = 0$ and therefore (28) can be written as:

$$(\tilde{\boldsymbol{\theta}}^{(1)} - \boldsymbol{\theta}^{0(1)}) = -\mathbf{Q}_1'(\boldsymbol{\theta}^{0(1)})[\mathbf{Q}_1''(\bar{\boldsymbol{\theta}}^{(1)})]^{-1} \tag{29}$$

$$\Rightarrow (\tilde{\boldsymbol{\theta}}^{(1)} - \boldsymbol{\theta}^{0(1)})(\sqrt{n}\mathbf{D}_1)^{-1} = [-\frac{1}{\sqrt{n}}\mathbf{Q}_1'(\boldsymbol{\theta}^{0(1)})\mathbf{D}_1][\mathbf{D}_1\mathbf{Q}_1''(\bar{\boldsymbol{\theta}}^{(1)})\mathbf{D}_1]^{-1} \tag{30}$$

Now let us calculate the right hand side explicitly. First consider the first derivative vector $\frac{1}{\sqrt{n}}\mathbf{Q}_1'(\boldsymbol{\theta}^{0(1)})\mathbf{D}_1$.

$$\frac{1}{\sqrt{n}}\mathbf{Q}_1'(\boldsymbol{\theta}^{0(1)})\mathbf{D}_1 = \begin{pmatrix} \frac{1}{n}\frac{\partial Q_1(\boldsymbol{\theta}^{0(1)})}{\partial A} & \frac{1}{n}\frac{\partial Q_1(\boldsymbol{\theta}^{0(1)})}{\partial B} & \frac{1}{n^2}\frac{\partial Q_1(\boldsymbol{\theta}^{0(1)})}{\partial \alpha} \end{pmatrix}$$

By straight forward calculations and using lemmas 3 and 4*(a)*, one can easily see that:

$$\frac{1}{\sqrt{n}}\mathbf{Q}_1'(\boldsymbol{\theta}^{0(1)})\mathbf{D}_1 \to 0 \text{ a.s.} \tag{31}$$

Now let us consider the second derivative matrix $\mathbf{D}_1\mathbf{Q}_1''(\bar{\boldsymbol{\theta}}^{(1)})\mathbf{D}_1$. Since $\tilde{\boldsymbol{\theta}}^{(1)} \xrightarrow{a.s.} \boldsymbol{\theta}^{0(1)}$ and $\bar{\boldsymbol{\theta}}^{(1)}$ is a point between them, we have:

$$\mathbf{D}_1\mathbf{Q}_1''(\bar{\boldsymbol{\theta}}^{(1)})\mathbf{D}_1 = \lim_{n\to\infty}\mathbf{D}_1\mathbf{Q}_1''(\boldsymbol{\theta}^{0(1)})\mathbf{D}_1$$

Again by routine calculations and using lemmas 1, 3 and 4*(a)* , one can evaluate each element of this 3×3 matrix, and get:

$$\lim_{n\to\infty}\mathbf{D}_1\mathbf{Q}_1''(\boldsymbol{\theta}^{0(1)})\mathbf{D}_1 = 2\boldsymbol{\Sigma}_1, \tag{32}$$

where $\boldsymbol{\Sigma}_1 = \begin{pmatrix} \frac{1}{2} & 0 & \frac{B^0}{4} \\ 0 & \frac{1}{2} & \frac{-A^0}{4} \\ \frac{B^0}{4} & \frac{-A^0}{4} & \frac{A^{0\,2}+B^{0\,2}}{6} \end{pmatrix} > 0$, a positive definite matrix. Hence combining (31) and (32), we get the desired result.

Using the above lemma, we get the following relationship between the sinusoidal component of the model and its estimate:

$$\tilde{A}\cos(\tilde{\alpha}t) + \tilde{B}\sin(\tilde{\alpha}t) = A^0\cos(\alpha^0 t) + B^0\sin(\alpha^0 t) + o(1) \tag{33}$$

Now to prove the consistency of $\tilde{\boldsymbol{\theta}}^{(1)} = (\tilde{C}, \tilde{D}, \tilde{\beta})$, we consider the following difference:

$$\begin{aligned}
&\frac{1}{n}(Q_2(\boldsymbol{\theta}^{(2)}) - Q_2(\boldsymbol{\theta}^{0(2)})) \\
&= \frac{1}{n}\left[\sum_{t=1}^{n}\left(y_1(t) - C\cos(\beta t^2) - D\sin(\beta t^2)\right)^2 - \left(y_1(t) - C^0\cos(\beta^0 t^2) - D^0\sin(\beta^0 t^2)\right)^2\right] \\
&= \frac{1}{n}\sum_{t=1}^{n}\left(C^0\cos(\beta^0 t^2) + D^0\sin(\beta^0 t^2) - C\cos(\beta t^2) - D\sin(\beta t^2)\right)^2 \\
&\quad + \frac{2}{n}\sum_{t=1}^{n}(A^0\cos(\alpha^0 t) + B^0\sin(\alpha^0 t^2) + X(t))\left(C^0\cos(\beta^0 t^2) + D^0\sin(\beta^0 t^2) - C\cos(\beta t^2) - D\sin(\beta t^2)\right) \\
&= f_2(\boldsymbol{\theta}^{(2)}) + g_2(\boldsymbol{\theta}^{(2)}).
\end{aligned}$$

Using lemmas 3 and 4, we have

$$\sup_{\boldsymbol{\theta}\in N_c}|g_2(\boldsymbol{\theta}^{(2)})| \xrightarrow{a.s.} 0,$$

and using straight forward, but lengthy calculations and splitting the set N_c, similar to the splitting of set M_c, before, it can be shown that $\liminf \inf_{\boldsymbol{\xi}\in N_c} f_2(\boldsymbol{\theta}^{(2)}) > 0$.

Thus, $\tilde{\boldsymbol{\theta}}^{(2)} \xrightarrow{a.s.} \boldsymbol{\theta}^{0(2)}$ as $n \to \infty$ by Lemma 7. Hence, the result.

□

Proof of Theorem 4: We first examine the asymptotic distribution of the sequential estimates of the sinusoidal component, that is $\tilde{\boldsymbol{\theta}}^{(1)}$ From 29, we have:

$$(\tilde{\boldsymbol{\theta}}^{(1)} - \boldsymbol{\theta}^{0\,(1)})\mathbf{D}_1^{-1} = -\mathbf{Q}_1'(\boldsymbol{\theta}^{0\,(1)})\mathbf{D}_1[\mathbf{D}_1\mathbf{Q}_1''(\bar{\boldsymbol{\theta}}^{(1)})\mathbf{D}_1]^{-1}.$$

First we show that $\mathbf{Q}_1'(\boldsymbol{\theta}^{0\,(1)})\mathbf{D}_1 \rightarrow N_3(0, 4\sigma^2 c\boldsymbol{\Sigma}_1)$. We compute the elements of the derivative vector $\mathbf{Q}_1'(\boldsymbol{\theta}^{0\,(1)})$ and using Conjecture 2 *(e), (f), (g)* and *(h)* (see section A), we obtain:

$$\mathbf{Q}_1'(\boldsymbol{\theta}^{0\,(1)})\mathbf{D}_1 \overset{a.eq.}{=} -2\begin{pmatrix} \frac{1}{\sqrt{n}}\sum_{t=1}^{n} X(t)\cos(\alpha^0 t) \\ \frac{1}{\sqrt{n}}\sum_{t=1}^{n} X(t)\sin(\alpha^0 t) \\ \frac{1}{n\sqrt{n}}\sum_{t=1}^{n} tX(t)(-A_1^0\sin(\alpha^0 t) + B^0\cos(\alpha^0 t)) \end{pmatrix}. \tag{34}$$

Here, $\overset{a.eq.}{=}$ means asymptotically equivalent. Now again using CLT, the right hand side of (34) tends to 3-variate Gaussian distribution with mean 0 and variance-covariance matrix, $4\sigma^2 c\boldsymbol{\Sigma}_1$. Using this and (32), we have the desired result.

Next we determine the asymptotic distribution of $\tilde{\boldsymbol{\theta}}^{(2)}$. For this, we consider the error sum of squares, $Q_2(\boldsymbol{\theta}^{(2)})$ as defined in (12). Let $\boldsymbol{Q}_2'(\boldsymbol{\theta}^{(2)})$ be the first derivative vector and $\boldsymbol{Q}_2''(\boldsymbol{\theta}^{(2)})$, the second derivative matrix of $Q_2(\boldsymbol{\theta}^{(2)})$.Using multivariate Taylor series expansion, we expand $\boldsymbol{Q}_2'(\tilde{\boldsymbol{\theta}}^{(2)})$ around the point $\boldsymbol{\theta}^{0\,(2)}$, and get:

$$(\tilde{\boldsymbol{\theta}}^{(2)} - \boldsymbol{\theta}^{0\,(2)}) = -\boldsymbol{Q}_2'(\boldsymbol{\theta}^{0\,(2)})[\boldsymbol{Q}_2''(\bar{\boldsymbol{\theta}}^{(2)})]^{-1}.$$

Multiplying both sides by the matrix $\boldsymbol{D}_2^{-1}$, where $\boldsymbol{D}_2 = diag(\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}, \frac{1}{n^2\sqrt{n}})$, we get:

$$(\tilde{\boldsymbol{\theta}}^{(2)} - \boldsymbol{\theta}^{0\,(2)})\mathbf{D}_2^{-1} = -\boldsymbol{Q}_2'(\boldsymbol{\theta}^{0\,(2)})\mathbf{D}_2[\mathbf{D}_2\boldsymbol{Q}_2''(\bar{\boldsymbol{\theta}}^{(2)})\mathbf{D}_2]^{-1}.$$

Now when we evaluate the first derivative vector $\boldsymbol{Q}_2'(\boldsymbol{\theta}^{0\,(2)})\mathbf{D}_2$, using Conjecture 2 *(a)* (see section A), we obtain:

$$\boldsymbol{Q}_2'(\boldsymbol{\theta}^{0\,(2)})\mathbf{D}_2 \overset{a.eq.}{=} -2\begin{pmatrix} \frac{1}{\sqrt{n}}\sum_{t=1}^{n} X(t)\cos(\beta^0 t^2) \\ \frac{1}{\sqrt{n}}\sum_{t=1}^{n} X(t)\sin(\beta^0 t^2) \\ \frac{1}{n^2\sqrt{n}}\sum_{t=1}^{n} tX(t)(-C^0\sin(\beta^0 t^2) + D^0\cos(\beta^0 t^2)) \end{pmatrix}. \tag{35}$$

Again using the CLT, the vector on the right hand side of (35) tends to $N_3(0, 4\sigma^2 c\boldsymbol{\Sigma}_2)$, where $\boldsymbol{\Sigma}_2 = \begin{pmatrix} \frac{1}{2} & 0 & \frac{D^0}{6} \\ 0 & \frac{1}{2} & \frac{-C^0}{6} \\ \frac{D^0}{6} & \frac{-C^0}{6} & \frac{C^{0\,2}+D^{0\,2}}{10} \end{pmatrix} > 0$.

Note that:

$$\lim_{n\to\infty} \mathbf{D}_2\boldsymbol{Q}_2''(\bar{\boldsymbol{\theta}}^{(2)})\mathbf{D}_2 = \lim_{n\to\infty} \mathbf{D}_2\boldsymbol{Q}_2''(\boldsymbol{\theta}^{0\,(2)})\mathbf{D}_2.$$

On computing the second derivative 3×3 matrix $\mathbf{D}_2\boldsymbol{Q}_2''(\boldsymbol{\theta}^{0\,(2)})\mathbf{D}_2$ and using lemmas 2, 3 and 4 *(b)*, we get:

$$\lim_{n\to\infty} \mathbf{D}_2\boldsymbol{Q}_2''(\boldsymbol{\theta}^{0\,(2)})\mathbf{D}_2 = 2\boldsymbol{\Sigma}_2. \tag{36}$$

Combining results (35) and (36), we get the stated asymptotic distribution of $\tilde{\boldsymbol{\theta}}^{(2)}$. Hence, the result.

□

C Multiple Component Chirp-like model

C.1 Proofs of the asymptotic properties of the LSEs

Proof of Theorm 6: Consider the error sum of squares, defined in (14). Let us denote $\mathbf{Q}'(\boldsymbol{\vartheta})$ as the $3(p+q)\times 1$ first derivative vector and $\mathbf{Q}''(\boldsymbol{\vartheta})$ as the $3(p+q)\times 3(p+q)$ second derivative matrix. Using multivariate Taylor series expansion, we have:

$$\mathbf{Q}'(\hat{\boldsymbol{\vartheta}}) - \mathbf{Q}'(\boldsymbol{\vartheta}^0) = (\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}^0)\mathbf{Q}''(\bar{\boldsymbol{\vartheta}}).$$

Here $\bar{\boldsymbol{\vartheta}}$ is a point between $\hat{\boldsymbol{\vartheta}}$ and $\boldsymbol{\vartheta}^0$. Now using the fact that $\mathbf{Q}'(\hat{\boldsymbol{\vartheta}}) = 0$ and multiplying both sides of the above equation by $\mathfrak{D}^{-1}$, we have:

$$(\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}^0)\mathfrak{D}^{-1} = -\mathbf{Q}'(\hat{\boldsymbol{\vartheta}})\mathfrak{D}[\mathfrak{D}\mathbf{Q}''(\bar{\boldsymbol{\vartheta}})\mathfrak{D}]^{-1}.$$

Also note that, $(\hat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}^0)\mathfrak{D}^{-1} = \Big((\hat{\boldsymbol{\theta}}_1^{(1)} - \boldsymbol{\theta}_1^{0\,(1)}), \cdots, (\hat{\boldsymbol{\theta}}_p^{(1)} - \boldsymbol{\theta}_p^{0\,(1)}), (\hat{\boldsymbol{\theta}}_1^{(2)} - \boldsymbol{\theta}_1^{0\,(2)}), \cdots, (\hat{\boldsymbol{\theta}}_q^{(2)} - \boldsymbol{\theta}_q^{0\,(2)})\Big)\mathfrak{D}^{-1}$.

Now we evaluate the elements of the vector $\mathbf{Q}'(\boldsymbol{\vartheta}^0)$ and the matrix $\mathbf{Q}''(\bar{\boldsymbol{\vartheta}})$:

$$\frac{\partial Q(\boldsymbol{\vartheta})}{\partial A_j}\Big|_{\boldsymbol{\vartheta}^0} = -2\sum_{t=1}^{n} X(t)\cos(\alpha_j^0 t), \quad \frac{\partial Q(\boldsymbol{\vartheta})}{\partial B_j}\Big|_{\boldsymbol{\vartheta}^0} = -2\sum_{t=1}^{n} X(t)\sin(\alpha_j^0 t), \text{ and}$$

$$\frac{\partial Q(\boldsymbol{\vartheta})}{\partial \alpha_j}\Big|_{\boldsymbol{\vartheta}^0} = -2\sum_{t=1}^{n} tX(t)\Big(-A_j^0\sin(\alpha_j^0 t) + B_j^0\cos(\alpha_j^0 t)\Big), \text{ for } j = 1, \cdots, p.$$

Similarly, for $k = 1, \cdots, q$, $\frac{\partial Q(\boldsymbol{\vartheta})}{\partial C_k}\Big|_{\boldsymbol{\vartheta}^0} = -2\sum_{t=1}^{n} X(t)\cos(\beta_k^0 t^2)$, $\frac{\partial Q(\boldsymbol{\vartheta})}{\partial D_k}\Big|_{\boldsymbol{\vartheta}^0} = -2\sum_{t=1}^{n} X(t)\sin(\beta_k^0 t^2)$ and

$$\frac{\partial Q(\boldsymbol{\vartheta})}{\partial \beta_k}\Big|_{\boldsymbol{\vartheta}^0} = -2\sum_{t=1}^{n} t^2 X(t)\Big(-C_k^0\sin(\beta_k^0 t^2) + D_k^0\cos(\beta_k^0 t)\Big).$$

$$\frac{\partial^2 Q(\boldsymbol{\vartheta})}{\partial A_j^2}\Big|_{\boldsymbol{\vartheta}^0} = 2\sum_{t=1}^{n}\cos^2(\alpha_j^0 t),\ \frac{\partial^2 Q(\boldsymbol{\vartheta})}{\partial B_j^2}\Big|_{\boldsymbol{\vartheta}^0} = 2\sum_{t=1}^{n}\sin^2(\alpha_j^0 t),\ j=1,\cdots,p,$$

$$\frac{\partial^2 Q(\boldsymbol{\vartheta})}{\partial C_k^2}\Big|_{\boldsymbol{\vartheta}^0} = 2\sum_{t=1}^{n}\cos^2(\beta_k^0 t^2) \text{ and } \frac{\partial^2 Q(\boldsymbol{\vartheta})}{\partial D_k^2}\Big|_{\boldsymbol{\vartheta}^0} = 2\sum_{t=1}^{n}\sin^2(\beta_k^0 t^2),\ k=1,\cdots,q.$$

$$\frac{\partial^2 Q(\boldsymbol{\vartheta})}{\partial A_j \partial B_j}\Big|_{\boldsymbol{\vartheta}^0} = 2\sum_{t=1}^{n}\sin(\alpha_j^0 t)\cos(\alpha_j^0 t),$$

$$\frac{\partial^2 Q(\boldsymbol{\vartheta})}{\partial A_j \partial \alpha_j}\Big|_{\boldsymbol{\vartheta}^0} = 2\sum_{t=1}^{n} tX(t)\sin(\alpha_j^0 t) - 2A_j^0\sum_{t=1}^{n} t\cos(\alpha_j^0 t)\sin(\alpha_j^0 t) + 2B_j^0\sum_{t=1}^{n} t\cos^2(\alpha_j^0 t),$$

$$\frac{\partial^2 Q(\boldsymbol{\vartheta})}{\partial A_j \partial C_k}\Big|_{\boldsymbol{\vartheta}^0} = 2\sum_{t=1}^{n}\cos(\beta_k^0 t^2)\cos(\alpha_j^0 t),\ \frac{\partial^2 Q(\boldsymbol{\vartheta})}{\partial A_j \partial D_k}\Big|_{\boldsymbol{\vartheta}^0} = 2\sum_{t=1}^{n}\sin(\beta_k^0 t^2)\cos(\alpha_j^0 t),$$

$$\frac{\partial^2 Q(\boldsymbol{\vartheta})}{\partial A_j \partial \beta_k}\Big|_{\boldsymbol{\vartheta}^0} = -2C_k^0\sum_{t=1}^{n} t^2\cos(\alpha_j^0 t)\sin(\beta_k^0 t^2) + 2D_k^0\sum_{t=1}^{n} t^2\cos(\alpha_j^0 t)\cos(\beta_k^0 t^2).$$

Similarly the rest of the partial derivatives can be computed and using lemmas 1, 2, 3 and 4, it can be shown that:

$$\mathfrak{D}\mathbf{Q}''(\bar{\boldsymbol{\vartheta}})\mathfrak{D} \to 2\mathcal{E}(\boldsymbol{\vartheta}^0).$$

Now, using CLT on the first derivative vector, $\mathbf{Q}'(\boldsymbol{\vartheta}^0)\mathfrak{D}$, it can be shown that it converges to a multivariate Gaussian distribution. Using routine calculations, and again using lemmas 1, 2, 3 and 4, we compute the asymptotic variances for each of the elements and their covariances and we get:

$$\mathbf{Q}'(\boldsymbol{\vartheta}^0)\mathfrak{D} \xrightarrow{d} N_{3(p+q)}(0, 4c\sigma^2\mathcal{E}(\boldsymbol{\vartheta}^0)).$$

Hence, the result.

□

C.2 Proofs of the asymptotic properties of the LSEs

To prove theorems 7 and 8, we need the following lemmas:

Lemma 9 *(a) Consider the set* $M_c^{(j)} = \{\boldsymbol{\theta}_j^{(1)} : |\boldsymbol{\theta}_j^{(1)} - \boldsymbol{\theta}_j^{0(1)}| \geqslant 3c; \boldsymbol{\theta}_j^{(1)} \in \boldsymbol{\Theta}^{(1)}\},\ j=1,\cdots,p.$ *If the following holds true:*

$$\liminf \inf_{M_c^{(j)}} \frac{1}{n}(Q_{2j-1}(\boldsymbol{\theta}_j^{(1)}) - Q_{2j-1}(\boldsymbol{\theta}_j^{0(1)})) > 0 \ a.s. \tag{37}$$

then $\tilde{\boldsymbol{\theta}}_j^{(1)} \xrightarrow{a.s.} \boldsymbol{\theta}_j^{0(1)}$ *as* $n \to \infty$

(b) Let us define the set $N_c^{(k)} = \{\boldsymbol{\theta}_k^{(2)} : \boldsymbol{\theta}_k^{(2)} \in \boldsymbol{\Theta}^{(2)};\ |\boldsymbol{\theta}_k^{(2)} - \boldsymbol{\theta}_k^{0(2)}| \geqslant 3c\},\ k=1,\cdots,q.$ *If for any* $c > 0$,

$$\liminf \inf_{\boldsymbol{\theta}_k^{(2)} \in N_c^{(k)}} \frac{1}{n}(Q_{2k}(\boldsymbol{\theta}_k^{(2)}) - Q_{2k}(\boldsymbol{\theta}_k^{0(2)})) > 0 \ a.s. \tag{38}$$

then $\tilde{\boldsymbol{\theta}}_k^{(2)} \xrightarrow{a.s.} \boldsymbol{\theta}_k^{0(2)}$ *as* $n \to \infty$.

Proof This can be proved by contradiction along the same lines as Lemma 5.

Lemma 10 *If the Assumptions 1, 3 and 4 are satisfied, then for* $j \leqslant p$ *and* $k \leqslant q$*:*

(a) $(\tilde{\boldsymbol{\theta}}_j - \boldsymbol{\theta}_j^0)(\sqrt{n}\boldsymbol{D}_1)^{-1} \xrightarrow{a.s.} 0.$

(b) $(\tilde{\boldsymbol{\xi}}_k - \boldsymbol{\xi}_k^0)(\sqrt{n}\boldsymbol{D}_2)^{-1} \xrightarrow{a.s.} 0.$

Here, $\boldsymbol{D}_1 = diag(\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}, \frac{1}{n\sqrt{n}})$ *and* $\boldsymbol{D}_2 = diag(\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}, \frac{1}{n^2\sqrt{n}})$.

Proof This proof can be obtained along the same lines as Lemma 8.

Now the proofs of theorems 7 and 8 can be obtained by using the above lemmas and following the same argument as in Theorem 3.

Next, we examine the situation when the number of components is overestimated (see Theorem 9). The proof of Theorem 9 will follow consequently from the below-stated lemmas:

Lemma 11 *If* $X(t)$*, is the error component as defined before, and if* $\tilde{A}$*,* $\tilde{B}$ *and* $\tilde{\alpha}$ *are obtained by minimizing the following function:*

$$Q_{p+q+1}(\boldsymbol{\theta}^{(1)}) = \frac{1}{n}\sum_{t=1}^{N}\Big(X(t) - A\cos(\alpha t) - B\sin(\alpha t)\Big)^2,$$

then $\tilde{A} \xrightarrow{a.s.} 0$ *and* $\tilde{B} \xrightarrow{a.s.} 0$.

Proof The sum of squares function $Q_{p+q+1}(\boldsymbol{\theta}^{(1)})$ can be written as:

$$\frac{1}{n}\sum_{t=1}^{n}X^2(t)-\frac{2}{n}\sum_{t=1}^{n}X(t)\Big(A\cos(\alpha t)+B\sin(\alpha t)\Big)+\frac{A^2+B^2}{2}+o(1)$$
$$=R(\boldsymbol{\theta}^{(1)})+o(1).$$

Since the difference between $Q_{p+q+1}(\boldsymbol{\theta}^{(1)})$ and $R(\boldsymbol{\theta}^{(1)})$ is $o(1)$, replacing former with latter will have negligible effect on the estimators. Thus, we have

$$\tilde{A}=\frac{2}{n}\sum_{t=1}^{n}X(t)\cos(\alpha t)+o(1) \text{ and } \tilde{B}=\frac{2}{n}\sum_{t=1}^{n}X(t)\sin(\alpha t)+o(1).$$

Now using Lemma 4 *(a)*, the result follows.

Lemma 12 *If $X(t)$, is the error component as defined before, and if $\tilde{C}$, $\tilde{D}$ and $\tilde{\beta}$ are obtained by minimizing the following function:*

$$\frac{1}{n}\sum_{t=1}^{N}\Big(X(t)-C\cos(\beta t^2)-D\sin(\beta t^2)\Big)^2,$$

then $\tilde{C}\xrightarrow{a.s.}0$ and $\tilde{D}\xrightarrow{a.s.}0$.

Proof The proof of this lemma follows along the same lines as Lemma 11.

Now we provide the proof of the fact that the sequential LSEs have the same asymptotic distribution as the LSEs.

Proof of Theorem 10: (a) By Taylor series expansion of $\boldsymbol{Q}_1'(\tilde{\boldsymbol{\theta}}_1^{(1)})$ around the point $\boldsymbol{\theta}_1^{0(1)}$, we have:

$$(\tilde{\boldsymbol{\theta}}_1^{(1)}-\boldsymbol{\theta}_1^{0(1)})=-\boldsymbol{Q}_1'(\boldsymbol{\theta}_1^{0(1)})[\boldsymbol{Q}_1''(\bar{\boldsymbol{\theta}}_1^{(1)})]^{-1}$$

Multiplying both sides by the matrix $\boldsymbol{D}_1^{-1}$, where $\boldsymbol{D}_1=diag(\frac{1}{\sqrt{n}},\frac{1}{\sqrt{n}},\frac{1}{n\sqrt{n}})$, we get:

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

First we show that $\boldsymbol{Q}_1'(\boldsymbol{\theta}_1^{0(1)})\boldsymbol{D}_1\to N_3(0,4\sigma^2c\boldsymbol{\Sigma}_1^{(1)})$.
To prove this, we compute the elements of the derivative vector $\boldsymbol{Q}_1'(\boldsymbol{\theta}_1^{0(1)})$:

$$\frac{\partial Q_1(\boldsymbol{\theta}_1^{0(1)})}{\partial A_1}=-2\sum_{t=1}^{n}\Big(\sum_{j=2}^{p}(A_j^0\cos(\alpha_j^0t)+B_j^0\sin(\alpha_j^0t))+\sum_{k=1}^{q}(C_k^0\cos(\beta_k^0t^2)+D_k^0\sin(\beta_k^0t^2))+X(t)\Big)\cos(\alpha_1^0t),$$
$$\frac{\partial Q_1(\boldsymbol{\theta}_1^{0(1)})}{\partial B_1}=-2\sum_{t=1}^{n}\Big(\sum_{j=2}^{p}(A_j^0\cos(\alpha_j^0t)+B_j^0\sin(\alpha_j^0t))+\sum_{k=1}^{q}(C_k^0\cos(\beta_k^0t^2)+D_k^0\sin(\beta_k^0t^2))+X(t)\Big)\sin(\alpha_1^0t),$$
$$\frac{\partial Q_1(\boldsymbol{\theta}_1^{0(1)})}{\partial \alpha_1}=-2\sum_{t=1}^{n}t\Big(\sum_{j=2}^{p}(A_j^0\cos(\alpha_j^0t)+B_j^0\sin(\alpha_j^0t))+\sum_{k=1}^{q}(C_k^0\cos(\beta_k^0t^2)+D_k^0\sin(\beta_k^0t^2))+X(t)\Big)\times$$
$$\Big(-A_1^0\sin(\alpha_1^0t)+B_1^0\cos(\alpha_1^0t)\Big).$$

Using Conjecture 2 (see section A), it can be shown that:

$$\boldsymbol{Q}_1'(\boldsymbol{\theta}_1^{0(1)})\boldsymbol{D}_1\overset{a.eq.}{=}-2\begin{pmatrix}\frac{1}{\sqrt{n}}\sum_{t=1}^{n}X(t)\cos(\alpha_1^0t)\\ \frac{1}{\sqrt{n}}\sum_{t=1}^{n}X(t)\sin(\alpha_1^0t)\\ \frac{1}{n\sqrt{n}}\sum_{t=1}^{n}tX(t)(-A_1^0\sin(\alpha_1^0t)+B_1^0\cos(\alpha_1^0t))\end{pmatrix}.$$

Now using CLT, we have:

$$\boldsymbol{Q}_1'(\boldsymbol{\theta}_1^{0(1)})\boldsymbol{D}_1\to N_3(0,4\sigma^2c\boldsymbol{\Sigma}_1^{(1)})$$

Next, we compute the elements of the second derivative matrix, $\boldsymbol{D}_1\boldsymbol{Q}_1''(\boldsymbol{\theta}_1^{0(1)})\boldsymbol{D}_1$. By straightforward calculations and using lemmas 1, 2, 3 and 4, it is easy to show that:

$$\boldsymbol{D}_1\boldsymbol{Q}_1''(\boldsymbol{\theta}_1^{0(1)})\boldsymbol{D}_1=2\boldsymbol{\Sigma}_1^{(1)}.$$

Thus, we have the desired result.

(b) Consider the error sum of squares $Q_2(\boldsymbol{\theta}^{(2)})=\sum_{t=1}^{n}\Big(y_1(t)-C\cos(\beta t^2)-D\sin(\beta t^2)\Big)^2$. Here $y_1(t)=y(t)-\tilde{A}\cos(\tilde{\alpha}t)-\tilde{B}\sin(\tilde{\alpha}t)$, $t=1,\cdots,n$. Let $\boldsymbol{Q}_2'(\boldsymbol{\theta}^{(2)})$ be the first derivative vector and $\boldsymbol{Q}_2''(\boldsymbol{\theta}^{(2)})$, the second derivative matrix of $Q_2(\boldsymbol{\theta}^{(2)})$. By Taylor series expansion of $\boldsymbol{Q}_2'(\tilde{\boldsymbol{\theta}}_1^{(2)})$ around the point $\boldsymbol{\theta}_1^{0(2)}$, we have:

$$(\tilde{\boldsymbol{\theta}}_1^{(2)}-\boldsymbol{\theta}_1^{0(2)})=-\boldsymbol{Q}_2'(\boldsymbol{\theta}_1^{0(2)})[\boldsymbol{Q}_2''(\bar{\boldsymbol{\theta}}_1^{(2)})]^{-1}$$

Multiplying both sides by the matrix $\boldsymbol{D}_2^{-1}$, where $\boldsymbol{D}_2 = diag(\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}, \frac{1}{n^2\sqrt{n}})$, we get:

$$(\tilde{\boldsymbol{\theta}}_1^{(2)} - \boldsymbol{\theta}_1^{0(2)})\boldsymbol{D}_2^{-1} = -\boldsymbol{Q}_2'(\boldsymbol{\theta}_1^{0(2)})\boldsymbol{D}_2[\boldsymbol{D}_2\boldsymbol{Q}_2''(\bar{\boldsymbol{\theta}}_1^{(2)})\boldsymbol{D}_2]^{-1}$$

Now using (33), and proceeding exactly as in part *(a)*, we get:

$$(\tilde{\boldsymbol{\theta}}_1^{(2)} - \boldsymbol{\theta}_1^{0(2)})\boldsymbol{D}_2^{-1} \xrightarrow{d} N_3(0, \sigma^2 c\boldsymbol{\Sigma}_1^{(2)^{-1}}).$$

Hence, the result.

□

Data Availability Statement

The datasets analyzed during the current study are available from the corresponding author on reasonable request.

References

1. Abatzoglou, T. J., 1986 Fast maximnurm likelihood joint estimation of frequency and frequency rate. IEEE Transactions on Aerospace and Electronic Systems, 6, pp. 708-715.
2. Bello, P., 1960 Joint estimation of delay, Doppler, and Doppler rate. IRE Transactions on Information Theory, Vol. 6, Issue 3, pp. 330-341.
3. Brent, R.P., 1973. Algorithms for Minimization without Derivatives, chap. 4.
4. Casazza, P.G. and Fickus, M., 2006. Fourier transforms of finite chirps. EURASIP Journal on Applied Signal Processing, Vol. 2006, Article ID 70204, pp. 1–7.
5. Christensen, M.G., Stoica, P., Jakobsson, A. and Jensen, S.H., 2008. Multi-pitch estimation. Signal Processing, 88(4), pp.972-983.
6. Djuric, P. M., and Kay, S. M., 1990 Parameter estimation of chirp signals. IEEE Transactions on Acoustics, Speech, and Signal Processing, 38(12), pp. 2118-2126.
7. Fuller, W.A., 2009. Introduction to statistical time series (Vol. 428). John Wiley & Sons.
8. Flandrin, P., 1988. Time-frequency processing of bat sonar signals. In Animal Sonar (pp. 797-802). Springer, Boston, MA.
9. Flandrin, P., 2001, March. Time frequency and chirps. In Wavelet Applications VIII. International Society for Optics and Photonics. Vol. 4391, pp. 161-176.
10. Gholami, S., Mahmoudi, A. and Farshidi, E., 2019. Two-stage estimator for frequency rate and initial frequency in LFM signal using linear prediction approach. Circuits, Systems, and Signal Processing, 38(1), pp.105-117.
11. Gini, F., Luise, M. and Reggiannini, R., 1998. Cramer-Rao bounds in the parametric estimation of fading radiotransmission channels. IEEE Transactions on communications, 46(10), pp.1390-1398.
12. Grover, R., Kundu, D. and Mitra, A., 2018, September. Chirp-like model and its parameters estimation. In 2018 International Conference on Computing, Power and Communication Technologies (GUCON) (pp. 1095-1100). IEEE.
13. Hassani, H. and Thomakos, D., 2010. A review on singular spectrum analysis for economic and financial time series. Statistics and its Interface, 3(3), pp.377-397.
14. Ikram, M. Z., Abed-Meraim, K. and Hua, Y., 1997 Fast quadratic phase transform for estimating the parameters of multi-component chirp signals. Digital Signal Processing, 7(2), pp. 127-135.
15. Jones, D.L. and Baraniuk, R.G., 1995. An adaptive optimal-kernel time-frequency representation. IEEE Transactions on Signal Processing, 43(10), pp.2361-2371.
16. Kay, S.M. and Marple, S.L., 1981. Spectrum analysis− a modern perspective. Proceedings of the IEEE, Vol. 69 , Issue 11, pp. 1380-1419.
17. Kelly, E. J., 1961 The radar measurement of range, velocity and acceleration. IRE Transactions on Military Electronics, Vol. 1051, Issue 2, pp. 51-57.
18. Kundu, D. and Nandi, S., 2008 Parameter estimation of chirp signals in presence of stationary noise. Statistica Sinica, pp. 187-201.
19. Kundu, D. and Nandi, S., 2012 Statistical Signal Processing: Frequency Estimation. New Delhi.
20. Lahiri, A. 2011. Estimators of Parameters of Chirp Signals and Their Properties. PhD thesis, Indian Institute of Technology, Kanpur.
21. Lahiri, A., Kundu, D. and Mitra, A., 2014 On least absolute deviation estimators for one-dimensional chirp model. Statistics, 48(2), pp. 405-420.
22. Lahiri, A., Kundu, D. and Mitra, A., 2015 Estimating the parameters of multiple chirp signals. Journal of Multivariate Analysis, 139, pp. 189-206.
23. Li, L. and Qiu, T., 2019. A robust parameter estimation of LFM signal based on sigmoid transform under the alpha stable distribution noise. Circuits, Systems, and Signal Processing, 38(7), pp.3170-3186.
24. Legendre, A.M., 1805. Nouvelles méthodes pour la détermination des orbites des cométes. F. Didot.
25. Ma, N. and Vray, D., 1998, May. Bottom backscattering coefficient estimation from wideband chirp sonar echoes by chirp adapted time-frequency representation. In Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP'98 (Cat. No. 98CH36181), Vol. 4, pp. 2461-2464. IEEE.
26. Mazumder, S., 2017 Single-step and multiple-step forecasting in one-dimensional single chirp signal using MCMC-based Bayesian analysis. Communications in Statistics-Simulation and Computation, 46(4), pp. 2529-2547.
27. Mboup, M. and Adal, T., 2012. A generalization of the Fourier transform and its application to spectral analysis of chirp-like signals. Applied and Computational Harmonic Analysis, Vol. 32, Issue 2, pp. 305-312.
28. McAulay, R. and Quatieri, T., 1986. Speech analysis/synthesis based on a sinusoidal representation. IEEE Transactions on Acoustics, Speech, and Signal Processing, 34(4), pp.744-754.
29. Montgomery, H.L., 1994. Ten lectures on the interface between analytic number theory and harmonic analysis (No. 84). American Mathematical Society.
30. Murthy, V.K., Haywood, L.J., Richardson, J., Kalaba, R., Salzberg, S., Harvey, G. and Vereeke, D., 1971. Analysis of power spectral densities of electrocardiograms. Mathematical Biosciences, 12(1-2), pp.41-51.

31. Nandi, S. and Kundu, D., 2004 Asymptotic properties of the least squares estimators of the parameters of the chirp signals. Annals of the Institute of Statistical Mathematics, 56(3), pp. 529-544.
32. Neuberg, J., Luckett, R., Baptie, B. and Olsen, K., 2000. Models of tremor and low-frequency earthquake swarms on Montserrat. Journal of Volcanology and Geothermal Research, 101(1-2), pp.83-104.
33. Peleg, S. and Porat, B., 1991 Linear FM signal parameter estimation from discrete-time observations. IEEE Transactions on Aerospace and Electronic Systems, 27(4), pp. 607-616.
34. Prasad, S., Chakraborty, M. and Parthasarathy, H., 1995. The role of statistics in signal processing- 'a brief review and some emerging trends. Indian Journal of Pure and Applied Mathematics, Vol. 26, pp. 547-578.
35. Prasad, A., Kundu, D. and Mitra, A., 2008 Sequential estimation of the sum of sinusoidal model parameters. Journal of Statistical Planning and Inference, 138(5), pp. 1297-1313.
36. Rice, J. A. and Rosenblatt, M., 1988 On frequency estimation. Biometrika, 75(3), pp. 477-484.
37. Richards, F. SG., 1962 A method of maximum-likelihood estimation. Journal of the Royal Statistical Society. Series B (Methodological), pp. 469-475.
38. Saha, S. and Kay, S. M., 2002 Maximum likelihood parameter estimation of superimposed chirps using Monte Carlo importance sampling. IEEE Transactions on Signal Processing, 50(2), pp. 224-230.
39. Song, J., Xu, Y., Liu, Y. and Zhang, Y., 2019. 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(12), pp.5861-5882.
40. Stoica, P., 1993. List of references on spectral line analysis. Signal Processing, Vol. 31, Issue 3, pp. 329-340.
41. VanderPlas, J.T. and Ivezic, Ž., 2015. Periodograms for multiband astronomical time series. The Astrophysical Journal, 812(1), p.18.
42. Wang, G. and Xia, X.G., 2000, May. An adaptive filtering approach to chirp estimation and ISAR imaging of maneuvering targets. In Record of the IEEE 2000 International Radar Conference [Cat. No. 00CH37037], pp. 481-486. IEEE.