IET Research Journals

Submission for IET Research Journal Paper

Estimation of Parameters of a Harmonic Chirp Model

ISSN 1751-8644
doi: 0000000000
www.ietdl.org

Rhythm Grover[1*], *Debasis Kundu*[1], *Amit Mitra*[1]

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

Abstract: In this paper, we address the problem of estimation of the parameters of a harmonic chirp model, often encountered in speech and music applications. This model was introduced recently by Christensen and Jensen [28] as an extension of a standard harmonic model. We propose two methods of estimation: the least squares estimation method and the approximate least squares estimation method. We establish the asymptotic properties of the least squares estimators as well as the approximate least squares estimators of the parameters of this model under the assumption of stationary errors. These asymptotic properties are proved analytically as well as corroborated through simulation experiments. We present two speech signal data sets and their analysis using both the estimation methods. The results show that the proposed methods perform reasonably well for estimating the unknown model parameters.

1 Introduction

Parameter estimation of a sinusoidal model is a long-established problem and has been given considerable importance in the literature. Mathematically, a sum of sinusoidal model can be described as:

$$y(t) = \sum_{k=1}^{p} A_k^0 \cos(\alpha_k^0 t) + B_k^0 \sin(\alpha_k^0 t) + X(t);\ t = 1, \ldots, n.$$

Here, the parameters A_k^0s, B_k^0s are the amplitudes, α_k^0s are the frequencies of the observed signal $y(t)$ and $X(t)$ is the additive random error component with mean zero. The explicit assumptions on $X(t)$ will be mentioned later. Many real-life periodic and even nearly periodic phenomena can be explained using this model and because of its increasingly ubiquitous applicability, it has been used extensively in the signal processing literature. For references, one may look into the works of Kay and Marple [5], Stoica [10], Prasad, Chakraborty and Parthasarathy [12] and Kundu and Nandi [27].

A special case of the sinusoidal model is a harmonic model, where the frequencies instead of being arbitrary, are integral multiples of a constant fundamental frequency. Following is the analytical expression of a harmonic model:

$$y(t) = \sum_{k=1}^{p} A_k^0 \cos(k\alpha^0 t) + B_k^0 \sin(k\alpha^0 t) + X(t)\ t = 1, \ldots, n, \tag{1}$$

where as before A_k^0s, B_k^0s are the amplitudes of the signal $y(t)$ and $X(t)$ is the additive error component. However, the frequencies of the sinusoids are multiples of a common, constant fundamental frequency, α^0. This model has several real life applications, particularly in analysing signals where exact periodicity is observed, for instance in the fields of speech processing, music processing and communications, see for example, Stylianou [15], Christensen, Stoica, Jakobsson and Jensen [23], Fletcher and Rossing [26] and the references cited therein. For references on some of the classical methods of parameter estimation of this model, one may see the works of Quinn and Thompson [9], Li, Stoica and Li [13], Quinn and Hannan [14], Nandi and Kundu [20] and Chan and So [18].

This paper deals with the parameter estimation of a harmonic chirp model which is a natural generalisation of the harmonic model (1) to take into account the time-variations in frequencies of the observed signals. Mathematically, the model is represented as:

$$y(t) = \sum_{k=1}^{p} A_k^0 \cos(k\alpha^0 t + k\beta^0 t^2) + B_k^0 \sin(k\alpha^0 t + k\beta^0 t^2) + X(t);\ t = 1, \ldots, n, \tag{2}$$

where the unknown frequencies are the harmonics of a fundamental frequency α^0 and the frequency rates are the harmonics of a fundamental chirp rate β^0 and the parameters A_k^0s and B_k^0s are the amplitudes of the observed signal $y(t)$. The error term X(t) is a stationary linear process, which is a more realistic assumption than the usual Gaussian noise assumption. Popular noise models such as AR, MA or ARMA models can be represented as a stationary linear process under certain conditions and hence this assumption covers a large class of random processes.

This model has been suggested recently by Christensen and Jensen [28] and thereafter some authors, for instance, Doweck, Amar and Cohen [30, 32], Nørholm, Jensen and Christensen [33], Sward, Brynolfsson, Jakobsson and Hansson-Sandsten [34], Jensen, Nielsen, Jensen, Christensen and Jensen [35] and Nielsen, Jensen, Jensen, Christensen and Jensen [36] have considered its parameter estimation. In real life applications, this model captures the deterministic component of the signals more accurately as it takes into account the modulations in the frequency of the observed signal. Apart from being an extension of a harmonic model, a harmonic chirp model can also be seen as a special case of a chirp model, where different harmonics have arbitrary frequencies and modulations. The latter is one of the fundamental models in signal processing literature and its parameter estimation has received considerable attention over the years. For references, see Abatzoglou [1], Djuric and Kay [7], Peleg and Porat [8], Saha and Kay [16], Myburg, Den Brinker and Van Eijndhoven [19], Weruaga and Kepesi [21] and Pantazis, Rosec, and Stylianou [25].

A brief review of the recent work on the parameter estimation of the harmonic chirp model: Doweck, Amar and Cohen [30] considered the joint estimation of the initial frequency and frequency rate of a complex harmonic chirp model under the assumption of complex Gaussian errors. They proposed the maximum likelihood estimation (MLE) method, the approximate maximum likelihood estimation method and also an alternative method to reduce the computational complexity involved in the first two methods. This alternative method is a two step estimation method, the first step is to separate the signal to its harmonic components and the second is to estimate the parameters using least squares method given the phases of the harmonic components. Doweck, Amar and Cohen [32] proposed another two methods of estimation. The first method is based on high-order ambiguity function (HAF) which reduces the two dimensional optimisation problems into two one-dimensional optimisation problems. The second method is a low-complexity method based on separate-estimate approach that they call Harmonic-SEES method. Through numerical experiments, they demostrate that the Harmonic-SEES method outperforms the one based on HAF approach. Nørholm, Jensen and Christensen [33] also consider the complex harmonic chirp model with additive complex Gaussian errors and proposed an iterative MLE method. They compared the performance of this model with the harmonic model based on the maximum a posteriori (MAP) model selection criterion for long segments. Sward, Brynolfsson, Jakobsson and Hansson-Sandsten [34] consider the same model and provide an iterative sparse reconstruction framework, which is an extension of the work done by Adalbjörnsson, Jakobsson and Christensen [29] on the harmonic model. Jensen, Nielsen, Jensen, Christensen and Jensen [35] consider a real-valued harmonic chirp model and provide an algorithm for computing an estimate to the ML estimator with much lower computational complexity using recursive matrix structures, the fast Fourier transform, and using a two-step approach that reduces the number of required function evaluations. Nielsen, Jensen, Jensen, Christensen and Jensen [36] proposed an accurate and robust fast method for harmonic chirp summation. In all the aforementioned works, it is shown that the proposed estimators achieve the Cramer Rao lower bound (CRLB), however there has been no study of the asymptotic properties of the proposed estimators.

The aim of this paper is two-fold. First, we consider a more general error structure, that is, we assume that $X(t)$s are from a stationary linear process. Because of this assumption, the model becomes more realistic as it takes into account the dependence structure that may be present in the errors in many practical situations. The explicit details of this assumption are provided in the next section. The second aim is to establish large sample properties of the studied estimators, namely consistency and asymptotic normality under the stationary assumptions on the error component.

Here, we mainly address the problem of parameter estimation of the model (2) using the least squares estimation method. These estimators are obtained by minimising the error sum of squares defined in (3). Note that the model under consideration is a non-linear time series regression model. Due to the highly non-linear nature of the model, apart from computational difficulty involved in finding the LSEs, one cannot study the finite sample statistical properties of these estimators as well. Thus, to assess the performance of the LSEs theoretically, one has to resort to their asymptotic behaviour under certain suitable assumptions. Moreover, deriving the asymptotic properties of these estimators is not quite straightforward as they do not satisfy the well-established results in literature, for instance, the sufficient conditions of consistency provided by Jennrich [3] or Wu [6]. Thus, the primary focus of this paper is to study the asymptotic properties of the LSEs under the assumption of stationary errors. These properties are proved analytically as well as corroborated through extensive simulation studies. It is also observed that under the special case of Gaussian errors, the asymptotic variances of the LSEs attain the asymptotic CRLBs.

We also discuss briefly the approximate least squares estimation method for the joint estimation of the fundamental frequency and the fundamental chirp rate. This method was first proposed by Walker [4] for estimating the parameters of a sinusoidal model. He obtained the asymptotic properties of approximate least squares estimators (ALSEs) under the assumption of independently and identically distributed (i.i.d.) errors with mean 0 and finite variance. Furthermore, Kundu and Nandi [17] and Grover, Kundu and Mitra [37, 38] extended this method to estimate the parameters of a 2-D sinusoidal model and a 1-D chirp model and 2-D chirp model, respectively. Under the assumption of stationary errors, they observed that the ALSEs of the unknown parameters have the same theoretical large sample properties as the corresponding LSEs. Moreover, the ALSEs have two remarkable features− firstly, they are computationally faster as compared to the LSEs and secondly, their consistency is obtained under weaker assumptions on the linear parameters than those required for the LSEs. Thus, it is indeed intriguing to investigate the properties of the ALSEs of the unknown parameters of the harmonic chirp model as well and see how they perform computationally in comparison with the LSEs. We observe that although theoretically the ALSEs have all the desirable properties of the LSEs, computationally the LSEs perform better than the ALSEs. This performance is compared in terms of the average estimates and the order of the biases and the variances of the estimates.

In the next section, we shall state the necessary model assumptions and develop statistical properties of the LSEs of the unknown parameters under these assumptions. In Section 3, we discuss the methodology to obtain the ALSEs and establish their statistical properties. We provide simulation results in Section 4 and in Section 5, we present two speech signal data sets and their analyses using the least squares estimation method as well as the approximate least squares estimation method. We conclude the paper in Section 6. All the proofs are provided in the appendices.

2 Least Squares Estimation Method

The least squares estimators of the parameters can be obtained by minimising the following error sum of squares:

$$Q(\boldsymbol{\theta}) = \sum_{t=1}^{n}\left(y(t) - \sum_{k=1}^{p}\{A_k\cos(k\alpha t + k\beta t^2) + B_k\sin(k\alpha t + k\beta t^2)\}\right)^2, \tag{3}$$

with respect to $\boldsymbol{\theta}$, where $\boldsymbol{\theta} = (A_1, B_1, \ldots, A_p, B_p, \alpha, \beta)$ is the parameter vector of the model. Henceforth, for brevity, we use matrix notation so that (3) can be rewritten as:

$$Q(\boldsymbol{\theta}) = (\mathbf{Y} - \mathbf{Z}(\alpha,\beta)\boldsymbol{\phi})^\top(\mathbf{Y} - \mathbf{Z}(\alpha,\beta)\boldsymbol{\phi}).$$

Here, $\mathbf{Y} = \begin{bmatrix} y(1) & \cdots & y(n)\end{bmatrix}^\top$ is the observed data vector, $\boldsymbol{\phi} = \begin{bmatrix} A_1 & B_1 & \cdots & A_p & B_p\end{bmatrix}^\top$ is the vector of linear parameters and the matrix $\mathbf{Z}(\alpha,\beta)$ is defined as follows:

$$\mathbf{Z}(\alpha,\beta) = \begin{bmatrix} \cos(\alpha+\beta) & \sin(\alpha+\beta) & \ldots & \cos(p\alpha+p\beta) & \sin(p\alpha+p\beta) \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ \cos(n\alpha+n^2\beta) & \sin(n\alpha+n^2\beta) & \ldots & \cos(pn\alpha+pn^2\beta) & \sin(pn\alpha+pn^2\beta) \end{bmatrix}. \tag{4}$$

Since $\boldsymbol{\phi}$ is a vector of linear parameters, for a given α and β they can be estimated using simple linear regression by the separable regression technique of Richards [2] as follows:

$$\widehat{\boldsymbol{\phi}}(\alpha,\beta) = [\mathbf{Z}(\alpha,\beta)^\top\mathbf{Z}(\alpha,\beta)]^{-1}\mathbf{Z}(\alpha,\beta)^\top\mathbf{Y}. \tag{5}$$

The above equation allows one to reduce the error sum of squares function $Q(\boldsymbol{\theta})$ as a function of only the non-linear parameters α and β. Thus, the reduced error sum of squares function can be written as:

$$\begin{aligned} R(\alpha,\beta) &= (\mathbf{Y} - \mathbf{Z}(\alpha,\beta)\widehat{\boldsymbol{\phi}})^\top(\mathbf{Y} - \mathbf{Z}(\alpha,\beta)\widehat{\boldsymbol{\phi}}) \\ &= \mathbf{Y}^\top(\mathbf{I} - \mathbf{P}_{\mathbf{Z}}(\alpha,\beta))\mathbf{Y}, \end{aligned} \tag{6}$$

where $\mathbf{P}_{\mathbf{Z}}(\alpha,\beta) = \mathbf{Z}(\alpha,\beta)[\mathbf{Z}(\alpha,\beta)^\top\mathbf{Z}(\alpha,\beta)]^{-1}\mathbf{Z}(\alpha,\beta)^\top$ is the projection matrix on the column space of the matrix $\mathbf{Z}(\alpha,\beta)$. The LSEs of α^0 and β^0 are obtained by minimizing $R(\alpha,\beta)$ with respect to α and β simultaneously and once we have the estimators, $\hat{\alpha}$ and $\hat{\beta}$, we can substitute them in (5) to get the linear parameter estimates.

In the subsequent subsections, we examine the asymptotic properties of these estimators under the following assumptions:

Assumption 1. *The error random sequence* $\{X(t)\}$ *is a stationary linear process of the form:*

$$X(t) = \sum_{j=-\infty}^{\infty} a(j)e(t-j). \tag{7}$$

Here $\{e(t); t \in \{0, \pm 1, \pm 2, \ldots\}\}$ *is a sequence of i.i.d. random variables with* $E(e(t)) = 0$, $V(e(t)) = \sigma^2$ *and finite fourth moment and* $a(j)$*s are real constants such that*

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

Equation (7) is a standard representation of a stationary linear process and covers a large class of stationary random processes.

Assumption 2. *The true parameter vector* $\boldsymbol{\theta^0}$ *is an interior point of* $\boldsymbol{\Theta}^{(1)} = [-K, K]^{2p} \times [0, \frac{\pi}{p}] \times [0, \frac{\pi}{p}]$. *Here,* $K > 0$ *is a real number.*

Theorem 1. *Under the assumptions* 1 *and* 2, $\hat{\boldsymbol{\theta}}$, *the LSE of* $\boldsymbol{\theta}^0$ *is a strongly consistent estimator of* $\boldsymbol{\theta}^0$, *that is,*

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

Proof: See Appendix 8.1.

□

Theorem 2. *Under the assumptions* 1 *and* 2, *if the number of components of the fitted model is more than the actual number of components, p, then for* $k > p$, $\hat{A}_k \to 0$ *a.s. and* $\hat{B}_k \to 0$ *a.s.*

Proof: See Appendix 8.1.

□

Theorem 3. *Under the assumptions* 1 *and* 2,

$$(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0)\boldsymbol{D}^{-1} \to \mathcal{N}_{2p+2}(0, 2c\sigma^2\boldsymbol{\Sigma}) \text{ as } n \to \infty.$$

Here, $\boldsymbol{D} = \text{diag}(\underbrace{\frac{1}{\sqrt{n}}, \ldots, \frac{1}{\sqrt{n}}}_{2p\ times}, \frac{1}{n\sqrt{n}}, \frac{1}{n^2\sqrt{n}})$,

$c = \sum\limits_{j=-\infty}^{\infty} a^2(j)$ *and the matrix*

$$\boldsymbol{\Sigma}^{-1} = \begin{bmatrix} \boldsymbol{I}_{2p} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22}. \end{bmatrix}.$$

Here, $\boldsymbol{I}_{2p}$ *is the identity matrix of order* $2p$,

$$\boldsymbol{\Sigma}_{12} = \begin{bmatrix} \frac{B_1^0}{2} & \cdots & p\frac{B_p^0}{2} & -\frac{A_1^0}{2} & \cdots & -p\frac{A_p^0}{2} \\ \frac{B_1^0}{3} & \cdots & p\frac{B_p^0}{3} & -\frac{A_1^0}{3} & \cdots & -p\frac{A_p^0}{3} \end{bmatrix}^\top,$$

$$\boldsymbol{\Sigma}_{21} = \boldsymbol{\Sigma}_{12}^\top$$

and

$$\boldsymbol{\Sigma}_{22} = \begin{bmatrix} \frac{\sum_{k=1}^{p} k^2({A_k^0}^2 + {B_k^0}^2)}{3} & \frac{\sum_{k=1}^{p} k^2({A_k^0}^2 + {B_k^0}^2)}{4} \\ \frac{\sum_{k=1}^{p} k^2({A_k^0}^2 + {B_k^0}^2)}{4} & \frac{\sum_{k=1}^{p} k^2({A_k^0}^2 + {B_k^0}^2)}{5} \end{bmatrix}.$$

Proof: See Appendix 8.2.

□

3 Approximate Least Squares Estimation Method

In this section, we consider another method of parameters estimation and we call it as the approximate least squares estimation method. This method is based on the idea of the periodogram estimators, one of the most popular methods of estimation of the frequencies of a sinusoidal model.

First we define the periodogram-type function and explain how to obtain the ALSEs and then we study different asymptotic properties of these estimators under the assumption of stationary errors.

Using the following lemma proved by Lahiri [31]:

Lemma 1. *For* $\omega_1, \omega_2, \omega_1', \omega_2' \in (0, \pi)$*:*

1. $\lim\limits_{n\to\infty} \frac{1}{n}\sum\limits_{t=1}^{n} \cos(\omega_1 t + \omega_2 t^2) = \lim\limits_{n\to\infty} \frac{1}{n}\sum\limits_{t=1}^{n} \sin(\omega_1 t + \omega_2 t^2) = 0$
2. $\lim\limits_{n\to\infty} \frac{1}{n^{k+1}}\sum\limits_{t=1}^{n} t^k \cos^2(\omega_1 t + \omega_2 t^2) = \frac{1}{2(k+1)}$
3. $\lim\limits_{n\to\infty} \frac{1}{n^{k+1}}\sum\limits_{t=1}^{n} t^k \sin^2(\omega_1 t + \omega_2 t^2) = \frac{1}{2(k+1)}$
4. $\lim\limits_{n\to\infty} \frac{1}{n^{k+1}}\sum\limits_{t=1}^{n} t^k \sin(\omega_1 t + \omega_2 t^2)\cos(\omega_1 t + \omega_2 t^2) = 0$

For $(\omega_1, \omega_2) \neq (\omega_1', \omega_2')$,

1. $\lim\limits_{n\to\infty} \frac{1}{n^{k+1}}\sum\limits_{t=1}^{n} t^k \sin(\omega_1 t + \omega_2 t^2)\sin(\omega_1' t + \omega_2' t^2) = 0$
2. $\lim\limits_{n\to\infty} \frac{1}{n^{k+1}}\sum\limits_{t=1}^{n} t^k \cos(\omega_1 t + \omega_2 t^2)\cos(\omega_1' t + \omega_2' t^2) = 0$
3. $\lim\limits_{n\to\infty} \frac{1}{n^{k+1}}\sum\limits_{t=1}^{n} t^k \sin(\omega_1 t + \omega_2 t^2)\cos(\omega_1' t + \omega_2' t^2) = 0$

equation (6) can be written as:

$$R(\alpha, \beta) = \mathbf{Y}^\top\mathbf{Y} - 2I(\alpha, \beta) + o(1)$$

where,

$$I(\alpha, \beta) = \frac{1}{n}\mathbf{Y}^\top\mathbf{Z}(\alpha, \beta)\mathbf{Z}(\alpha, \beta)^\top\mathbf{Y}. \tag{8}$$

Since the difference between the reduced error sum of squares function, $R(\alpha, \beta)$, and the periodogram-type function $I(\alpha, \beta)$ is $O(1)$ as $n \to \infty$, minimising the former is equivalent to maximising the latter. This result is an extension of the one obtained by Walker [4] for the sinusoidal model. Here, the function $f = o(1)$ means $f \to 0$ a.s. and $f = O(1)$ means $f \leqslant c_0$ a.s., where c_0 is some finite constant. The estimators obtained by maximising $I(\alpha, \beta)$ with respect to α and β simultaneously are called the ALSEs of α^0 and β^0. The linear parameter estimates can be obtained by plugging-in the non-linear parameter estimates and using Lemma 1 in (5).

Under the same assumption on the error component $X(t)$ (see Assumption 1) and the following assumption on the parameters:

Assumption 3. *The true parameter vector,* $\boldsymbol{\theta^0}$ *is an interior point of* $\boldsymbol{\Theta}^{(2)} = (-\infty, \infty)^{2p} \times [0, \frac{\pi}{p}] \times [0, \frac{\pi}{p}]$,

we examine the asymptotic properties of the ALSEs and the obtained results are provided below.

Theorem 4. *Under the assumptions* 1 *and* 3, $\tilde{\boldsymbol{\theta}}$, *the ALSE of* $\boldsymbol{\theta}^0$ *is a strongly consistent estimator of* $\boldsymbol{\theta}^0$, *that is,*

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

Proof: See Appendix 8.3.

□

Theorem 5. *Under the assumptions* 1 *and* 2*, if the number of components of the fitted model is more than the actual number of components, p, then for* $k > p$*,* $\tilde{A}_k \to 0$ *a.s. and* $\tilde{B}_k \to 0$ *a.s.*

Proof: This proof follows from the proof of Theorem 2. □

Theorem 6. *Under the assumptions* 1 *and* 3*,*

$$(\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}^0)\boldsymbol{D}^{-1} \stackrel{d}{=} (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0)\boldsymbol{D}^{-1} \text{ as } n \to \infty.$$

Here the diagonal matrix, ***D*** *and* c *are same as defined in Theorem* 3.

Proof: See Appendix 8.4. □

4 Simulation Studies

We conduct extensive simulation studies for the joint estimation of the fundamental frequency and fundamental chirp rate using the suggested estimation methods. Here, we report the results obtained for the following simulated harmonic chirp model:

$$y(t) = \sum_{k=1}^{p} A_k^0 \cos(k\alpha^0 t + k\beta^0 t^2) + B_k^0 \sin(k\alpha^0 t + k\beta^0 t^2) + X(t);\ t = 1, \ldots, n,$$

with the following true values of the amplitudes:

$$A_1^0 = 3.5,\ A_2^0 = 3.2,\ A_3^0 = 2.5,\ A_4^0 = 1.75,$$
$$B_1^0 = 2.8,\ B_2^0 = 2.5,\ B_3^0 = 1.5,\ B_4^0 = 1.$$

We consider two different error structures for data generation:

$$X(t) = \epsilon(t) + 0.8\epsilon(t-1),$$
$$X(t) = 0.8X(t-1) + \epsilon(t),$$

where $\epsilon(t) \sim \mathcal{N}(0, \sigma^2)$. For the first set of simulation experiments, we choose $\alpha = 0.03$, $\beta = 0.003$, $\sigma = 0.5$ and consider different sample lengths, $n = 125, 250, 500, 1000$ and 2000. In the second set of experiments, we fix the sample length at $n = 500$ and with the same values of α and β as in the previous set of simulations, we vary standard deviation, $\sigma = 0.01, 0.1, 0.5, 1$ and 2. Next, we assess the performance of the estimators for different values of $\alpha = 0.01, 0.02, 0.03, 0.04$ and 0.05. In these set of numericals, we choose $\beta = 0.003$, $\sigma = 0.5$ and $n = 500$. And lastly, for fixed values of $\alpha = 0.03, \sigma = 0.5$ and $n = 500$, we perform the experiments by varying $\beta = 0.001, 0.002, 0.003, 0.004$ and 0.005. The number of runs for each set of experiments is chosen to be 1000. Apart from the above set of experiments, we evaluated the performance of the proposed estimators on the same model with several other pre-fixed values of the parameters. These results are reported in the form of root mean square errors (RMSEs). Due to space constraint, the results are reported for some of them, however it is important to note that for different set of experiments, the experimental results are consistent with those reported here and are at par with the theoretically derived results.

Figures 1, 2, 3, 4 and 5 show the effect of sample length on the performance of the LSEs as well as the ALSEs of the parameters of the model. It can be seen that as n increases, the RMSEs of the estimates decrease, thus validating the consistency property of the estimators. The RMSEs of the LSEs match quite well with the square root of theoretical asymptotic variances for both linear and non-linear parameters. The RMSEs of ALSEs of the non-linear parameters are reasonably close to those of LSEs and hence to the corresponding asymptotic standard deviations. However, for the linear parameters, the performance of ALSEs is unstable.

In Figures 6, 7, 8, 9 and 10, we assess the performance of the estimates with varying signal-to-noise ratio (SNR) on the x-axis. As seen from the plots, the RMSEs of both the estimates decrease as the SNR increases, which is expected. Also, the RMSEs of the LSEs are found to be in good agreement with the asymptotic results. It is also observed that the RMSEs of the ALSEs do not change significantly with increasing SNR, particularly for the linear parameters.

Finally, in Figures 11, 13, 14, 15, 16 and 17, 18, 19, 20, 21, we evaluate the performance of the estimators with respect to varying α and β on the x-axis, respectively. These plots show that the estimation accuracy of the LSEs for different values of α and β is quite balanced and at par with the corresponding asymptotic results. Also, the ALSEs perform well but the behaviour of the ALSEs seems erratic with varying α and β.

It is evident that the LSEs outperforms the ALSEs across all sample sizes, SNR values and the parameter values. Moreover, the RMSEs of LSEs almost coincide with the square root of theoretically derived asymptotic variances which is used here as the benchmark for the estimation accuracy.

5 Data Analysis

In this section, we present the analyses of two real life data sets. These data sets are speech signal data sets, 'u:' and 'a:' (vowel sounds) produced by a sound instrument at the Speech Signal Processing laboratory of Indian Institute of Technology, Kanpur. We fit a harmonic chirp model to these data sets using both the least squares and the approximate least squares estimation methods. To estimate the number of chirps present in the signal, we use the following Bayesian Information Criterion (BIC):

$$BIC(k) = n\ln(\mathrm{SS}_{res}(k)) + 2(2k+2)\ln(n);\ k = 1, \ldots, J.$$

Here, $\mathrm{SS}_{res}(k)$, is the residual sum of squares when there are k components present in the signal, which decreases as k increases and the second term, $2(2k+2)\ln(n)$, is the penalty term, and for a fixed n, it increases as k increases. Assuming that the maximum number of components can be J, we choose the estimate of p corresponding to the minimum value of $BIC(k)$. In Figure 22, we plot the $BIC(k)$ values for $k = 3, \ldots, 12$. The purple solid line in the figure represents the BIC curve for the data set 'u:' when the estimates are obtained using least squares estimation method and the pink dashed line represents the BIC curve for the same data set when approximate least squares estimation method is used. It is evident from the plot, that we choose the estimate of p as 9 when the LSEs are used and 8 when the ALSEs are used.
Similarly, for the second data set 'a:', we plot the BIC curve (see Figure 23) and we choose the estimate of p as 9 in case of the LSEs and 7 in case of the ALSEs. Once we have the estimate of the number of components, we estimate the other parameters of the deterministic component of the model, the amplitudes, the fundamental frequency and the fundamental frequency rate using the LSEs and the ALSEs. Using these estimates, we obtain the fitted values.

In Figure 24, the first subplot represents the original signal 'a:' along with the estimated signals using both the least squares and the approximate least squares estimation method in the time domain. It can be seen that both the fittings estimate the original signal quite accurately, except at certain peaks and depressions. The subplot to its right, shows the amplitude versus frequency plot for the same data set. Although certain low frequencies are not resolved, the estimated models capture the intermediate frequencies very well. The plots in the next row of the figure, represent the corresponding signals for data set 'u:'. It can be seen that the fittings match very well with the original signal in the time domain. Moreover, the estimated models capture the lower order frequencies quite precisely.

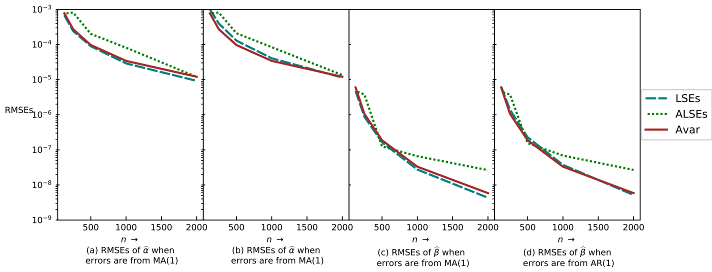

Fig. 1: In each sub-plot the solid line represents the theoretical asymptotic variances of the estimators, the dashed line represents the RMSEs of the LSEs and the dotted line represents the RMSEs of the ALSEs of the non-linear parameters of the underlying simulated multiple-component model versus the sample length n.

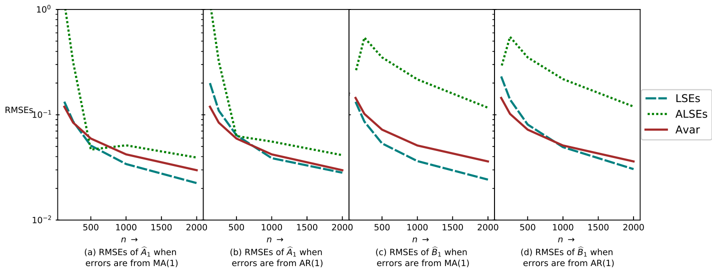

Fig. 2: In each sub-plot the solid line represents the theoretical asymptotic variances of the estimators, the dashed line represents the RMSEs of the LSEs and the dotted line represents the RMSEs of the ALSEs of parameters A_1^0 and B_1^0 of the underlying simulated multiple-component model versus the sample length n.

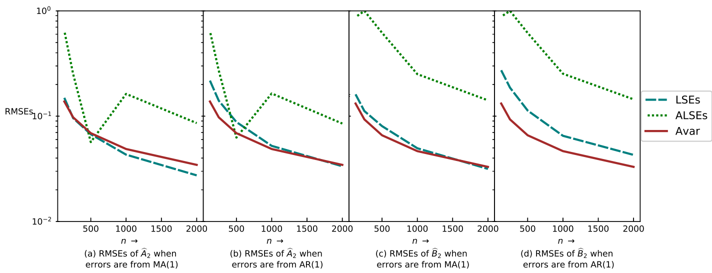

Fig. 3: In each sub-plot the solid line represents the theoretical asymptotic variances of the estimators, the dashed line represents the RMSEs of the LSEs and the dotted line represents the RMSEs of the ALSEs of parameters A_2^0 and B_2^0 of the underlying simulated multiple-component model versus the sample length n.

6 Conclusion

The aim of this paper has been two-fold. First, we considered a more general error structure than the usual Gaussian assumption. Because of this assumption, the model is more realistic as it takes into account the dependence structure that may be present in the errors in many practical situations. The second aim was to establish large-sample properties of the LSEs and the ALSEs under the stationary assumptions on the error component. We observed that both types of estimators are strongly consistent. It is also observed that the order of asymptotic variances of the amplitudes estimators is n^{-1} and that of the fundamental frequency and fundamental chirp rate estimators are n^{-3} and n^{-5} respectively. In fact, both the LSEs and the ALSEs have the same asymptotic distribution. Moreover, under the special case of Gaussian errors, the asymptotic variances of the proposed estimators attain the CRLB. By means of numerical simulations, we evaluate the performance of LSEs and ALSEs and observe that LSEs outperform ALSEs. We have also analysed two speech signal data sets to exemplify the implementation and effectiveness of the two methods to fit a harmonic chirp model in a practical situation. The outcomes of these data analyses are quite encouraging.

There are many other methods for dealing with this problem of parameter estimation. For example, Tsakonas et al. [22] proposed a filtering method for this purpose. However, they assume that the errors are white Gaussian which is usually unrealistic in practice and no theoretical properties such as consistency are established for the proposed method. It will be interesting to extend the method for colored noise and establish theoretical properties of their method.

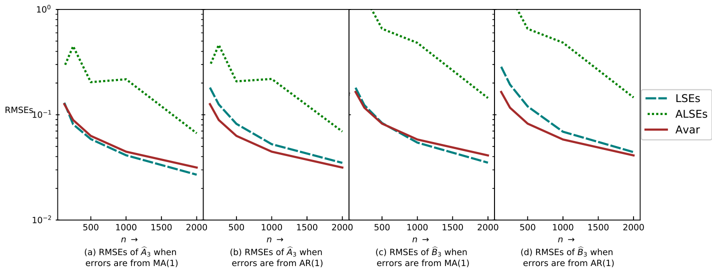

Fig. 4: In each sub-plot the solid line represents the theoretical asymptotic variances of the estimators, the dashed line represents the RMSEs of the LSEs and the dotted line represents the RMSEs of the ALSEs of parameters A_3^0 and B_3^0 of the underlying simulated multiple-component model versus the sample length n.

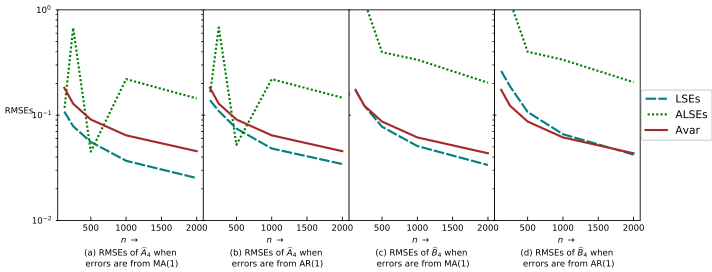

Fig. 5: In each sub-plot the solid line represents the theoretical asymptotic variances of the estimators, the dashed line represents the RMSEs of the LSEs and the dotted line represents the RMSEs of the ALSEs of parameters A_4^0 and B_4^0 of the underlying simulated multiple-component model versus the sample length n.

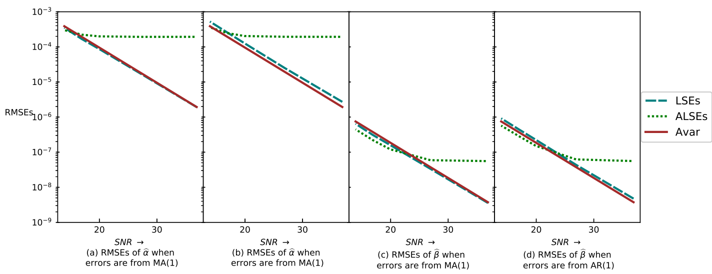

Fig. 6: In each sub-plot the solid line represents the theoretical asymptotic variances of the estimators, the dashed line represents the RMSEs of the LSEs and the dotted line represents the RMSEs of the ALSEs of the non-linear parameters of the underlying simulated multiple-component model versus SNR.

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.

8 Appendices

8.1 Proof of consistency of the LSEs

To prove Theorem 1, we need the following lemma:

Lemma 2. *Consider the following set:*

$$S_\delta = \{\boldsymbol{\theta} : \boldsymbol{\theta} \in \boldsymbol{\Theta}^{(1)};\ |A_k - A_k^0| > \delta \text{ or } |B_k - B_k^0| > \delta \text{ or } |\alpha - \alpha^0| > \delta \text{ or } |\beta - \beta^0| > \delta \text{ for any } k = 1, \ldots, p\}.$$

If for some $\delta > 0$,

$$\liminf_{n\to\infty} \inf_{\boldsymbol{\theta}} \frac{1}{n}(Q(\boldsymbol{\theta}) - Q(\boldsymbol{\theta}^0)) > 0 \tag{9}$$

then $\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} Q(\boldsymbol{\theta})$, *is a strongly consistent estimator of* $\boldsymbol{\theta}^0$. *The notations* $\liminf$ *and* $\inf$ *denote limit infimum and infimum of a function, respectively.*

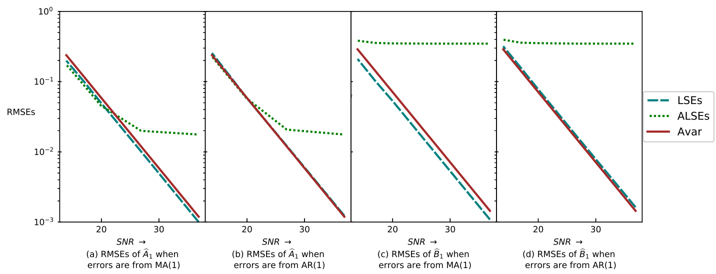

Fig. 7: In each sub-plot the solid line represents the theoretical asymptotic variances of the estimators, the dashed line represents the RMSEs of the LSEs and the dotted line represents the RMSEs of the ALSEs of parameters A_1^0 and B_1^0 of the underlying simulated multiple-component model versus SNR.

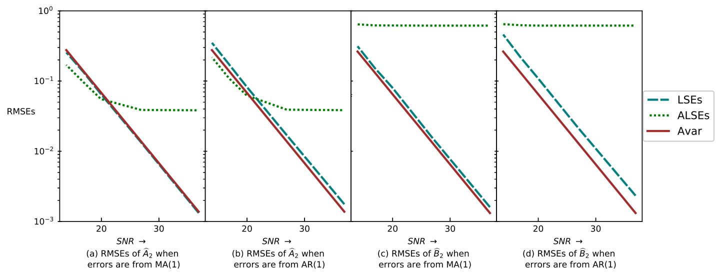

Fig. 8: In each sub-plot the solid line represents the theoretical asymptotic variances of the estimators, the dashed line represents the RMSEs of the LSEs and the dotted line represents the RMSEs of the ALSEs of parameters A_2^0 and B_2^0 of the underlying simulated multiple-component model versus SNR.

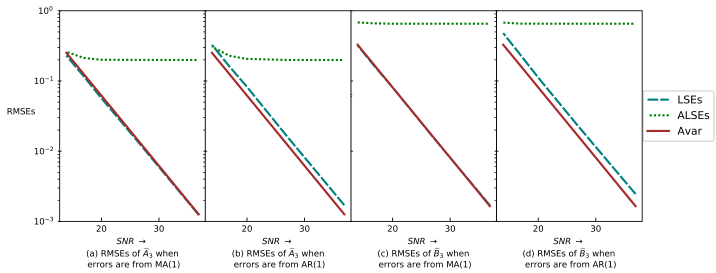

Fig. 9: In each sub-plot the solid line represents the theoretical asymptotic variances of the estimators, the dashed line represents the RMSEs of the LSEs and the dotted line represents the RMSEs of the ALSEs of parameters A_3^0 and B_3^0 of the underlying simulated multiple-component model versus SNR.

Proof: Here in this proof we denote $Q(\boldsymbol{\theta})$ by $Q_n(\boldsymbol{\theta})$ and $\boldsymbol{\theta}$ by $\boldsymbol{\theta}_n$ to emphasize the point that they depend on n. Now, suppose that (9) is true and $\hat{\boldsymbol{\theta}}$ is not a consistent estimator of $\boldsymbol{\theta}^0$, therefore there exists a $\delta > 0$ and a subsequence $\{n_j\}$ of $\{n\}$ such that:

$$\hat{\boldsymbol{\theta}}_{n_j} \in S_\delta \; \forall \; j = 1, 2, \ldots.$$

Note that when $n = n_j$, $\hat{\boldsymbol{\theta}}_{n_j}$ is the least squares estimators of $\boldsymbol{\theta}^{\mathbf{0}}$, that is, it minimises $Q(\boldsymbol{\theta})$, therefore:

$$Q_{n_j}(\hat{\boldsymbol{\theta}}_{n_j}) \leqslant Q_{n_j}(\boldsymbol{\theta}^0). \tag{10}$$

However, since (9) holds true, there exists a J_0, such that $\forall \; j > J_0$

$$Q_{n_j}(\hat{\boldsymbol{\theta}}_{n_j}) - Q_{n_j}(\boldsymbol{\theta}^0) > 0 \; a.s.$$

which is a contradiction to (10).

□

Proof of Theorem 1*:* Consider the following difference:

$$\frac{1}{n}(Q(\boldsymbol{\theta}) - Q(\boldsymbol{\theta}^{\mathbf{0}})) = \frac{1}{n}\sum_{t=1}^{n}\Bigg(\sum_{k=1}^{p}\{A_k^0\cos(k{\alpha^0}t + k{\beta^0}t^2)$$
$$+ B_k^0\sin(k{\alpha^0}t + k{\beta^0}t^2)\} - \sum_{k=1}^{p}\{A_k\cos(k\alpha t + k\beta t^2)$$
$$+ B_k\sin(k\alpha t + k\beta t^2)\}\Bigg)^2 + \frac{2}{n}\sum_{t=1}^{n}X(t)\Bigg(\sum_{k=1}^{p}\{A_k^0$$

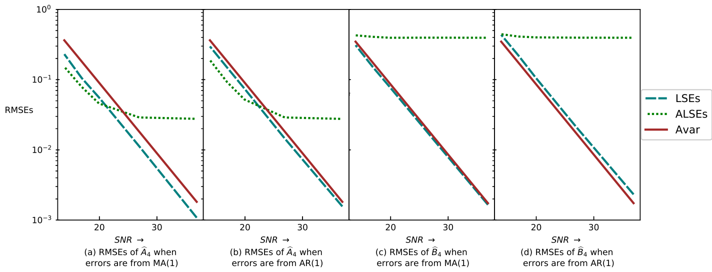

Fig. 10: In each sub-plot the solid line represents the theoretical asymptotic variances of the estimators, the dashed line represents the RMSEs of the LSEs and the dotted line represents the RMSEs of the ALSEs of parameters A_4^0 and B_4^0 of the underlying simulated multiple-component model versus SNR.

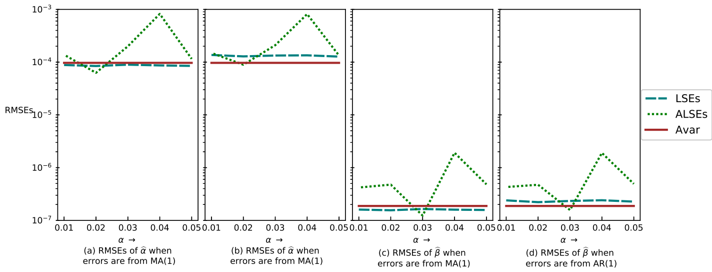

Fig. 11: In each sub-plot the solid line represents the theoretical asymptotic variances of the estimators, the dashed line represents the RMSEs of the LSEs and the dotted line represents the RMSEs of the ALSEs of the non-linear parameters of the underlying simulated multiple-component model versus different values of frequency parameter α^0.

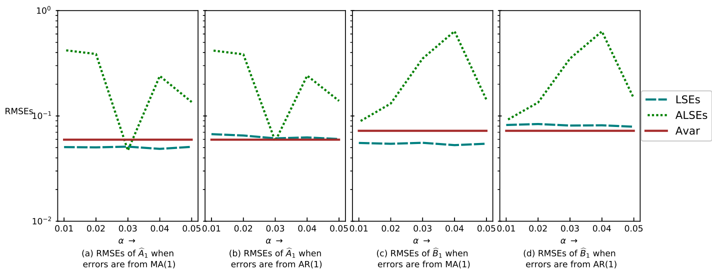

Fig. 12: In each sub-plot the solid line represents the theoretical asymptotic variances of the estimators, the dashed line represents the RMSEs of the LSEs and the dotted line represents the RMSEs of the ALSEs of parameters A_1^0 and B_1^0 of the underlying simulated multiple-component model versus different values of frequency parameter α^0.

$$
\begin{aligned}
&\cos(k\alpha^0 t + k\beta^0 t^2) + B_k^0 \sin(k\alpha^0 t + k\beta^0 t^2)\} \\
&- \sum_{k=1}^{p}\{A_k \cos(k\alpha t + k\beta t^2) + B_k \sin(k\alpha t + k\beta t^2)\}\Bigg) \\
&= f_n(\boldsymbol{\theta}) + g_n(\boldsymbol{\theta}), \text{ where} \\
&f_n(\boldsymbol{\theta}) = \frac{1}{n}\sum_{t=1}^{n}\Bigg(\sum_{k=1}^{p}\{A_k^0 \cos(k\alpha^0 t + k\beta^0 t^2) \\
&+ B_k^0 \sin(k\alpha^0 t + k\beta^0 t^2)\} - \sum_{k=1}^{p}\{A_k \cos(k\alpha t + k\beta t^2) \\
&+ B_k \sin(k\alpha t + k\beta t^2)\}\Bigg)^2,
\end{aligned}
$$

$$
\begin{aligned}
g_n(\boldsymbol{\theta}) = \frac{2}{n}\sum_{t=1}^{n} X(t)\Bigg(&\sum_{k=1}^{p}\{A_k^0 \cos(k\alpha^0 t + k\beta^0 t^2) \\
&+ B_k^0 \sin(k\alpha^0 t + k\beta^0 t^2)\} - \sum_{k=1}^{p}\{A_k \cos(k\alpha t + k\beta t^2) \\
&+ B_k \sin(k\alpha t + k\beta t^2)\}\Bigg).
\end{aligned}
$$

Using Lemma 2 of Lahiri [31], it is easy to see that:

$$
\liminf_{n\to\infty} \inf_{\boldsymbol{\theta}\in S_\delta} g_n(\boldsymbol{\theta}) = 0 \text{ a.s..}
$$

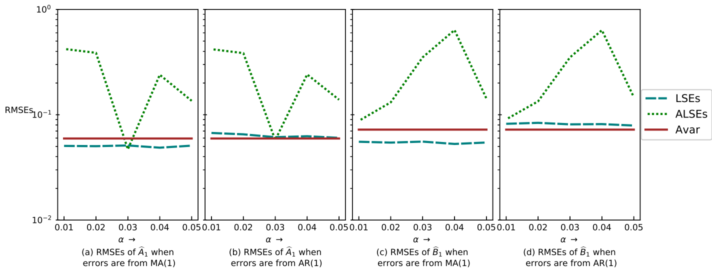

Fig. 13: In each sub-plot the solid line represents the theoretical asymptotic variances of the estimators, the dashed line represents the RMSEs of the LSEs and the dotted line represents the RMSEs of the ALSEs of parameters A_1^0 and B_1^0 of the underlying simulated multiple-component model versus different values of frequency parameter α^0.

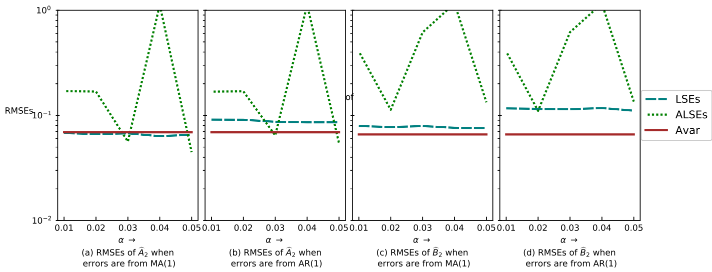

Fig. 14: In each sub-plot the solid line represents the theoretical asymptotic variances of the estimators, the dashed line represents the RMSEs of the LSEs and the dotted line represents the RMSEs of the ALSEs of parameters A_2^0 and B_2^0 of the underlying simulated multiple-component model versus different values of frequency parameter α^0.

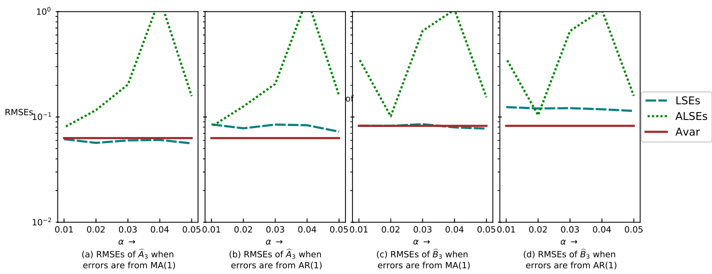

Fig. 15: In each sub-plot the solid line represents the theoretical asymptotic variances of the estimators, the dashed line represents the RMSEs of the LSEs and the dotted line represents the RMSEs of the ALSEs of parameters A_3^0 and B_3^0 of the underlying simulated multiple-component model versus different values of frequency parameter α^0.

Thus it is sufficient to show that:

$$\liminf_{n\to\infty} \inf_{\boldsymbol{\theta} \in S_\delta} f_n(\boldsymbol{\theta}) > 0 \text{ a.s.,} \tag{11}$$

to get the desired result. To show this, we proceed as follows.
We split the set S_δ as follows:

$$S_\delta = S_\delta^{A_1} \cup \ldots \cup S_\delta^{A_p} \cup S_\delta^{B_1} \cup \ldots \cup S_\delta^{B_p} \cup S_\delta^{\alpha} \cup S_\delta^{\beta}, \text{ where}$$

$$\begin{aligned}
S_\delta^{A_j} &= \{\boldsymbol{\theta} : \boldsymbol{\theta} \in \boldsymbol{\Theta}^{(1)};\ |A_j - A_j^0| > \delta\};\ j = 1, \ldots, p,\\
S_\delta^{B_j} &= \{\boldsymbol{\theta} : \boldsymbol{\theta} \in \boldsymbol{\Theta}^{(1)};\ |B_j - B_j^0| > \delta\};\ j = 1, \ldots, p,\\
S_\delta^{\alpha} &= \{\boldsymbol{\theta} : \boldsymbol{\theta} \in \boldsymbol{\Theta}^{(1)};\ |\alpha - \alpha^0| > \delta\}, \text{ and}\\
S_\delta^{\beta} &= \{\boldsymbol{\theta} : \boldsymbol{\theta} \in \boldsymbol{\Theta}^{(1)};\ |\beta - \beta^0| > \delta\}.
\end{aligned}$$

For notational simplicity, let us assume that $p = 2$. Now consider the set $S_\delta^{A_1}$ which can be further split as follows:

$$\begin{aligned}
S_\delta^{A_1} &= \{\boldsymbol{\theta} : \boldsymbol{\theta} \in \boldsymbol{\Theta}^{(1)};\ |A_1 - A_1^0| > \delta; (\alpha, \beta) = (\alpha^0, \beta^0)\}\\
&\cup \{\boldsymbol{\theta} : \boldsymbol{\theta} \in \boldsymbol{\Theta}^{(1)};\ |A_1 - A_1^0| > \delta; (\alpha, \beta) \neq (\alpha^0, \beta^0)\}
\end{aligned}$$

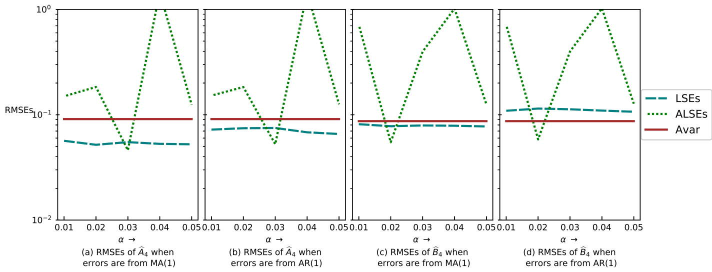

Fig. 16: In each sub-plot the solid line represents the theoretical asymptotic variances of the estimators, the dashed line represents the RMSEs of the LSEs and the dotted line represents the RMSEs of the ALSEs of parameters A_4^0 and B_4^0 of the underlying simulated multiple-component model versus different values of frequency parameter α^0.

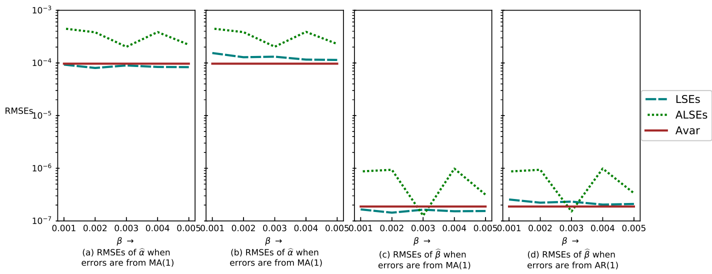

Fig. 17: In each sub-plot the solid line represents the theoretical asymptotic variances of the estimators, the dashed line represents the RMSEs of the LSEs and the dotted line represents the RMSEs of the ALSEs of the non-linear parameters of the underlying simulated multiple-component model versus different values of frequency rate parameter β^0.

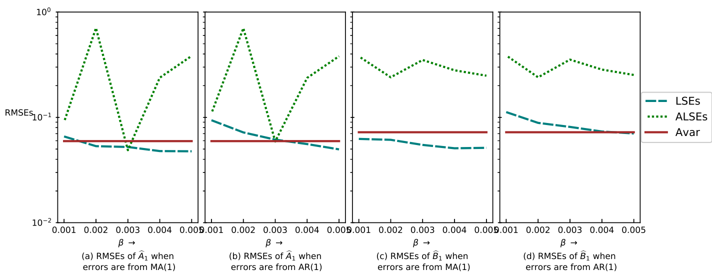

Fig. 18: In each sub-plot the solid line represents the theoretical asymptotic variances of the estimators, the dashed line represents the RMSEs of the LSEs and the dotted line represents the RMSEs of the ALSEs of parameters A_1^0 and B_1^0 of the underlying simulated multiple-component model versus different values of frequency parameter β^0.

$$= S_\delta^{A_1(1)} \cup S_\delta^{A_1(2)}$$

Then, $\liminf_{n\to\infty} \inf_{\boldsymbol{\theta} \in S_\delta^{A_1(1)}} f_n(\boldsymbol{\theta})$

$$= \liminf_{n\to\infty} \inf_{\boldsymbol{\theta} \in S_\delta^{A_1(1)}} \frac{1}{n}\sum_{t=1}^{n}\left(\sum_{k=1}^{2}\Big\{(A_k^0 - A_k)\cos(k\alpha^0 t + k\beta^0 t^2) + (B_k^0 - B_k)\sin(k\alpha^0 t + k\beta^0 t^2)\Big\}\right)^2$$

$$= \frac{1}{2}\Big\{(A_1^0 - A_1)^2 + (A_2^0 - A_2)^2 + (B_1^0 - B_1)^2 + (B_2^0 - B_2)^2\Big\} > 0 \; a.s.,$$

The last equality follows from Lemma 1, part (b)-(g).

Similarly, $\liminf_{n\to\infty} \inf_{\boldsymbol{\theta} \in S_\delta^{A_1(2)}} f_n(\boldsymbol{\theta})$

$$= \liminf_{n\to\infty} \inf_{\boldsymbol{\theta} \in S_\delta^{A_1(2)}} \frac{1}{n}\sum_{t=1}^{n}\left(\sum_{k=1}^{2}\Big\{A_k^0\cos(k\alpha^0 t + k\beta^0 t^2) - A_k\cos(k\alpha t + k\beta t^2) + B_k^0\sin(k\alpha^0 t + k\beta^0 t^2) - B_k\sin(k\alpha t + k\beta t^2)\Big\}\right)^2 = \frac{1}{2}\Big\{A_1^{0\,2} + {A_1}^2 + A_2^{0\,2} + {A_2}^2 + B_1^{0\,2} + {B_1}^2 + B_2^{0\,2} + {B_2}^2\Big\} > 0 \; a.s.,$$

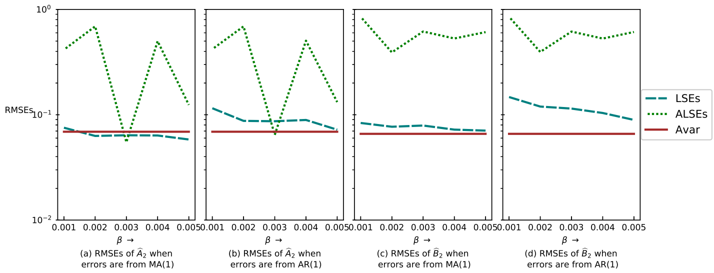

Fig. 19: In each sub-plot the solid line represents the theoretical asymptotic variances of the estimators, the dashed line represents the RMSEs of the LSEs and the dotted line represents the RMSEs of the ALSEs of parameters A_2^0 and B_2^0 of the underlying simulated multiple-component model versus different values of frequency parameter β^0.

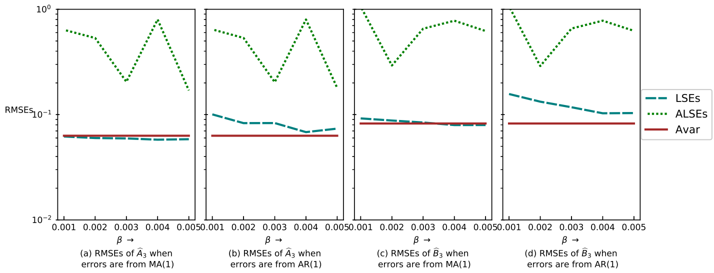

Fig. 20: In each sub-plot the solid line represents the theoretical asymptotic variances of the estimators, the dashed line represents the RMSEs of the LSEs and the dotted line represents the RMSEs of the ALSEs of parameters A_3^0 and B_3^0 of the underlying simulated multiple-component model versus different values of frequency parameter β^0.

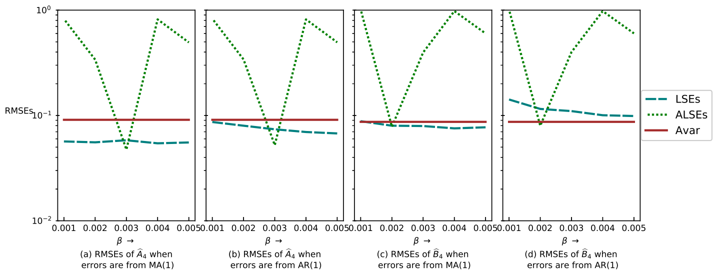

Fig. 21: In each sub-plot the solid line represents the theoretical asymptotic variances of the estimators, the dashed line represents the RMSEs of the LSEs and the dotted line represents the RMSEs of the ALSEs of parameters A_4^0 and B_4^0 of the underlying simulated multiple-component model versus different values of frequency parameter β^0.

On combining, we have $\liminf_{n\to\infty} \inf_{\boldsymbol{\theta} \in S_\delta^{A_1}} f_n(\boldsymbol{\theta}) > 0$. Along similar lines, it can be shown that $\liminf_{n\to\infty} \inf_{\boldsymbol{\theta} \in S} f_n(\boldsymbol{\theta}) > 0$, where S can be $S_\delta^{A_2}$, $S_\delta^{B_j}$, S_δ^{α} or S_δ^{β}. Hence using Lemma 2, the result follows. □

Proof of Theorem 2: Suppose the fitted model has $k = p+1$, number of components. Then

$$\begin{aligned}\widehat{\boldsymbol{\phi}} &= \begin{bmatrix} A_1 & B_1 & \cdots & A_{p+1} & B_{p+1} \end{bmatrix}^\top \\ &= [\mathbf{Z}(\alpha,\beta)^\top \mathbf{Z}(\alpha,\beta)]^{-1}\mathbf{Z}(\alpha,\beta)^\top \mathbf{Y}\end{aligned}$$

where, the matrix $\mathbf{Z}(\alpha,\beta)$ can be redefined by replacing p by $p+1$ in (4). Using Lemma 1, parts (b)-(g), it can be easily seen that the vector of LSEs of linear parameters:

$$\widehat{\boldsymbol{\phi}} = \widetilde{\boldsymbol{\phi}} + o(1)$$

where,

$$\widetilde{\boldsymbol{\phi}} = \frac{2}{n}\mathbf{Z}(\alpha,\beta)^\top \mathbf{Y}$$

is the vector of ALSEs of the linear parameters. Now the result follows from Theorem 5. □

8.2 *Proof of asymptotic normality of the LSEs*

Proof of Theorem 3: Let $\boldsymbol{Q}'(\boldsymbol{\theta})$ be the first derivative vector and $\boldsymbol{Q}''(\boldsymbol{\theta})$ be the second derivative matrix of the error sum of squares

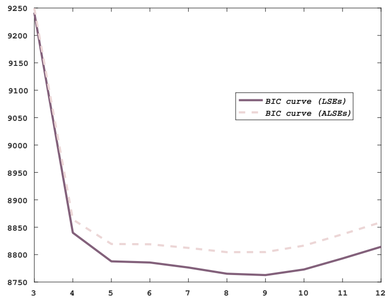

Fig. 22: BIC curves for the data 'u:'.

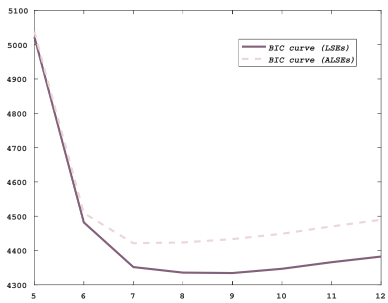

Fig. 23: BIC curves for the data 'a:'.

function $Q(\boldsymbol{\theta})$ defined in (3).

Let us first consider the first derivative vector defined as follows:

$$\boldsymbol{Q}'(\boldsymbol{\theta})_{1\times 2p+2} = \begin{bmatrix} \frac{\partial Q(\boldsymbol{\theta})}{\partial A_1} & \frac{\partial Q(\boldsymbol{\theta})}{\partial B_1} & \cdots & \frac{\partial Q(\boldsymbol{\theta})}{\partial A_p} & \frac{\partial Q(\boldsymbol{\theta})}{\partial B_p} & \frac{\partial Q(\boldsymbol{\theta})}{\partial \alpha} & \frac{\partial Q(\boldsymbol{\theta})}{\partial \beta} \end{bmatrix}.$$

At $\boldsymbol{\theta} = \boldsymbol{\theta}^0$,

$$\frac{\partial Q(\boldsymbol{\theta}^0)}{\partial A_j} = -2\sum_{t=1}^{n} X(t)\cos(j\alpha^0 t + j\beta^0 t^2);\ j = 1, \ldots, p,$$

$$\frac{\partial Q(\boldsymbol{\theta}^0)}{\partial B_j} = -2\sum_{t=1}^{n} X(t)\sin(j\alpha^0 t + j\beta^0 t^2);\ j = 1, \ldots, p,$$

$$\frac{\partial Q(\boldsymbol{\theta}^0)}{\partial \alpha} = -2\sum_{t=1}^{n} tX(t)\bigg(\sum_{k=1}^{p}\{-A_k^0\sin(k\alpha^0 t + k\beta t^2) + B_k^0\cos(k\alpha^0 t + k\beta^0 t^2)\}\bigg),$$

$$\frac{\partial Q(\boldsymbol{\theta}^0)}{\partial \beta} = -2\sum_{t=1}^{n} t^2 X(t)\bigg(\sum_{k=1}^{p}\{-A_k^0\sin(k\alpha^0 t + k\beta^0 t^2) + B_k^0\cos(k\alpha^0 t + k\beta^0 t^2)\}\bigg).$$

The second derivative matrix $\boldsymbol{Q}''(\boldsymbol{\theta})$ is a $(2p+2)\times(2p+2)$ matrix defined as follows:

$$\boldsymbol{Q}''(\boldsymbol{\theta}) = \begin{bmatrix} \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial A_1^2} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial A_1 \partial B_1} & \cdots & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial A_1 \partial A_p} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial A_1 \partial B_p} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial A_1 \partial \alpha} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial A_1 \partial \beta} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial A_p \partial A_1} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial A_p \partial B_1} & \cdots & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial A_p^2} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial A_p \partial B_p} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial A_p \partial \alpha} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial A_p \partial \beta} \\ \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial B_1 \partial A_1} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial B_1^2} & \cdots & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial B_1 \partial A_p} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial B_1 \partial B_p} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial B_1 \partial \alpha} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial B_1 \partial \beta} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial B_p \partial A_1} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial B_p \partial B_1} & \cdots & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial B_p \partial A_p} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial B_p^2} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial B_p \partial \alpha} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial B_p \partial \beta} \\ \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial \alpha \partial A_1} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial \alpha \partial B_1} & \cdots & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial \alpha \partial A_p} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial \alpha \partial B_p} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial \alpha^2} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial \alpha \partial \beta} \\ \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial \beta \partial A_1} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial \beta \partial B_1} & \cdots & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial \beta \partial A_p} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial \beta \partial B_p} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial \beta \partial \alpha} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial \beta^2} \end{bmatrix}.$$

At $\boldsymbol{\theta} = \boldsymbol{\theta}^0$, for $j, l = 1, \ldots p$:

$$\frac{\partial^2 Q(\boldsymbol{\theta}^0)}{\partial A_j^2} = 2\sum_{t=1}^{n}\cos^2(j\alpha^0 t + j\beta^0 t^2),$$

$$\frac{\partial^2 Q(\boldsymbol{\theta}^0)}{\partial B_j^2} = 2\sum_{t=1}^{n}\sin^2(j\alpha^0 t + j\beta^0 t^2),$$

$$\frac{\partial^2 Q(\boldsymbol{\theta}^0)}{\partial A_j \partial A_l} = 2\sum_{t=1}^{n}\cos(j\alpha^0 t + j\beta^0 t^2)\cos(l\alpha^0 t + l\beta^0 t^2);\ j \neq l,$$

$$\frac{\partial^2 Q(\boldsymbol{\theta}^0)}{\partial B_j \partial B_l} = 2\sum_{t=1}^{n}\sin(j\alpha^0 t + j\beta^0 t^2)\sin(l\alpha^0 t + l\beta^0 t^2);\ j \neq l,$$

$$\frac{\partial^2 Q(\boldsymbol{\theta}^0)}{\partial A_j \partial B_l} = 2\sum_{t=1}^{n}\cos(j\alpha^0 t + j\beta^0 t^2)\sin(l\alpha^0 t + l\beta^0 t^2),$$

$$\frac{\partial^2 Q(\boldsymbol{\theta}^0)}{\partial \alpha \partial A_j} = 2\sum_{t=1}^{n} t\cos(j\alpha t + j\beta t^2)\bigg\{\sum_{k=1}^{p} k(-A_k^0\sin(k\alpha^0 t + k\beta^0 t^2) + B_k^0\cos(k\alpha^0 t + k\beta^0 t^2))\bigg\} + 2j\sum_{t=1}^{n} tX(t)\sin(j\alpha t + j\beta t^2),$$

$$\frac{\partial^2 Q(\boldsymbol{\theta}^0)}{\partial \beta \partial A_j} = 2\sum_{t=1}^{n} t^2\cos(j\alpha t + j\beta t^2)\bigg\{\sum_{k=1}^{p} k(-A_k^0\sin(k\alpha^0 t + k\beta^0 t^2) + B_k^0\cos(k\alpha^0 t + k\beta^0 t^2))\bigg\} + 2j\sum_{t=1}^{n} t^2 X(t)\sin(j\alpha t + j\beta t^2),$$

$$\frac{\partial^2 Q(\boldsymbol{\theta}^0)}{\partial \alpha \partial B_j} = 2\sum_{t=1}^{n} t\sin(j\alpha t + j\beta t^2)\bigg\{\sum_{k=1}^{p} k (-A_k^0\sin(k\alpha^0 t + k\beta^0 t^2) + B_k^0\cos(k\alpha^0 t + k\beta^0 t^2))\bigg\} + 2j\sum_{t=1}^{n} tX(t)\cos(j\alpha t + j\beta t^2),$$

$$\frac{\partial^2 Q(\boldsymbol{\theta}^0)}{\partial \beta \partial A_j} = 2\sum_{t=1}^{n} t^2\sin(j\alpha t + j\beta t^2)\bigg\{\sum_{k=1}^{p} k (-A_k^0\sin(k\alpha^0 t + k\beta^0 t^2) + B_k^0\cos(k\alpha^0 t + k\beta^0 t^2))\bigg\} + 2j\sum_{t=1}^{n} t^2 X(t)\cos(j\alpha t + j\beta t^2)$$

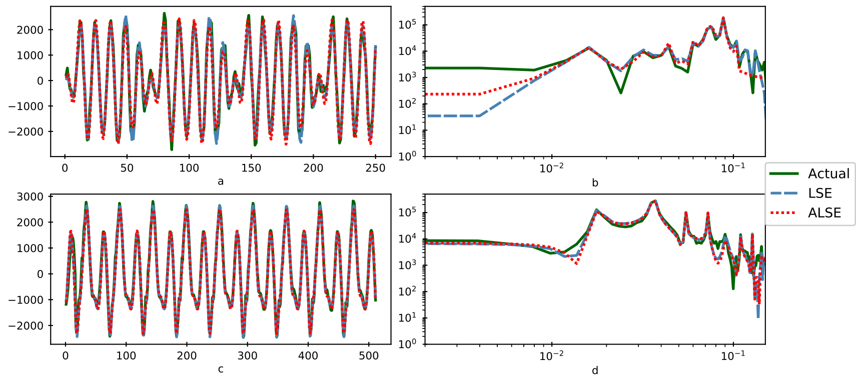

Fig. 24: First subplot: the estimated signals using LSEs and ALSEs along with the original signal 'a:' in the time domain; second subplot: the amplitude versus frequency plot of the estimated signals along with the original signal 'a:'; third subplot: the estimated signals using LSEs and ALSEs along with the original signal 'u:' in the time domain; fourth subplot: the amplitude versus frequency plot of the estimated signals along with the original signal 'u:'.

$$\frac{\partial^2 Q(\boldsymbol{\theta}^0)}{\partial \alpha^2} = 2\sum_{t=1}^{n} t^2 X(t)\Bigg\{\sum_{k=1}^{p} k^2 (A_k^0 \cos(k{\alpha^0} t + k{\beta^0} t^2) + B_k^0 \sin(k{\alpha^0} t + k{\beta^0} t^2))\Bigg\} + 2\sum_{t=1}^{n} t^2 \Bigg\{\sum_{k=1}^{p} k(-A_k^0 \sin(k{\alpha^0} t + k{\beta^0} t^2) + B_k^0 \cos(k{\alpha^0} t + k{\beta^0} t^2))\Bigg\},$$

$$\frac{\partial^2 Q(\boldsymbol{\theta}^0)}{\partial \alpha \partial \beta} = 2\sum_{t=1}^{n} t^3 X(t)\Bigg\{\sum_{k=1}^{p} k^2 (A_k^0 \cos(k{\alpha^0} t + k{\beta^0} t^2) + B_k^0 \sin(k{\alpha^0} t + k{\beta^0} t^2))\Bigg\} + 2\sum_{t=1}^{n} t^3 \Bigg\{\sum_{k=1}^{p} k(-A_k^0 \sin(k{\alpha^0} t + k{\beta^0} t^2) + B_k^0 \cos(k{\alpha^0} t + k{\beta^0} t^2))\Bigg\}$$

$$,\frac{\partial^2 Q(\boldsymbol{\theta}^0)}{\partial \beta^2} = 2\sum_{t=1}^{n} t^4 X(t)\Bigg\{\sum_{k=1}^{p} k^2 (A_k^0 \cos(k{\alpha^0} t + k{\beta^0} t^2) + B_k^0 \sin(k{\alpha^0} t + k{\beta^0} t^2))\Bigg\} + 2\sum_{t=1}^{n} t^4 \Bigg\{\sum_{k=1}^{p} k(-A_k^0 \sin(k{\alpha^0} t + k{\beta^0} t^2) + B_k^0 \cos(k{\alpha^0} t + k{\beta^0} t^2))\Bigg\}.$$

We expand $\boldsymbol{Q}'(\hat{\boldsymbol{\theta}})$ around the point $\boldsymbol{\theta}^0$ using Taylor series expansion as follows:

$$\boldsymbol{Q}'(\hat{\boldsymbol{\theta}}) - \boldsymbol{Q}'(\boldsymbol{\theta}^0) = (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0)\boldsymbol{Q}''(\bar{\boldsymbol{\theta}}), \tag{12}$$

where $\bar{\boldsymbol{\theta}}$ is a point between $\hat{\boldsymbol{\theta}}$ and $\boldsymbol{\theta}^0$. Since, $\hat{\boldsymbol{\theta}}$ is the LSE of $\boldsymbol{\theta}^0$, $\boldsymbol{Q}'(\hat{\boldsymbol{\theta}}) = 0$ and therefore (12) can be rewritten as:

$$(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0) = -\boldsymbol{Q}'(\boldsymbol{\theta}^0)[\boldsymbol{Q}''(\bar{\boldsymbol{\theta}})]^{-1}. \tag{13}$$

Multiplying both sides of (13) by a diagonal matrix $\mathbf{D} = \text{diag}(\underbrace{\frac{1}{\sqrt{n}}, \ldots, \frac{1}{\sqrt{n}}}_{2p\ times}, \frac{1}{n\sqrt{n}}, \frac{1}{n^2\sqrt{n}})$ we get:

$$(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0)\mathbf{D}^{-1} = -\boldsymbol{Q}'(\boldsymbol{\theta}^0)\mathbf{D}[\mathbf{D}\boldsymbol{Q}''(\bar{\boldsymbol{\theta}})\mathbf{D}]^{-1}. \tag{14}$$

Now using Central Limit theorem of a linear process, see Fuller [24], it can be seen that:

$$-\boldsymbol{Q}'(\boldsymbol{\theta}^0)\mathbf{D} \xrightarrow{L} \mathcal{N}_{2p+2}(0, 2c\sigma^2\boldsymbol{\Sigma}^{-1}). \tag{15}$$

Since $\bar{\boldsymbol{\theta}}$ is a point between $\hat{\boldsymbol{\theta}}$ and $\boldsymbol{\theta}^0$ and $\hat{\boldsymbol{\theta}} \xrightarrow{a.s.} \boldsymbol{\theta}^0$,

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

Using Lemma 1, it is straightforward to show that:

$$\lim_{n\to\infty} \mathbf{D}\boldsymbol{Q}''(\boldsymbol{\theta}^0)\mathbf{D} = \boldsymbol{\Sigma}^{-1}. \tag{16}$$

Combining the results (15), (16) and (14), we get:

$$(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0)\mathbf{D}^{-1} \xrightarrow{L} \mathcal{N}_{2p+2}(0, 2c\sigma^2\boldsymbol{\Sigma}).$$

Hence, the result.

□

8.3 *Proof of consistency of the ALSEs*

We need the following lemmas to prove Theorem 4:

Lemma 3. *Consider the set:*

$$M_\delta = \{(\alpha, \beta) : (\alpha, \beta) \in \left[0, \frac{\pi}{p}\right] \times \left[0, \frac{\pi}{p}\right]; |\alpha - \alpha^0| > \delta \text{ or } |\beta - \beta^0| > \delta\}.$$

If for any $\delta > 0$,

$$\limsup_{n\to\infty} \sup_{(\alpha,\beta)\in M_\delta} \frac{1}{n}(I(\alpha, \beta) - I(\alpha^0, \beta^0)) < 0 \ a.s.$$

then $\tilde{\alpha}$ is a consistent estimator of α^0 and $\tilde{\beta}$ is a consistent estimator of β^0.

Proof: This proof follows along the same lines as Lemma 2.

□

Lemma 4. *The ALSEs of the non-linear parameters α^0 and β^0 are strongly consistent, that is,*

$$\tilde{\alpha} \xrightarrow{a.s.} \alpha^0 \text{ as } n \to \infty \qquad \tilde{\beta} \xrightarrow{a.s.} \beta^0 \text{ as } n \to \infty.$$

Proof: Consider the following difference:

$$\begin{aligned}
&\frac{1}{n}\Big(I(\alpha,\beta) - I(\alpha^0,\beta^0)\Big) \\
&= \frac{1}{n^2}\sum_{k=1}^{p}\Bigg[\Big\{\sum_{t=1}^{n} y(t)\cos(k\alpha t + k\beta t^2)\Big\}^2 + \Big\{\sum_{t=1}^{n} y(t)\sin(k\alpha t + k\beta t^2)\Big\}^2 \\
&- \Big\{\sum_{t=1}^{n} y(t)\cos(k\alpha^0 t + k\beta^0 t^2)\Big\}^2 + \Big\{\sum_{t=1}^{n} y(t)\sin(k\alpha^0 t + k\beta^0 t^2)\Big\}^2\Bigg] \\
&= \sum_{k=1}^{p}\Bigg[\Bigg\{\frac{1}{n}\sum_{t=1}^{n}\Big(\sum_{j=1}^{p} A_j^0\cos(j\alpha^0 t + j\beta^0 t^2) + B_j^0\sin(j\alpha^0 t + j\beta^0 t^2) + X(t)\Big)\cos(k\alpha t + k\beta t^2)\Bigg\}^2 \\
&+ \sum_{k=1}^{p}\Bigg[\Bigg\{\frac{1}{n}\sum_{t=1}^{n}\Big(\sum_{j=1}^{p} A_j^0\cos(j\alpha^0 t + j\beta^0 t^2) + B_j^0\sin(j\alpha^0 t + j\beta^0 t^2) + X(t)\Big)\sin(k\alpha t + k\beta t^2)\Bigg\}^2 \\
&- \sum_{k=1}^{p}\Bigg[\Bigg\{\frac{1}{n}\sum_{t=1}^{n}\Big(\sum_{j=1}^{p} A_j^0\cos(j\alpha^0 t + j\beta^0 t^2) + B_j^0\sin(j\alpha^0 t + j\beta^0 t^2) + X(t)\Big)\cos(k\alpha^0 t + k\beta^0 t^2)\Bigg\}^2 \\
&- \sum_{k=1}^{p}\Bigg[\Bigg\{\frac{1}{n}\sum_{t=1}^{n}\Big(\sum_{j=1}^{p} A_j^0\cos(j\alpha^0 t + j\beta^0 t^2) + B_j^0\sin(j\alpha^0 t + j\beta^0 t^2) + X(t)\Big)\sin(k\alpha^0 t + k\beta^0 t^2)\Bigg\}^2.
\end{aligned}$$

Using Lemma 1, parts (b)-(g), we have:

$$\limsup_{n\to\infty}\sup_{(\alpha,\beta)\in M_\delta}\frac{1}{n}\Big(I(\alpha,\beta) - I(\alpha^0,\beta^0)\Big) = -\sum_{k=1}^{p}\frac{{A_k^0}^2 + {B_k^0}^2}{4} < 0 \text{ a.s.}$$

Therefore, using Lemma 3,we have the desired result.

□

Lemma 5. *If $\tilde{\alpha}$ and $\tilde{\beta}$ are the ALSEs of α^0 and β^0 respectively, then:*

$$n(\tilde{\alpha} - \alpha^0) \xrightarrow{a.s.} 0 \qquad n^2(\tilde{\beta} - \beta^0) \xrightarrow{a.s.} 0.$$

Proof: Let $\boldsymbol{I}'(\alpha,\beta)$ be the first derivative vector and $\boldsymbol{I}''(\alpha,\beta)$ be the second derivative matrix of the periodogram-type function $I(\alpha,\beta)$ defined in (8). Using Taylor series expansion, we expand $\boldsymbol{I}'(\tilde{\alpha},\tilde{\beta})$ around the point (α^0,β^0) as follows:

$$\boldsymbol{I}'(\tilde{\alpha},\tilde{\beta}) - \boldsymbol{I}'(\alpha^0,\beta^0) = \big((\tilde{\alpha}-\alpha^0) \quad (\tilde{\beta}-\beta^0)\big)\boldsymbol{I}''(\bar{\alpha},\bar{\beta}). \tag{17}$$

Here, $(\bar{\alpha},\bar{\beta})$ is a point between $(\tilde{\alpha},\tilde{\beta})$ and (α^0,β^0). Since, $(\tilde{\alpha},\tilde{\beta})$ maximises the function $I(\alpha,\beta)$, therefore $\boldsymbol{I}'(\tilde{\alpha},\tilde{\beta}) = 0$, and (17) can be rewritten as:

$$\big((\tilde{\alpha}-\alpha^0) \quad (\tilde{\beta}-\beta^0)\big) = -\boldsymbol{I}'(\alpha^0,\beta^0)[\boldsymbol{I}''(\bar{\alpha},\bar{\beta})]^{-1}.$$

Multiplying the above equation with $\frac{1}{\sqrt{n}}\mathbf{D}_1^{-1}$ where diagonal matrix $\mathbf{D}_1 = \text{diag}(\frac{1}{n\sqrt{n}}, \frac{1}{n^2\sqrt{n}})$, we have:

$$\begin{aligned}
&\big((\tilde{\alpha}-\alpha^0) \quad (\tilde{\beta}-\beta^0)\big)\frac{1}{\sqrt{n}}\mathbf{D}_1^{-1} \\
&= -\frac{1}{\sqrt{n}}\boldsymbol{I}'(\alpha^0,\beta^0)\mathbf{D}_1[\mathbf{D}_1\boldsymbol{I}''(\bar{\alpha},\bar{\beta})\mathbf{D}_1]^{-1}.
\end{aligned} \tag{18}$$

Since $(\tilde{\alpha},\tilde{\beta})$ is a consistent estimator of (α^0,β^0) (see Lemma 4),

$$\lim_{n\to\infty}\mathbf{D}_1\boldsymbol{I}''(\bar{\alpha},\bar{\beta})\mathbf{D}_1 = \lim_{n\to\infty}\mathbf{D}_1\boldsymbol{I}''(\alpha^0,\beta^0)\mathbf{D}_1. \tag{19}$$

Consider the first derivative vector:

$$\boldsymbol{I}'(\alpha^0,\beta^0) = \begin{bmatrix}\frac{\partial I(\alpha,\beta)}{\partial\alpha} & \frac{\partial I(\alpha,\beta)}{\partial\beta}\end{bmatrix}\Bigg|_{\alpha=\alpha^0,\ \beta=\beta^0}$$

Here,

$$\begin{aligned}
\frac{1}{n^2}\frac{\partial I(\alpha^0,\beta^0)}{\partial\alpha} &= \frac{2}{n^3}\sum_{k=1}^{p} k\Bigg[\Big\{\sum_{t=1}^{n} y(t)\cos(k\alpha^0 t + k\beta^0 t^2)\Big\}\Big\{-\sum_{t=1}^{n} t y(t)\sin(k\alpha^0 t + k\beta^0 t^2)\Big\} \\
&\Big\{\sum_{t=1}^{n} y(t)\sin(k\alpha^0 t + k\beta^0 t^2)\Big\}\Big\{\sum_{t=1}^{n} t y(t)\cos(k\alpha^0 t + k\beta^0 t^2)\Big\}\Bigg] \\
\frac{1}{n^3}\frac{\partial I(\alpha^0,\beta^0)}{\partial\beta} &= \frac{2}{n^3}\sum_{k=1}^{p} k\Bigg[\Big\{\sum_{t=1}^{n} y(t)\cos(k\alpha^0 t + k\beta^0 t^2)\Big\}\Big\{-\sum_{t=1}^{n} t^2 y(t)\sin(k\alpha^0 t + k\beta^0 t^2)\Big\} \\
&\Big\{\sum_{t=1}^{n} y(t)\sin(k\alpha^0 t + k\beta^0 t^2)\Big\}\Big\{\sum_{t=1}^{n} t^2 y(t)\cos(k\alpha^0 t + k\beta^0 t^2)\Big\}\Bigg].
\end{aligned}$$

Using Lemma 1, parts (b) - (g), we have:

$$\frac{1}{\sqrt{n}}\boldsymbol{I}'(\alpha^0,\beta^0)\mathbf{D}_1 \xrightarrow{a.s.} 0 \ n \to \infty. \tag{20}$$

The second derivative matrix is defined as follows:

$$\boldsymbol{I}''(\alpha^0,\beta^0) = \begin{bmatrix}\frac{\partial^2 I(\alpha,\beta)}{\partial\alpha^2} & \frac{\partial^2 I(\alpha,\beta)}{\partial\alpha\partial\beta} \\ \frac{\partial^2 I(\alpha,\beta)}{\partial\beta\partial\alpha} & \frac{\partial^2 I(\alpha,\beta)}{\partial\beta^2}\end{bmatrix}\Bigg|_{\alpha=\alpha^0,\ \beta=\beta^0}$$

$$\frac{1}{n^3}\frac{\partial^2 I(\alpha,\beta)}{\partial\alpha^2} = \frac{2}{n^4}\sum_{k=1}^{p} k^2\Bigg[\Bigg\{\sum_{t=1}^{n} y(t)\cos(k\alpha^0 t + k\beta^0 t^2)\Bigg\}$$
$$\Bigg\{-\sum_{t=1}^{n} t^2 y(t)\cos(k\alpha^0 t + k\beta^0 t^2)\Bigg\}$$
$$+\Bigg\{\sum_{t=1}^{n} t y(t)\sin(k\alpha^0 t + k\beta^0 t^2)\Bigg\}^2$$
$$+\Bigg\{\sum_{t=1}^{n} t y(t)\cos(k\alpha^0 t + k\beta^0 t^2)\Bigg\}^2$$
$$+\Bigg\{\sum_{t=1}^{n} y(t)\sin(k\alpha^0 t + k\beta^0 t^2)\Bigg\}$$
$$\Bigg\{-\sum_{t=1}^{n} t^2 y(t)\sin(k\alpha^0 t + k\beta^0 t^2)\Bigg\}\Bigg]$$

$$\frac{1}{n^5}\frac{\partial^2 I(\alpha,\beta)}{\partial\beta^2} = \frac{2}{n^6}\sum_{k=1}^{p} k^2\Bigg[\Bigg\{\sum_{t=1}^{n} y(t)\cos(k\alpha^0 t + k\beta^0 t^2)\Bigg\}$$
$$\Bigg\{-\sum_{t=1}^{n} t^4 y(t)\cos(k\alpha^0 t + k\beta^0 t^2)\Bigg\}$$
$$+\Bigg\{\sum_{t=1}^{n} t^2 y(t)\sin(k\alpha^0 t + k\beta^0 t^2)\Bigg\}^2$$
$$+\Bigg\{\sum_{t=1}^{n} t^2 y(t)\cos(k\alpha^0 t + k\beta^0 t^2)\Bigg\}^2$$
$$+\Bigg\{\sum_{t=1}^{n} y(t)\sin(k\alpha^0 t + k\beta^0 t^2)\Bigg\}$$
$$\Bigg\{-\sum_{t=1}^{n} t^4 y(t)\sin(k\alpha^0 t + k\beta^0 t^2)\Bigg\}\Bigg]$$
$$\frac{1}{n^4}\frac{\partial^2 I(\alpha,\beta)}{\partial\alpha\partial\beta} = \frac{2}{n^5}\sum_{k=1}^{p} k^2\Bigg[\Bigg\{\sum_{t=1}^{n} y(t)\cos(k\alpha^0 t + k\beta^0 t^2)\Bigg\}$$
$$\Bigg\{-\sum_{t=1}^{n} t^3 y(t)\cos(k\alpha^0 t + k\beta^0 t^2)\Bigg\}$$
$$+\Bigg\{-\sum_{t=1}^{n} t^2 y(t)\sin(k\alpha^0 t + k\beta^0 t^2)\Bigg\}$$
$$\Bigg\{-\sum_{t=1}^{n} t y(t)\sin(k\alpha^0 t + k\beta^0 t^2)\Bigg\}$$
$$+\Bigg\{\sum_{t=1}^{n} t^2 y(t)\cos(k\alpha^0 t + k\beta^0 t^2)\Bigg\}$$
$$\Bigg\{\sum_{t=1}^{n} t y(t)\cos(k\alpha^0 t + k\beta^0 t^2)\Bigg\}$$
$$+\Bigg\{\sum_{t=1}^{n} y(t)\sin(k\alpha^0 t + k\beta^0 t^2)\Bigg\}$$
$$\Bigg\{-\sum_{t=1}^{n} t^3 y(t)\sin(k\alpha^0 t + k\beta^0 t^2)\Bigg\}\Bigg]$$

Using Lemma 1, parts (b)-(g), we get:

$$\lim_{n\to\infty} \mathbf{D}_1 \boldsymbol{I}''(\alpha^0,\beta^0)\mathbf{D}_1 = -\mathbf{S} \tag{21}$$

where,

$$\mathbf{S} = \begin{bmatrix} \frac{\sum_{k=1}^{p} k^2({A_k^0}^2+{B_k^0}^2)^2}{24} & \frac{\sum_{k=1}^{p} k^2({A_k^0}^2+{B_k^0}^2)^2}{24} \\ \frac{\sum_{k=1}^{p} k^2({A_k^0}^2+{B_k^0}^2)^2}{24} & 2\frac{\sum_{k=1}^{p} k^2({A_k^0}^2+{B_k^0}^2)^2}{45} \end{bmatrix}$$

is a positive definite matrix. On combining (18), (19), (20) and (21), we get the desired result.

□

Proof of Theorem 4*:* The strong consistency of the non-linear parameters of $\tilde{\alpha}$ and $\tilde{\beta}$ follows from Lemma 4. Now we need to prove the consistency of the linear parameters.

For $j = 1,\ldots,p$:

$$\tilde{A}_j = \frac{2}{n}\sum_{t=1}^{n} y(t)\cos(j\tilde{\alpha}t + j\tilde{\beta}t^2)$$
$$\tilde{B}_j = \frac{2}{n}\sum_{t=1}^{n} y(t)\sin(j\tilde{\alpha}t + j\tilde{\beta}t^2)$$

Consider the ALSE $\tilde{A}_j$ of A_j^0:

$$\tilde{A}_j = \frac{2}{n}\sum_{t=1}^{n}\Bigg(\sum_{k=1}^{p} A_k^0\cos(k\alpha^0 t + k\beta^0 t^2) + B_k^0\sin(k\alpha^0 t + k\beta^0 t^2)$$
$$+ X(t)\Bigg)\cos(j\tilde{\alpha}t + j\tilde{\beta}t^2)$$
$$= \frac{2}{n}\sum_{t=1}^{n}\Bigg(\sum_{k=1}^{p} A_k^0\cos(k\alpha^0 t + k\beta^0 t^2) + B_k^0\sin(k\alpha^0 t + k\beta^0 t^2)$$
$$+ X(t)\Bigg)\Bigg(\cos(j\alpha^0 t + j\beta^0 t^2)$$
$$- jt(\tilde{\alpha}-\alpha^0)\sin(j\bar{\alpha}t + j\bar{\beta}t^2) - jt^2(\tilde{\beta}-\beta^0)\sin(j\bar{\alpha}t + j\bar{\beta}t^2)\Bigg)$$

(by expanding $\cos(j\tilde{\alpha}t + j\tilde{\beta}t^2)$ around (α^0,β^0) using Taylor series)

$$\Rightarrow \tilde{A}_j \to A_j^0 \text{ a.s. as } n\to\infty;\ j = 1,\ldots p$$

using Lemma 1, parts (b)-(g) and Lemma 5. Similarly, it can be proved that $\tilde{B}_j^0$ is a strongly consistent estimator of B_j^0 for all $j = 1,\ldots,p$. Hence, the result.

□

Proof of Theorem 5*:* Suppose the fitted model has $k = p+1$, number of components. Now consider the ALSE of A_{p+1}^0:

$$\tilde{A}_{p+1} = \frac{2}{n}\sum_{t=1}^{n} y(t)\cos((p+1)\tilde{\alpha}t + (p+1)\tilde{\beta}t^2)$$
$$= \frac{2}{n}\sum_{t=1}^{n}\Bigg(\sum_{k=1}^{p} A_k^0\cos(k\alpha^0 t + k\beta^0 t^2)$$
$$+ B_k^0\sin(k\alpha^0 t + k\beta^0 t^2)\Bigg)$$
$$\Bigg(\cos((p+1)\alpha^0 t + (p+1)\beta^0 t^2) - (p+1)t(\tilde{\alpha}-\alpha^0)$$
$$\sin((p+1)\bar{\alpha}t + (p+1)\bar{\beta}t^2) - (p+1)t^2(\tilde{\beta}-\beta^0)$$
$$\sin((p+1)\bar{\alpha}t + (p+1)\bar{\beta}t^2)\Bigg)$$

$$+\frac{2}{n}\sum_{t=1}^{n}X(t)\cos((p+1)\tilde{\alpha}t+(p+1)\tilde{\beta}t^2)$$

in the first term, we expand $\cos((p+1)\tilde{\alpha}t+(p+1)\tilde{\beta}t^2)$ around (α^0,β^0) using Taylor series)

Using Lemma 1, parts (e)-(g), Lemma 4 and 5, we have:

$$\tilde{A}_{p+1}=\frac{2}{n}\sum_{t=1}^{n}X(t)\cos((p+1)\tilde{\alpha}t+(p+1)\tilde{\beta}t^2)+o(1).$$

Now lastly using Lemma 2 of Lahiri [31], we get $\tilde{A}_{p+1}\to 0$ a.s. Similarly, it can be shown that $\tilde{B}_{p+1}\to 0$ a.s. and thereby, the result follows.

8.4 Proof of asymptotic normality of the ALSEs

Proof of Theorem 6*:* Consider the following:

$$\frac{1}{n}Q(\boldsymbol{\theta})=\frac{1}{n}\sum_{t=1}^{n}\Bigg(y(t)-\sum_{k=1}^{p}(A_k\cos(k\alpha t+k\beta t^2)$$

$$+B_k\sin(k\alpha t+k\beta t^2))\Bigg)^2$$

$$=\frac{1}{n}\sum_{t=1}^{n}y(t)^2-\sum_{k=1}^{p}\left\{\frac{2}{n}\sum_{t=1}^{n}y(t)\{A_k\cos(k\alpha t+k\beta t^2)\right.$$

$$\left.+B_k\sin(k\alpha t+k\beta t^2)\}\right\}$$

$$+\sum_{k=1}^{p}\left\{\frac{1}{n}\sum_{t=1}^{n}(A_k\cos(k\alpha t+k\beta t^2)\right.$$

$$\left.+B_k\sin(k\alpha t+k\beta t^2))^2\right\}$$

$$=\frac{1}{n}\sum_{t=1}^{n}y(t)^2-\sum_{k=1}^{p}\left\{\frac{2}{n}\sum_{t=1}^{n}y(t)\{A_k\cos(k\alpha t+k\beta t^2)\right.$$

$$\left.+B_k\sin(k\alpha t+k\beta t^2)\}\right\}+\sum_{k=1}^{p}\frac{1}{2}\left(A_k^2+B_k^2\right)+o(1)$$

$$=C-\frac{1}{n}J(\boldsymbol{\theta})+o(1).$$

Here, $C=\frac{1}{n}\sum_{t=1}^{n}y^2(t)$ and

$$\frac{1}{n}J(\boldsymbol{\theta})=\sum_{k=1}^{p}\left\{\frac{2}{n}\sum_{t=1}^{n}y(t)\{A_k\cos(k\alpha t+k\beta t^2)\right.$$

$$\left.+B_k\sin(k\alpha t+k\beta t^2)\}\right\}-\sum_{k=1}^{p}\frac{1}{2}\left(A_k^2+B_k^2\right).$$

Note that at $A_k=\tilde{A}_k$ and $B_k=\tilde{B}_k$:

$$J(\tilde{A}_1,\tilde{B}_1,\ldots,\tilde{A}_p,\tilde{B}_p,\alpha,\beta)=2I(\alpha,\beta).$$

Hence the estimator of $\boldsymbol{\theta}$ which maximizes $J(\boldsymbol{\theta})$ is equivalent to $\tilde{\boldsymbol{\theta}}$, the ALSE of $\boldsymbol{\theta}^{\mathbf{0}}$. Therefore, by Taylor series expansion and using the fact that $\mathbf{J}'(\tilde{\boldsymbol{\theta}})=0$:

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

Post-multiplying the above equation with $\mathbf{D}^{-1}$, we have:

$$(\tilde{\boldsymbol{\theta}}-\boldsymbol{\theta}^{\mathbf{0}})\mathbf{D}^{-1}=-\boldsymbol{J}'(\boldsymbol{\theta}^{\mathbf{0}})\mathbf{D}[\mathbf{D}\boldsymbol{J}''(\bar{\boldsymbol{\theta}}))\mathbf{D}]^{-1}. \quad (22)$$

Based on the following famous number theoretic conjecture (see Montgomery [11] or Lahiri [31]):

Conjecture : If $\theta_1,\theta_2,\theta_1',\theta_2'\in(0,\pi)$, then except for a countable number of points:

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

Additionally, if $\theta_2\neq\theta_2'$

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

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

$k=0,1,2,\ldots$, it can be shown that:

$$\lim_{n\to\infty}\boldsymbol{Q}'(\boldsymbol{\theta}^{\mathbf{0}})\mathbf{D}=-\lim_{n\to\infty}\boldsymbol{J}'(\boldsymbol{\theta}^{\mathbf{0}})\mathbf{D},\text{ and} \quad (23)$$

$$\lim_{n\to\infty}\mathbf{D}\boldsymbol{Q}''(\boldsymbol{\theta}^{\mathbf{0}})\mathbf{D}=-\lim_{n\to\infty}\mathbf{D}\boldsymbol{J}''(\boldsymbol{\theta}^{\mathbf{0}})\mathbf{D}. \quad (24)$$

Using (23) and (24) in (22), we have the following asymptotic equivalence:

$$(\tilde{\boldsymbol{\theta}}-\boldsymbol{\theta}^{\mathbf{0}})\mathbf{D}^{-1}\overset{a.eq.}{=}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}^{\mathbf{0}})\mathbf{D}^{-1}.$$

Hence, the result.

□

9 References

1 Abatzoglou, T. J., 1986 Fast maximnurm likelihood joint estimation of frequency and frequency rate. IEEE Transactions on Aerospace and Electronic Systems, Vol. 6, pp. 708-715.
2 Richards, F. SG., 1962. A method of maximum-likelihood estimation. Journal of the Royal Statistical Society. Series B (Methodological), pp.469-475.
3 Jennrich, R.I., 1969. Asymptotic properties of non-linear least squares estimators. The Annals of Mathematical Statistics, 40(2), pp.633-643.
4 Walker, A.M., 1971. On the estimation of a harmonic component in a time series with stationary independent residuals. Biometrika, 58(1), pp.21-36.
5 Kay, S.M. and Marple, S.L., 1981. Spectrum analysis— a modern perspective. Proceedings of the IEEE, 69(11), pp.1380-1419.
6 Wu, C.F., 1981. Asymptotic theory of nonlinear least squares estimation. The Annals of Statistics, pp.501-513.
7 Djuric, P. M., and Kay, S. M., 1990 Parameter estimation of chirp signals. IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. 38, Issue 12, pp. 2118-2126.
8 Peleg, S. and Porat, B., 1991 Linear FM signal parameter estimation from discrete-time observations. IEEE Transactions on Aerospace and Electronic Systems, Vol. 27, Issue 4, pp. 607-616.
9 Quinn, B.G. and Thomson, P.J., 1991. Estimating the frequency of a periodic function. Biometrika, Vol. 78, Issue 1, pp. 65-74.
10 Stoica, P., 1993. List of references on spectral line analysis. Signal Processing, 31(3), pp.329-340.
11 Montgomery, H.L., 1994. Ten lectures on the interface between analytic number theory and harmonic analysis (No. 84). American Mathematical Society.
12 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, 26, pp.547-578.
13 Li, H., Stoica, P. and Li, J., 2000. Computationally efficient parameter estimation for harmonic sinusoidal signals. Signal Processing, Vol. 80, Issue 9, pp. 1937-1944.
14 Quinn, B.G. and Hannan, E.J., 2001. The estimation and tracking of frequency (Vol. 9). Cambridge University Press.
15 Stylianou, Y., 2001. Applying the harmonic plus noise model in concatenative speech synthesis. IEEE Transactions on Speech and Audio Processing, 9(1), pp.21-29.
16 Saha, S. and Kay, S. M., 2002 Maximum likelihood parameter estimation of superimposed chirps using Monte Carlo importance sampling. IEEE Transactions on Signal Processing, Vol. 50, Issue 2, pp. 224-230.
17 Kundu, D. and Nandi, S., 2003. Determination of discrete spectrum in a random field. Statistica Neerlandica, 57(2), pp.258-284.
18 Chan, K.W. and So, H.C., 2004. Accurate frequency estimation for real harmonic sinusoids. IEEE Signal Processing Letters, Vol. 11, Issue 7, pp.609-612.
19 Myburg, F., Den Brinker, A.C. and Van Eijndhoven, S., 2004, April. Multicomponent chirp analysis in parametric audio coding. In Fourth IEEE Benelux Signal Processing Symposium.

20 Nandi, S. and Kundu, D., 2004. Estimating the fundamental frequency of a periodic function. Statistical Methods and Applications, Vol. 12, Issue 3, pp. 341-360.

21 Weruaga, L. and Képesi, M., 2007. The fan-chirp transform for non-stationary harmonic signals. Signal Processing, 87(6), pp.1504-1522.

22 E. E. Tsakonas, N. D. Sidiropoulos and A. Swami, 2008. Optimal Particle Filters for Tracking a Time-Varying Harmonic or Chirp Signal. IEEE Transactions on Signal Processing, vol. 56, no. 10, pp. 4598-4610, Oct. 2008, doi: 10.1109/TSP.2008.927462. s

23 Christensen, M.G., Stoica, P., Jakobsson, A. and Jensen, S.H., 2008. Multi-pitch estimation. Signal Processing, 88(4), pp.972-983.

24 Fuller, W.A., 2009. Introduction to statistical time series (Vol. 428). John Wiley & Sons Inc., 2nd ed. New York.

25 Pantazis, Y., Rosec, O. and Stylianou, Y., 2009, April. Chirp rate estimation of speech based on a time-varying quasi-harmonic model. In 2009 IEEE International Conference on Acoustics, Speech and Signal Processing (pp. 3985-3988). IEEE.

26 Fletcher, N.H. and Rossing, T.D., 2012. The physics of musical instruments. Springer Science & Business Media. Springer-Verlag New York.

27 Kundu, D. and Nandi, S., 2012. Statistical signal processing: frequency estimation. Springer Science & Business Media. Springer India.

28 Christensen, M.G. and Jensen, J.R., 2014, November. Pitch estimation for non-stationary speech. In 2014 48th Asilomar Conference on Signals, Systems and Computers. pp. 1400-1404.

29 Adalbjörnsson, S.I., Jakobsson, A. and Christensen, M.G., 2015. Multi-pitch estimation exploiting block sparsity. Signal Processing, 109, pp.236-247.

30 Doweck, Y., Amar, A. and Cohen, I., 2015. Joint Model Order Selection and Parameter Estimation of Chirps With Harmonic Components. IEEE Transactions on Signal Processing, 63(7), pp.1765-1778.

31 Lahiri, A., Kundu, D. and Mitra, A., 2015. Estimating the parameters of multiple chirp signals. Journal of Multivariate Analysis, 139, pp.189-206.

32 Doweck, Y., Amar, A. and Cohen, I., 2015. Fundamental initial frequency and frequency rate estimation of random-amplitude harmonic chirps. IEEE Transactions on Signal Processing, vol. 63, no. 23, pp. 6213-6228.

33 Nørholm, S.M., Jensen, J.R. and Christensen, M.G., 2016. Instantaneous fundamental frequency estimation with optimal segmentation for nonstationary voiced speech. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 24(12), pp.2354-2367.

34 Sward, J., Brynolfsson, J., Jakobsson, A. and Hansson-Sandsten, M., 2016. Sparse Semi-Parametric Estimation of Harmonic Chirp Signals. IEEE Transactions on Signal Processing, 64(7), pp.1798-1807.

35 Jensen, T.L., Nielsen, J.K., Jensen, J.R., Christensen, M.G. and Jensen, S.H., 2017. A fast algorithm for maximum-likelihood estimation of harmonic chirp parameters. IEEE Transactions on Signal Processing, 65(19), pp.5137-5152.

36 Nielsen, J.K., Jensen, T.L., Jensen, J.R., Christensen, M.G. and Jensen, S.H., 2017. Fast harmonic chirp summation. IEEE International Conference on, pp. 416-420.

37 Grover, R., Kundu, D. and Mitra, A., 2018. On approximate least squares estimators of parameters of one-dimensional chirp signal. Statistics, 52(5), pp.1060-1085.

38 Grover, R., Kundu, D. and Mitra, A., 2018. Approximate least squares estimators of a two-dimensional chirp model and their asymptotic properties. Journal of Multivariate Analysis, 168, pp.211-220.