On Chirp and Some Related Signals Analysis: A Brief Review and Some New Results *

Debasis Kundu [†]and Swagata Nandi [‡]

Abstract

Chirp signals have played an important role in the statistical signal processing literature. An extensive amount of work has been done in analyzing different one dimensional chirp, two dimensional chirp and some related signal processing models. The main aim of this article is to introduce the challenges associated with these problems to the statistical community with a hope that it will generate enough interests among the statisticians to contribute in this area of research. We provide a comprehensive review of different one dimensional, two dimensional, generalized chirp signals and some related models. We discuss some new chirp signal like models which can be used quite effectively for analyzing real life signals. Several open problems are discussed through out the paper for future research.

Key Words and Phrases: Chirp signals, sinusoidal signals; periodogram function; maximum likelihood estimators; MCMC model; least squares estimators; consistency; asymptotic normality.

AMS Subject Classifications: 62F10, 62F03, 62H12.

*This paper has been dedicated to Professor C.R.Rao on his birth centenary

[†]Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Pin 208016, India. E-mail: kundu@iitk.ac.in, Phone no. 91-512-2597141, Fax no. 91-512-2597500.

[‡]Theoretical Statistics and Mathematics Division, Indian Statistical Institute Delhi, 7 S.J.S. Sansanwal Marg, New Delhi 110016

1 Introduction

A real valued one-dimensional chirp signal in additive noise can be written mathematically as follows:

$$y(n) = A\cos(\alpha n + \beta n^2) + B\sin(\alpha n + \beta n^2) + X(n); \quad n = 1, \ldots, N. \tag{1}$$

Here $\{y(n)\}$ is a real valued signal observed at $n = 1, \ldots, N$, A, B are amplitudes, α, β are frequency and frequency rate, respectively. The additive noise $X(n)$ has mean zero, the explicit structure of $X(n)$ will be discussed later. This model arises in many applications of signal processing; one of the most important being the radar problem. For instance, consider a radar illuminating a target. Thus the transmitted signal will undergo a phase shift induced by the distance and relative motion between the target and the receiver. Assuming this motion to be continuous and differentiable, the phase shift can be adequately modelled as $\phi(t) = a_0 + a_1 t + a_2 t^2$, where the parameters a_1 and a_2 are either related to speed and acceleration or range and speed depending on what the radar is intended for, and on the kind of waveform transmitted, see for example Rihaczek [62] (page 56 – 65).

Chirp signals are encountered in many different engineering applications, particularly in radar, active sonar and passive sonar systems. The problem of parameter estimation of chirp signals has received a considerable amount of attention mainly in the engineering literature; see for example Abatzoglou [1], Djurić and Kay [12], Saha and Kay [64], Gini et al. [23], Besson et al. [7], Volcker and Otterstern [68], Lu et al. [48], Wang and Yang [69], Guo et al. [28], Yaron, Alon and Israel [76], Yang, Liu and Jiang [75], Fourier et al. [18], and see the references cited therein. A limited amount of work can be found in the statistical literature, interested readers are referred to the following; Nandi and Kundu [52], Kundu and Nandi [38], Lahiri, Kundu and Mitra [42, 44, 45], Lahiri and Kundu [41], Mazumder [49] and Grover, Kundu and Mitra [24, 25], Robertson, Gray and Woodward [63].

An alternative formulation of model (1) is also available mainly in the engineering literature, and that is as follows:

$$y(n) = Ae^{i(\alpha n + \beta n^2)} + X(n); \qquad n = 1, \ldots, N. \tag{2}$$

Here $y(n)$ is the complex valued signal; the amplitude A and the error component $X(n)$'s are also complex valued; α, β are same as defined before and $i = \sqrt{-1}$. Although, all physical signals are real valued, it might be advantageous from an analytical, a notational or an algorithmic point of view to work with signals in their analytic form which is complex valued, see for example Gabor [22]. For a real valued continuous signal, its analytic form can be easily obtained using the Hilbert transformation. Hence, it is equivalent to work either with model (1) or (2). In this paper, we concentrate on real valued chirp model only.

Observe that the chirp model (1) is a generalization of the well known sinusoidal frequency model;

$$y(n) = A\cos(\alpha n) + B\sin(\alpha n) + X(n); \quad n = 1, \ldots, N, \tag{3}$$

where, A, B, α and $X(n)$ are same as in (1). In sinusoidal frequency model, the frequency α is constant over time and does not change like chirp signal model (1). In chirp signal model, frequency changes linearly over time. A more general form of (3) can be written as

$$y(n) = \sum_{k=1}^{p} [A_k \cos(\alpha_k n) + B_k \sin(\alpha_k n)] + X(n); \qquad n = 1, \ldots, N. \tag{4}$$

The multicomponent sinusoidal frequency model (4) is mainly used to analyze any periodic signal. The main theoretical justification of using the sinusoidal model (4) to analyze any periodic signal comes from the Fourier decomposition of a periodic function. An extensive amount of work has been done on this model in time series and different other fields. Several books and papers have appeared on this model. For a comprehensive review on different statistical aspects associated with this model see the recent monograph by Nandi and Kundu [55]. It is clear that model (3) can be written as a special case of the model (1). Moreover, model (1) can be treated as a frequency modulated sinusoidal model of (3) also.

It may be mentioned that an alternative formulation of the model (4) similar to the chirp model also exists in the Statistical Signal Processing literature, and it is as follows:

$$y(n) = \sum_{k=1}^{p} A_k e^{i\alpha_k n} + X(n); \qquad n = 1, \ldots, N. \tag{5}$$

Here also $y(n)$ is the complex valued signal; the amplitude A_k's and the error component $X(n)$'s are complex valued similar to the complex chirp model and α_k's are the frequencies similar to the multiple sinusoidal model. The idea of using complex model is same as mentioned in case of chirp model. Professor Rao and his co-workers have worked quite extensively with this model in the mid 90s. They provided the consistency and asymptotic normality of the least squares estimators (LSEs) of the unknown parameters when the errors are independent identically distributed (i.i.d.) complex normal random variables. They also provided an efficient algorithm in Bai et al. [3], which can be used to compute the estimators in finite number of steps, and these estimators have the same efficiency as the least squares estimators. This is the first time an algorithm has been proposed which guarantees convergence in a finite number of steps. They have further provided in Bai et al. [4] a simultaneous estimation procedure of the number of components and the frequencies of the model (5) when some observations are missing. The detailed contributions of Professor Rao in the area of Statistical Signal Processing can be found in a recent article by Kundu [36].

The main aim of this article is to introduce the one-component chirp model (1), the more general multicomponent chirp model as defined below;

$$y(n) = \sum_{k=1}^{p} \left[A_k \cos(\alpha_k n + \beta_k n^2) + B_k \sin(\alpha_k n + \beta_k n^2)\right] + X(n); \quad n = 1, \ldots, N, \tag{6}$$

and some related models. We discuss different issues related with these models, mainly related to the estimation of the unknown parameters and establishing their properties. One major difficulty associated with chirp signal is the estimation of the unknown parameters. Although, model (1) can be seen as a non-linear regression model, the model does not satisfy the sufficient conditions of Jennrich [33] or Wu [71] required for establishing the consistency and asymptotic normality of LSEs of the unknown parameters of a standard non-linear regression model. Therefore, the consistency and the asymptotic normality of the LSEs or the maximum likelihood estimators (MLEs) are not immediate. Moreover, the chirp signal model is a highly non-linear model in its parameters, hence finding the MLEs or the LSEs becomes a non-trivial problem. An extensive amount of work has been done to compute different efficient estimators, and deriving their properties under different error assumptions. We provide different classical and Bayesian estimation procedures available till date and discuss their properties. We discuss some of the related models like random amplitude chirp model and harmonic chirp model which have received some attention in recent years. We further discuss one dimensional chirp like signal and it is observed that the chirp like model can conveniently be used in place of multicomponent chirp model in practice and is quite useful for analyzing real life signals. It seems the proposed chirp like model can be used as alternatives to one dimensional multicomponent chirp model.

The rest of the paper is organized as follows. In Section 2, we provide details of one dimensional chirp model. Two dimensional chirp and two dimensional polynomial phase models are presented in Section 3 and Section 4, respectively. In Section 5, some chirp related models are discussed. Finally, we conclude the paper in Section 6.

2 One Dimensional Chirp Model

In this section, first we discuss different issues involved with single chirp model and then the multiple chirp model will be discussed.

2.1 Single Chirp Model

The one dimensional single chirp model can be written as follows:

$$y(n) = A^0 \cos(\alpha^0 n + \beta^0 n^2) + B^0 \sin(\alpha^0 n + \beta^0 n^2) + X(n); \quad n = 1, \ldots, N. \tag{7}$$

Here $y(n)$ is the real valued signal as mentioned before, and it is observed at $n = 1, \ldots, N$. A^0, B^0 are real valued amplitudes; α^0, β^0 are frequency and frequency rate, respectively. The problem is to estimate the unknown parameters namely A^0, B^0, α^0 and β^0, based on the observed sample. Different methods have been proposed in the literature. We provide different methods of estimation and discuss theoretical properties of different estimators. First, it is assumed that the errors $X(n)$s are i.i.d. normal random variables with mean 0 and variance σ^2. A more general form of the error random variable $X(n)$ will be considered later on.

2.1.1 Maximum Likelihood Estimators

We need to obtain the log-likelihood function of the observed data to compute the MLEs. The log-likelihood function without the additive constant can be written as

$$l(\mathbf{\Theta}) = -\frac{N}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{n=1}^{N} \left(y(n) - A\cos(\alpha n + \beta n^2) - B \sin(\alpha n + \beta n^2)\right)^2, \tag{8}$$

where $\mathbf{\Theta} = (A, B, \alpha, \beta, \sigma^2)^\top$. Hence, the MLEs of A, B, α, β and σ^2, say $\widehat{A}$, $\widehat{B}$, $\widehat{\alpha}$, $\widehat{\beta}$ and $\widehat{\sigma}^2$, respectively, can be obtained by maximizing $l(\mathbf{\Theta})$, with respect to the unknown parameters. It is immediate that $\widehat{A}$, $\widehat{B}$, $\widehat{\alpha}$ and $\widehat{\beta}$ can be obtained by minimizing $Q(\mathbf{\Gamma})$, where, $\mathbf{\Gamma} = (A, B, \alpha, \beta)^\top$ and

$$Q(\mathbf{\Gamma}) = \sum_{n=1}^{N} \left(y(n) - A\cos(\alpha n + \beta n^2) - B \sin(\alpha n + \beta n^2)\right)^2, \tag{9}$$

with respect to the unknown parameters. Therefore, as it is expected under normality assumption, $\widehat{A}$, $\widehat{B}$, $\widehat{\alpha}$ and $\widehat{\beta}$ are the LSEs of the corresponding parameters also. Once, $\widehat{A}$, $\widehat{B}$,

$\widehat{\alpha}$ and $\widehat{\beta}$ are obtained, $\widehat{\sigma}^2$ can be obtained as

$$\widehat{\sigma}^2 = \frac{1}{N}\sum_{n=1}^{N}\left(y(n) - \widehat{A}\cos(\widehat{\alpha}n + \widehat{\beta}n^2) - \widehat{B}\sin(\widehat{\alpha}n + \widehat{\beta}n^2)\right)^2. \tag{10}$$

In the following, we first provide the procedure how to obtain $\widehat{A}$, $\widehat{B}$, $\widehat{\alpha}$ and $\widehat{\beta}$ and then we discuss about their asymptotic properties. Observe that, $Q(\boldsymbol{\Gamma})$ can be written as

$$Q(\boldsymbol{\Gamma}) = [\mathbf{Y} - \mathbf{W}(\alpha,\beta)\boldsymbol{\theta}]^\top [\mathbf{Y} - \mathbf{W}(\alpha,\beta)\boldsymbol{\theta}], \tag{11}$$

using the notation $\mathbf{Y} = (y(1),\ldots,y(N))^\top$ as the data vector, $\boldsymbol{\theta} = (A,B)^\top$, the linear parameter vector and

$$\mathbf{W}(\alpha,\beta) = \begin{bmatrix} \cos(\alpha+\beta) & \sin(\alpha+\beta) \\ \cos(2\alpha+4\beta) & \sin(2\alpha+4\beta) \\ \vdots & \vdots \\ \cos(N\alpha+N^2\beta) & \sin(N\alpha+N^2\beta) \end{bmatrix}, \tag{12}$$

the matrix with non-linear parameters. From (11), it is immediate that for fixed (α,β), the MLE of $\boldsymbol{\theta}$, say $\widehat{\boldsymbol{\theta}}(\alpha,\beta)$, can be obtained as

$$\widehat{\boldsymbol{\theta}}(\alpha,\beta) = \left[\mathbf{W}^\top(\alpha,\beta)\mathbf{W}(\alpha,\beta)\right]^{-1}\mathbf{W}^\top(\alpha,\beta)\mathbf{Y}, \tag{13}$$

and the MLEs of α and β can be obtained as

$$(\widehat{\alpha},\widehat{\beta}) = \arg\min_{\alpha,\beta} Q(\widehat{A}(\alpha,\beta),\widehat{B}(\alpha,\beta),\alpha,\beta). \tag{14}$$

Replacing $\boldsymbol{\theta}$ by $\widehat{\boldsymbol{\theta}}(\alpha,\beta)$ in (11), it can be easily seen that

$$(\widehat{\alpha},\widehat{\beta}) = \arg\max_{\alpha,\beta}\mathbf{Z}(\alpha,\beta), \tag{15}$$

where

$$\mathbf{Z}(\alpha,\beta) = \mathbf{Y}^\top\mathbf{W}(\alpha,\beta)[\mathbf{W}^\top(\alpha,\beta)\mathbf{W}(\alpha,\beta)]^{-1}\mathbf{W}^\top(\alpha,\beta)\mathbf{Y} = \mathbf{Y}^\top\mathbf{P}_{\mathbf{W}(\alpha,\beta)}\mathbf{Y}. \tag{16}$$

Here, $\mathbf{P}_{\mathbf{W}(\alpha,\beta)}$ is the projection matrix on the column space of $\mathbf{W}(\alpha,\beta)$. Then, the MLE of $\boldsymbol{\theta}$ can be obtained as $\widehat{\boldsymbol{\theta}} = \widehat{\boldsymbol{\theta}}(\widehat{\alpha},\widehat{\beta})$. Clearly, $\widehat{\alpha}$ and $\widehat{\beta}$ cannot be obtained analytically. Different numerical methods may be used to compute $\widehat{\alpha}$ and $\widehat{\beta}$. Saha and Kay [64] proposed to use the method of Pincus [59] in computing $\widehat{\alpha}$ and $\widehat{\beta}$, and they can be described as follows. Using the main theorem of Pincus [59], it follows that

$$\widehat{\alpha} = \lim_{c\to\infty} \frac{\int_0^\pi \int_0^\pi \alpha \exp(c\mathbf{Z}(\alpha,\beta)) d\beta d\alpha}{\int_0^\pi \int_0^\pi \exp(c\mathbf{Z}(\alpha,\beta)) d\beta d\alpha} \quad \text{and} \quad \widehat{\beta} = \lim_{c\to\infty} \frac{\int_0^\pi \int_0^\pi \beta \exp(c\mathbf{Z}(\alpha,\beta)) d\beta d\alpha}{\int_0^\pi \int_0^\pi \exp(c\mathbf{Z}(\alpha,\beta)) d\beta d\alpha}. \tag{17}$$

Therefore, one needs to compute two dimensional integration to compute the MLEs. Alternatively, importance sampling technique can be used quite effectively in this case to compute the MLEs of α and β using the following algorithm.

The following simple algorithm of the importance sampling method can be used to compute $\widehat{\alpha}$ and $\widehat{\beta}$ as defined in (17).

Algorithm 1:

Step 1: Generate $\alpha_1, \ldots, \alpha_M$ from uniform(0, π) and similarly, generate $\beta_1, \ldots, \beta_M$ from uniform(0, π).

Step 2: Consider a sequence of $\{c_k\}$, such that $c_1 < c_2 < c_3 < \ldots,$. For fixed $c = c_k$, compute

$$\widehat{\alpha}(c) = \frac{\frac{1}{M}\sum_{i=1}^M \alpha_i \exp(c\mathbf{Z}(\alpha_i,\beta_i))}{\frac{1}{M}\sum_{i=1}^M \exp(c\mathbf{Z}(\alpha_i,\beta_i))} \quad \text{and} \quad \widehat{\beta}(c) = \frac{\frac{1}{M}\sum_{i=1}^M \beta_i \exp(c\mathbf{Z}(\alpha_i,\beta_i))}{\frac{1}{M}\sum_{i=1}^M \exp(c\mathbf{Z}(\alpha_i,\beta_i))}.$$

Stop the iteration if the convergence takes place.

It may be noted that one can use the same α_n's and β_n's to compute $\widehat{\alpha}(c_k)$ and $\widehat{\beta}(c_k)$ for each k. Other methods like Newton-Raphson, Gauss-Newton or gradient simplex methods may be used to compute the MLEs in this case. But one needs very good initial guesses for any iterative process to converge. The $\mathbf{Z}(\alpha,\beta)$ surface for $0 < \alpha, \beta < \pi$ has several local maxima, hence any iterative process without very good initial guesses often converges to a local maximum rather than a global maximum.

Now we discuss the properties of the MLEs of the unknown parameters. Model (7) can be seen as a typical non-linear regression model with an additive error, where the errors are i.i.d. normal random variables. For different non-linear regression models, see, for example, Seber and Wild [65]. Jennrich [33] and Wu [71] developed several sufficient conditions under which the MLEs of the unknown parameters in a nonlinear regression model with an additive Gaussian errors are consistent and asymptotically normally distributed. Unfortunately, the sufficient conditions proposed by Jennrich [33] or Wu [71] do not hold here. Therefore, the consistency and the asymptotic normality of the MLEs are not immediate.

Nandi and Kundu [52] developed the consistency and asymptotic normality results of the MLEs under certain regularity conditions and they will be presented below. Interestingly, it is observed that the rate of convergence of the non-linear parameters are quite different than the linear ones. The results are presented in the following theorems, and for the proofs of these theorems, see Nandi and Kundu [52].

THEOREM 2.1 *If there exists a* K*, such that* $0 < |A^0|+|B^0| < K$, $0 < \alpha^0, \beta^0 < \pi$, *and* $\sigma > 0$, *then* $\widehat{\boldsymbol{\theta}} = (\widehat{A}, \widehat{B}, \widehat{\alpha}, \widehat{\beta}, \widehat{\sigma}^2)^\top$ *is a strongly consistent estimate of* $\boldsymbol{\theta}^0 = (A^0, B^0, \alpha^0, \beta^0, \sigma^2)^\top$. ∎

THEOREM 2.2 *Under the same assumptions as in Theorem 2.1,*

$$\left(N^{1/2}(\widehat{A} - A^0), N^{1/2}(\widehat{B} - B^0), N^{3/2}(\widehat{\alpha} - \alpha^0), N^{5/2}(\widehat{\beta} - \beta^0)\right)^\top \xrightarrow{d} \mathcal{N}_4(\mathbf{0}, 2\sigma^2\boldsymbol{\Sigma}), \qquad (18)$$

here $\xrightarrow{d}$ *means convergence in distribution,* $\mathcal{N}_4(\mathbf{0}, \boldsymbol{\Sigma})$ *means a 4-variate normal distribution with mean vector* $\mathbf{0}$ *and the dispersion matrix*

$$\boldsymbol{\Sigma} = \frac{2}{A^{0^2} + B^{0^2}} \begin{bmatrix} \frac{1}{2}(A^{0^2} + 9B^{0^2}) & -4A^0B^0 & -18B^0 & 15B^0 \\ -4A^0B^0 & \frac{1}{2}(9A^{0^2} + B^{0^2}) & 18A^0 & -15A^0 \\ -18B^0 & 18A^0 & 96 & -90 \\ 15B^0 & -15A^0 & -90 & 90 \end{bmatrix}. \qquad (19)$$

∎

Theorem 2.1 establishes the consistency of the MLEs, whereas Theorem 2.2 states the asymptotic distribution along with the rate of convergence of the MLEs. It is interesting to observe that the rates of the convergence of the MLEs of the linear parameters are significantly different than the corresponding non-linear parameters. The non-linear parameters are estimated more efficiently than the linear parameters for a given sample size. Moreover, as $A^{0^2}+B^{0^2}$ decreases, the asymptotic variances of the MLEs increase. Therefore, it may be noted that the asymptotic distribution of the MLEs can be easily used to construct asymptotic confidence intervals of the unknown parameters, or developing testing of hypotheses also.

Theorem 2.1 provides the consistency results of the MLEs under the boundedness assumptions on the linear parameters. The following open problem might be of interest.

Open Problem 1: Develop the consistency properties of the MLEs without any boundedness assumptions on the linear parameters. ∎

Theorem 2.2 provides the asymptotic distributions of the MLEs. The asymptotic distribution may be used to construct approximate confidence intervals of the unknown parameters, and also to develop testing of hypotheses problems. So far the construction of confidence intervals or testing of hypotheses problems have not been considered in the literature. The following open problem has some practical applications.

Open Problem 2: Construct different bootstrap confidence intervals like percentile bootstrap, biased corrected bootstrap etc. and compare them with the approximate confidence intervals based on asymptotic distributions of the MLEs. ∎

From Theorem 2.2, it is observed that the rate of convergence of the MLEs of the linear parameters is $N^{-1/2}$, whereas the rate of convergence of the MLEs for the frequency and frequency rate are, $N^{-3/2}$ and $N^{-5/2}$, respectively. Hence, the variances of the MLEs for the frequency and frequency rate converge faster than the corresponding variances of the linear parameters. It implies that the variance of the MLE of the frequency is larger than the variance of the MLE of the frequency rate. Hence, it is possible to estimate the frequency rate more efficiently than the frequency. Moreover, both these parameters can be estimated more efficiently than the linear parameters. These results are quite different than the usual rate of convergence of the MLEs for a general non-linear regression model, see for example Seber and Wild [65]. Intuitively, it can be explained as follows. Observe that the model (7) can be written as follows:

$$y(n) = f_n(A^0, B^0, \alpha^0, \beta^0) + X(n); \qquad n = 1, \ldots, N, \tag{20}$$

where

$$f_n(A^0, B^0, \alpha^0, \beta^0) = A^0 \cos(\alpha^0 n + \beta^0 n^2) + B^0 \sin(\alpha^0 n + \beta^0 n^2).$$

Now to obtain the asymptotic variances of the MLEs we need proper normalization factors of the following quantities:

$$\sum_{n=1}^{N} \frac{\partial f_n(A, B, \alpha, \beta)}{\partial A}, \quad \sum_{n=1}^{N} \frac{\partial f_n(A, B, \alpha, \beta)}{\partial B}, \quad \sum_{n=1}^{N} \frac{\partial f_n(A, B, \alpha, \beta)}{\partial \alpha}, \quad \sum_{n=1}^{N} \frac{\partial f_n(A, B, \alpha, \beta)}{\partial \beta},$$

so that they converge to non-zero limits. Now in the first two cases the normalizing factor is N^{-1}, where as the third and fourth cases they are N^{-3} and N^{-5}, respectively. Due to this reason the variances of the MLEs of A^0 and B^0 are of the order N^{-1}, where as for α^0 and β^0, they are N^{-3} and N^{-5}, respectively. It may be mentioned for a typical non-linear regression model the normalizing factor is usually N^{-1}, hence the asymptotic variance of the MLE becomes of the order N^{-1}, see Seber and Wild [65] in this respect. Although, we have provided the results for the MLEs when the errors are normally distributed, it is clear that the consistency and the asymptotic normality results are valid for the general LSEs when the errors are i.i.d. random variables with mean zero and finite variance σ^2. Now we discuss the properties of the LSEs for more general errors.

2.1.2 Least Squares Estimators

In the last section, we have discussed about the estimation of the unknown parameters when the errors are normally distributed. It may be mentioned that when the errors are normally distributed, then the LSEs and the MLEs are same. In this section, we discuss about the estimation of the unknown parameters when the errors may not be normally distributed. The following assumption on $X(n)$ has been made.

Assumption 1 *Suppose*

$$X(n) = \sum_{j=-\infty}^{\infty} a(j)e(n-j), \tag{21}$$

where $\{e(n)\}$ is a sequence of i.i.d. random variables with mean zero and finite fourth moment, and

$$\sum_{j=-\infty}^{\infty} |a(j)| < \infty. \tag{22}$$

It may be mentioned that Assumption 1 is a standard assumption for a stationary linear process, and any finite dimensional stationary auto regressive (AR), moving average (MA) or auto regressive moving average (ARMA) process can be represented as (21), when $a(j)$s are absolutely summable, that is, satisfy (22).

Kundu and Nandi [38] first discussed the estimation of the unknown parameters of model (1) when $X(n)$ satisfies Assumption 1. Clearly, the MLEs are not possible to obtain as the exact distribution of $X(n)$ is not known. Kundu and Nandi [38] considered the LSEs of the unknown parameters and they can be obtained by minimizing $Q(\mathbf{\Gamma})$ as defined in (9), with respect to $\mathbf{\Gamma}$. Hence, the same computational procedure which has been used to compute the MLEs can be used here also. But the consistency and the asymptotic normality of the LSEs need to be developed independently. Kundu and Nandi [38] developed the following two results similar to Theorems 2.1 and 2.2.

THEOREM 2.3 *If there exists a K, such that $0 < |A^0| + |B^0| < K$, $0 < \alpha^0, \beta^0 < \pi$, $\sigma^2 > 0$ and $X(n)$ satisfies Assumption 1, then $\widehat{\mathbf{\Gamma}} = (\widehat{A}, \widehat{B}, \widehat{\alpha}, \widehat{\beta})^\top$ is a strongly consistent estimate of $\mathbf{\Gamma}^0 = (A^0, B^0, \alpha^0, \beta^0)^\top$.* ∎

THEOREM 2.4 *Under the same assumptions as in Theorem 2.3,*

$$\Big(N^{1/2}(\widehat{A} - A^0), N^{1/2}(\widehat{B} - B^0), N^{3/2}(\widehat{\alpha} - \alpha^0), N^{5/2}(\widehat{\beta} - \beta^0)\Big)^\top \xrightarrow{d} \mathcal{N}_4(\mathbf{0}, 2c\sigma^2\mathbf{\Sigma}), \tag{23}$$

here $\mathbf{\Sigma}$ is same as defined in Theorem 2.2, and $c = \sum_{j=-\infty}^{\infty} a^2(j)$. ∎

Note that Theorem 2.4 provides the rates of convergence of the LSEs under the assumption of stationary linear process on the error. This asymptotic distribution stated in Theorem 2.4 can also be used to construct asymptotic confidence intervals of the unknown parameters and in developing testing of hypotheses provided one can obtain a good estimate of $c\sigma^2$. In this respect, the method proposed by Nandi and Kundu [53] may be used. The following open problems will be of interest.

Open Problem 3: Construct confidence intervals based on the asymptotic distributions of the LSEs and develop different bootstrap confidence intervals also. ∎

Open Problem 4: Develop different testing of hypotheses based on the LSEs associated with the chirp model parameters. ∎

Now we will discuss the properties of the least absolute deviation (LAD) estimators, which are more robust than the LSEs, if the errors are heavy tailed or there are outliers in the data.

2.1.3 Least Absolute Deviation Estimators

In this section, we discuss about LAD estimators of A^0, B^0, α^0 and β^0. The LAD estimators of the unknown parameters can be obtained by minimizing

$$R(\mathbf{\Gamma}) = \sum_{n=1}^{N} \left| y(n) - A\cos(\alpha n + \beta n^2) - B\sin(\alpha n + \beta n^2) \right| . \tag{24}$$

Here $\mathbf{\Gamma}$ is same as defined before. With the abuse of notations, let us denote the LAD estimators of A^0, B^0, α^0 and β^0, by $\widehat{A}, \widehat{B}, \widehat{\alpha}$ and $\widehat{\beta}$, respectively. They can be obtained as

$$(\widehat{A}, \widehat{B}, \widehat{\alpha}, \widehat{\beta}) = \arg\min R(\mathbf{\Gamma}). \tag{25}$$

The minimization of $R(\mathbf{\Gamma})$ with respect to unknown parameters is a challenging problem. Even when α and β are known, unlike the LSEs of A and B, the LAD estimators of A and B cannot be obtained in closed form. For known α and β, the LAD estimators of A and B can be obtained by solving a $2N$ dimensional linear programming problem, which is quite intensive computationally, see for example Kennedy and Gentle [35]. Hence, finding the LAD estimators of $\mathbf{\Gamma}$ efficiently, is a challenging problem. Assume that $-K \leq A, B \leq K$, for some $K > 0$, then grid search method may be used to compute the LAD estimators in this case. But clearly, it is also a very computationally intensive method. The importance sampling technique as it was used to find the MLEs can be used here also, and the LAD estimators can be obtained as follows.

$$\begin{aligned}
\widehat{A} &= \lim_{c\to\infty} \frac{\int_{-K}^{K}\int_{-K}^{K}\int_0^{\pi}\int_0^{\pi} A\exp(cR(\mathbf{\Gamma}))d\beta d\alpha dA dB}{\int_{-K}^{K}\int_{-K}^{K}\int_0^{\pi}\int_0^{\pi} \exp(cR(\mathbf{\Gamma}))d\beta d\alpha dA dB} \\
\widehat{B} &= \lim_{c\to\infty} \frac{\int_{-K}^{K}\int_{-K}^{K}\int_0^{\pi}\int_0^{\pi} B\exp(cR(\mathbf{\Gamma}))d\beta d\alpha dA dB}{\int_{-K}^{K}\int_{-K}^{K}\int_0^{\pi}\int_0^{\pi} \exp(cR(\mathbf{\Gamma}))d\beta d\alpha dA dB} \\
\widehat{\alpha} &= \lim_{c\to\infty} \frac{\int_{-K}^{K}\int_{-K}^{K}\int_0^{\pi}\int_0^{\pi} \alpha\exp(cR(\mathbf{\Gamma}))d\beta d\alpha dA dB}{\int_{-K}^{K}\int_{-K}^{K}\int_0^{\pi}\int_0^{\pi} \exp(cR(\mathbf{\Gamma}))d\beta d\alpha dA dB} \\
\widehat{\beta} &= \lim_{c\to\infty} \frac{\int_{-K}^{K}\int_{-K}^{K}\int_0^{\pi}\int_0^{\pi} \beta\exp(cR(\mathbf{\Gamma}))d\beta d\alpha dA dB}{\int_{-K}^{K}\int_{-K}^{K}\int_0^{\pi}\int_0^{\pi} \exp(cR(\mathbf{\Gamma}))d\beta d\alpha dA dB}.
\end{aligned}$$

Therefore, in this case also, using similar algorithm as Algorithm 1, one can obtain the LAD estimators of the unknown parameters. It will be interesting to see performance of the above estimators based on simulations. More work is needed in that direction. Efficient technique of computing the LAD estimators is not available till date, but Lahiri, Kundu and Mitra [44] established the asymptotic properties of the LAD estimators under some regularity conditions. The results can be stated as follows, the proofs are available in Lahiri, Kundu and Mitra [44] and Lahiri [40].

Assumption 2 *The error random variable $\{X(n)\}$ is a sequence of i.i.d. random variables with mean zero, variance σ^2, and it has a probability density function (PDF) $f(x)$. The PDF $f(x)$ is symmetric and differentiable in $(0, \epsilon)$ and $(-\epsilon, 0)$, for some $\epsilon > 0$, and $f(0) > 0$.*

THEOREM 2.5 *If there exists an K, such that $0 < |A^0| + |B^0| < K$, $0 < \alpha^0, \beta^0 < \pi$, $\sigma^2 > 0$, and $\{X(n)\}$ satisfies Assumption 2, then $\widehat{\mathbf{\Gamma}} = (\widehat{A}, \widehat{B}, \widehat{\alpha}, \widehat{\beta})^\top$ is a strongly consistent estimate of $\mathbf{\Gamma}^0$.* ∎

THEOREM 2.6 *Under the same assumptions as in Theorem 2.5,*

$$\Big(N^{1/2}(\widehat{A} - A^0), N^{1/2}(\widehat{B} - B^0), N^{3/2}(\widehat{\alpha} - \alpha^0), N^{5/2}(\widehat{\beta} - \beta^0)\Big)^\top \xrightarrow{d} \mathcal{N}_4(\mathbf{0}, \frac{1}{f^2(0)}\mathbf{\Sigma}), \qquad (26)$$

here $\mathbf{\Sigma}$ is same as defined in (18). ∎

Therefore, it is observed that the LSEs and LAD estimators both provide consistent estimators of the unknown parameters, and they have the same rate of convergence. Similarly, as in Theorem 2, Theorem 6 also can be used to construct approximate confidence intervals of the unknown parameters and to develop different testing of hypotheses provided one can obtain a good estimate of $f^2(0)$. It will be important to see the performances of the approximate confidence intervals in terms of the coverage percentages and confidence lengths. The following problems related to LAD estimators will be of interest.

Open Problem 5: Develop an efficient algorithm to compute the LAD estimators for the chirp parameters under different error assumptions.

Open Problem 6: Construct confidence intervals of the unknown parameters of the chirp model and develop the testing of hypotheses based on the LAD estimators.

Open Problem 7: Develop the properties of the LAD estimators for multicomponent chirp model.

2.1.4 Finite Step Efficient Algorithm

It has been observed that the MLEs, LSEs and LAD estimators provide consistent estimators of the frequency and frequency rate, but all of them need to be computed by using some optimization technique or the importance sampling method described before. Any optimization method involving non-linear model needs to be solved using some iterative procedure, and any iterative procedure has its own problem of convergence. In this section, we present an algorithm where the frequency and frequency rate estimators can be obtained by using an iterative procedure. The interesting point of this algorithm is that it is known that it converges in a fixed number of iterations, and at the same time both the frequency and frequency rate estimators attain the same rate of convergence as the MLEs or LAD estimators. The main idea of this algorithm came from a finite step efficient algorithm in case of sinusoidal signals, proposed by Bai et al. [3]. It is observed that if we start the initial guesses of α^0 and β^0 with convergence rates $O_p(N^{-1})$ and $O_p(N^{-2})$, respectively, then after four iterations, the algorithm produces an estimate of α^0 with convergence rate $O_p(N^{-3/2})$, and an estimate of β^0 with convergence rate $O_p(N^{-5/2})$. Before providing the algorithm in details first we show how to improve the estimators of α^0 and β^0. Then, the exact computational algorithm for practical implementation will be provided. If $\widetilde{\alpha}$ is an estimator of α^0, such that for $\delta_1 > 0$, $\widetilde{\alpha} - \alpha^0 = O_p(N^{-1-\delta_1})$, and $\widetilde{\beta}$ is an estimator of β^0, such that for $\delta_2 > 0$,

$\widetilde{\beta} - \beta^0 = O_p(N^{-2-\delta_2})$, then the improved estimators of α^0 and β^0, can be obtained as

$$\widetilde{\widetilde{\alpha}} = \widetilde{\alpha} + \frac{48}{N^2} Im\left(\frac{P_N^\alpha}{Q_N}\right) \tag{27}$$

$$\widetilde{\widetilde{\beta}} = \widetilde{\beta} + \frac{45}{N^4} Im\left(\frac{P_N^\beta}{Q_N}\right), \tag{28}$$

respectively, where

$$P_N^\alpha = \sum_{n=1}^{N} y(n)\left(n - \frac{N}{2}\right) e^{-i(\widetilde{\alpha}n + \widetilde{\beta}n^2)},$$

$$P_N^\beta = \sum_{n=1}^{N} y(n)\left(n^2 - \frac{N^2}{3}\right) e^{-i(\widetilde{\alpha}n + \widetilde{\beta}n^2)},$$

$$Q_N^\alpha = \sum_{n=1}^{N} y(n) e^{-i(\widetilde{\alpha}n + \widetilde{\beta}n^2)}.$$

Here if C is a complex number, then $Im(C)$ means the imaginary part of C.

The following two theorems provide the justification for the improved estimators, and the proofs of these two theorems can be obtained in Lahiri, Kundu and Mitra [42].

THEOREM 2.7 *If* $\widetilde{\alpha} - \alpha^0 = O_p(N^{-1-\delta_1})$ *for* $\delta_1 > 0$, *then*

$$(a)\ (\widetilde{\widetilde{\alpha}} - \alpha^0) = O_p(N^{-1-2\delta_1}) \quad if \quad \delta_1 \le 1/4,$$
$$(b)\ N^{3/2}(\widetilde{\widetilde{\alpha}} - \alpha^0) \xrightarrow{d} \mathcal{N}(0, \sigma_1^2) \quad if \quad \delta_1 > 1/4,$$

where $\sigma_1^2 = \dfrac{384\sigma^2}{A^{02} + B^{02}}$. ∎

THEOREM 2.8 *If* $\widetilde{\beta} - \beta^0 = O_p(N^{-2-\delta_2})$ *for* $\delta_2 > 0$, *then*

$$(a)\ (\widetilde{\widetilde{\beta}} - \beta^0) = O_p(N^{-2-2\delta_2}) \quad if \quad \delta_2 \le 1/4,$$
$$(b)\ N^{5/2}(\widetilde{\widetilde{\beta}} - \beta^0) \xrightarrow{d} \mathcal{N}(0, \sigma_2^2) \quad if \quad \delta_1 > 1/4,$$

where $\sigma_2^2 = \dfrac{360\sigma^2}{A^{02} + B^{02}}$. ∎

Now we show that starting from initial guesses $\widetilde{\alpha}$, $\widetilde{\beta}$ with convergence rates $\widetilde{\alpha} - \alpha^0 = O_p(N^{-1})$ and $\widetilde{\beta} - \beta^0 = O_p(N^{-2})$, respectively, how the above procedure can be used to obtain efficient estimators. It may be noted that finding initial guesses with the above convergence rates is not difficult. It can be obtained by finding the minimum of $Q(\widehat{A}(\alpha,\beta),\widehat{B}(\alpha,\beta),\alpha,\beta)$ over the grid $\left(\frac{\pi j}{N},\frac{\pi k}{N^2}\right)$; $j=1,\ldots,N$ and $k=1,\ldots,N^2$. The main idea is not to use the whole sample at the beginning, as it was originally suggested by Bai et al. [3]. A part of the sample is used at the beginning, and we gradually proceed towards the complete sample. With varying sample size, more and more data points are used with the increasing number of iteration. The algorithm can be described as follows. Denote the estimates of α^0 and β^0 obtained at the j-th iteration as $\widetilde{\alpha}^{(j)}$ and $\widetilde{\beta}^{(j)}$, respectively.

Algorithm 2:

Step 1: Choose $N_1 = N^{8/9}$. Therefore, $\widetilde{\alpha}^{(0)} - \alpha^0 = O_p(N^{-1}) = O_p(N_1^{-1-1/8})$ and $\widetilde{\beta}^{(0)} - \beta^0 = O_p(N^{-2}) = O_p(N_1^{-2-1/4})$. Perform steps (27) and (28). Therefore, after 1-st iteration, we have

$$\widetilde{\alpha}^{(1)} - \alpha^0 = O_p(N_1^{-1-1/4}) = O_p(N^{-10/9}) \quad \text{and} \quad \widetilde{\beta}^{(1)} - \beta^0 = O_p(N_1^{-2-1/2}) = O_p(N^{-20/9}).$$

Step 2: Choose $N_2 = N^{80/81}$. Therefore, $\widetilde{\alpha}^{(1)} - \alpha^0 = O_p(N_2^{-1-1/8})$ and $\widetilde{\beta}^{(1)} - \beta^0 = O_p(N_2^{-2-1/4})$. Perform steps (27) and (28). Therefore, after 2-nd iteration, we have

$$\widetilde{\alpha}^{(2)} - \alpha^0 = O_p(N_2^{-1-1/4}) = O_p(N^{-100/81}) \quad \text{and} \quad \widetilde{\beta}^{(2)} - \beta^0 = O_p(N_2^{-2-1/2}) = O_p(N^{-200/81}).$$

Step 3: Choose $N_3 = N$. Therefore, $\widetilde{\alpha}^{(2)} - \alpha^0 = O_p(N_3^{-1-19/81})$ and $\widetilde{\beta}^{(2)} - \beta^0 = O_p(N_3^{-2-38/81})$. Perform steps (27) and (28). Therefore, after 3-rd iteration, we have

$$\widetilde{\alpha}^{(3)} - \alpha^0 = O_p(N^{-1-38/81}) \quad \text{and} \quad \widetilde{\beta}^{(3)} - \beta^0 = O_p(N^{-2-76/81}).$$

Step 4: Choose $N_4 = N$ and perform steps (27) and (28). Now we obtain the required convergence rates, i.e.

$$\widetilde{\alpha}^{(4)} - \alpha^0 = O_p(N^{-3/2}) \quad \text{and} \quad \widetilde{\beta}^{(4)} - \beta^0 = O_p(N^{-5/2}).$$

The above algorithm can be used quite efficiently to compute estimators in four steps which are equivalent to the LSEs. It may be mentioned that the fraction of the sample sizes which have been used in each step is not unique. It is possible to obtain equivalent estimators with different choices. Although they are asymptotically equivalent, the finite sample performances might be different. It might be interesting to compare the finite sample performances of the different estimators by extensive Monte Carlo simulations.

2.1.5 Approximate Least Squares Estimators

Recently, Grover, Kundu and Mitra [24] proposed a periodogram like estimators of the unknown parameters of a chirp signal. Consider the periodogram like function defined for the chirp signal as follows;

$$I(\alpha,\beta) = \frac{2}{N}\left\{\left(\sum_{n=1}^{N} y(n)\cos(\alpha n + \beta n^2)\right)^2 + \left(\sum_{n=1}^{N} y(n)\sin(\alpha n + \beta n^2)\right)^2\right\}. \tag{29}$$

Recall that the periodogram function for the sinusoidal signal (3) is defined as follows

$$I(\alpha) = \frac{2}{N}\left\{\left(\sum_{n=1}^{N} y(n)\cos(\alpha n)\right)^2 + \left(\sum_{n=1}^{N} y(n)\sin(\alpha n)\right)^2\right\}. \tag{30}$$

The periodogram estimator or the approximate least squares estimator (ALSE) of α can be obtained by maximizing $I(\alpha)$ as defined in (30) over the range $(0,\pi)$. If $y(n)$ satisfies the sinusoidal model assumption (3), then the ALSE of α is consistent and asymptotically equivalent to the LSE of α of model (3). Therefore, (29) is a natural generalization of (30) to compute the estimators of the frequency and the frequency rate of the one component chirp model (1).

Grover, Kundu and Mitra [24] proposed the approximate maximum likelihood estimators (AMLEs) of α and β as

$$(\widetilde{\alpha},\widetilde{\beta}) = \text{arg max}_{\alpha,\beta} I(\alpha,\beta), \tag{31}$$

under the normality assumption on the error $X(n)$. Once $\widetilde{\alpha}$ and $\widetilde{\beta}$ are obtained, $\widetilde{A}$ and $\widetilde{B}$ can be calculated as

$$\widetilde{A} = \frac{2}{N}\sum_{n=1}^{N}\cos(\widetilde{\alpha}n + \widetilde{\beta}n^2), \quad \widetilde{B} = \frac{2}{N}\sum_{n=1}^{N}\sin(\widetilde{\alpha}n + \widetilde{\beta}n^2). \tag{32}$$

A and B can also be estimated using $\widehat{\boldsymbol{\theta}}(\widetilde{\alpha}, \widetilde{\beta})$. In order to compute the AMLEs one can use Algorithm 1 by replacing $\mathbf{Z}(\alpha, \beta)$ with $I(\alpha, \beta)$. Grover, Kundu and Mitra [24] established the asymptotic properties of the ALSEs. It is observed that the ALSEs also have the consistency and asymptotic normality properties as the LSEs but in a slightly weaker condition. The following results can be found in Grover, Kundu and Mitra [24].

THEOREM 2.9 *If* $0 < |A^0| + |B^0|$, $0 < \alpha^0, \beta^0 < \pi$, *and* $X(n)$ *satisfies Assumption 1, then* $\widetilde{\boldsymbol{\Gamma}} = (\widetilde{A}, \widetilde{B}, \widetilde{\alpha}, \widetilde{\beta})^\top$ *is a strongly consistent estimate of* $\boldsymbol{\Gamma}^0 = (A^0, B^0, \alpha^0, \beta^0)^\top$. ∎

THEOREM 2.10 *Under the same assumptions as in Theorem 2.9,*

$$\Big(N^{1/2}(\widetilde{A} - A^0), N^{1/2}(\widetilde{B} - B^0), N^{3/2}(\widetilde{\alpha} - \alpha^0), N^{5/2}(\widetilde{\beta} - \beta^0)\Big)^\top \xrightarrow{d} \mathcal{N}_4(\mathbf{0}, 2\sigma^2 c\boldsymbol{\Sigma}). \tag{33}$$

here c *and* $\boldsymbol{\Sigma}$ *are same as defined in Theorem 2.4.* ∎

Comparing Theorem 2.3 and Theorem 2.9, it may be observed that in Theorem 2.9, the boundedness condition on the linear parameters has been removed. By extensive simulation experiments it has been observed that computationally the ALSEs have slight advantage over the LSEs, although the mean squared errors (MSEs) of the ALSEs are slightly more than the corresponding LSEs. Again, as in the case of LSEs, Theorem 2.10 can also be used for construction of confidence intervals of the unknown parameters and for different testing of hypotheses provided one can obtain a good estimate of $\sigma^2 c$. The following problem will be of interest.

Open Problem 8: Construct confidence intervals of the unknown parameters based on LSEs and ALSEs and compare them in terms of average lengths of the confidence intervals and their respective coverage percentages.

2.1.6 Bayes Estimates

Recently, Mazumder [49] considered model (1) when $X(n)$'s are i.i.d. normal random variables and provided a Bayesian solution. The author provided the Bayes estimates and the associated credible intervals of the unknown parameters. The following transformations and assumptions have been made.

$$A^0 = r^0 \cos(\theta^0), \quad B^0 = r^0 \sin(\theta^0), \quad r^0 \in (0, K], \quad \theta^0 \in [0, 2\pi], \quad \alpha, \beta \in (0, \pi).$$

The following prior assumptions distributions have been made on the above unknown parameters.

$$r \sim \text{uniform}(0, K) \tag{34}$$

$$\theta \sim \text{uniform}(0, 2\pi) \tag{35}$$

$$\alpha \sim \text{vonMises}(a_0, a_1) \tag{36}$$

$$\beta \sim \text{vonMises}(b_0, b_1) \tag{37}$$

$$\sigma^2 \sim \text{inverse gamma}(c_0, c_1). \tag{38}$$

It may be mentioned that r and θ have non-informative priors, σ^2 has a conjugate prior. In this case, 2α and 2β are circular random variables, that is why, von Misses distribution, the natural analog of the normal distribution in circular data has been considered. Let us denote the prior densities of r, θ, α, β, σ^2 as $[r]$, $[\theta]$, $[\alpha]$, $[\beta]$, $[\sigma^2]$, respectively, and $\boldsymbol{Y}$ is the data vector as defined in Section 2.1.1. If it is assumed that the priors are independently distributed, then the joint posterior density function of r, θ, α, β, σ^2 can be obtained as

$$[r, \theta, \alpha, \beta, \sigma^2 | \boldsymbol{Y}] \propto [r][\theta][\alpha][\beta][\sigma^2][\boldsymbol{Y} | r, \theta, \alpha, \beta, \sigma^2].$$

Now to compute the Bayes estimates of the unknown parameters using the Gibbs sampling technique one needs to compute conditional distribution of each parameter given all the parameters, known as the full conditional distribution, denoted by $[\cdot|\ldots]$, and they are given by

$$[r|\ldots] \propto [r][\boldsymbol{Y}|r,\theta,\alpha,\beta,\sigma^2]$$

$$[\theta|\ldots] \propto [\theta][\boldsymbol{Y}|r,\theta,\alpha,\beta,\sigma^2]$$

$$[\alpha|\ldots] \propto [\alpha][\boldsymbol{Y}|r,\theta,\alpha,\beta,\sigma^2]$$

$$[\beta|\ldots] \propto [\beta][\boldsymbol{Y}|r,\theta,\alpha,\beta,\sigma^2]$$

$$[\sigma^2|\ldots] \propto [\sigma^2][\boldsymbol{Y}|r,\theta,\alpha,\beta,\sigma^2].$$

The closed form expression of the full conditionals cannot be obtained. Mazumder [49] proposed to use the random walk Markov Chain Monte Carlo (MCMC) technique to update these parameters. The author has used this method to predict future observation also.

2.1.7 Testing of Hypothesis

Recently, Kundu and Das [11] considered the following testing of hypothesis problem for single chirp model. They have considered the model (1) and it is assumed that $X(n)$'s are i.i.d. random variables with mean zero and finite variance σ^2. Some more assumptions are needed in developing the properties of the test statistics, and those will be explicitly mentioned later. If we denote the vector $\boldsymbol{\theta}^0 = (A^0, B^0, \alpha^0, \beta^0)$, then we want to test the following hypothesis

$$H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}^0 \qquad vs. \qquad H_1 : \boldsymbol{\theta} \neq \boldsymbol{\theta}^0. \tag{39}$$

This is a typical testing of hypothesis problem, and it mainly tests whether the data are coming from a specific chirp model or not.

Four different tests to test the hypothesis (39) have been proposed by Dhar, Kundu and Das [11] based on the following test statistics:

$$T_{N,1} = ||\mathbf{D}^{-1}(\widehat{\boldsymbol{\theta}}_{N,LSE} - \boldsymbol{\theta}^0||_2^2, \tag{40}$$

$$T_{N,2} = ||\mathbf{D}^{-1}(\widehat{\boldsymbol{\theta}}_{N,LAD} - \boldsymbol{\theta}^0||_2^2, \tag{41}$$

$$T_{N,3} = ||\mathbf{D}^{-1}(\widehat{\boldsymbol{\theta}}_{N,LSE} - \boldsymbol{\theta}^0||_1^2, \tag{42}$$

$$T_{N,4} = ||\mathbf{D}^{-1}(\widehat{\boldsymbol{\theta}}_{N,LAD} - \boldsymbol{\theta}^0||_1^2. \tag{43}$$

Here, $\widehat{\boldsymbol{\theta}}_{N,LSE}$ and $\widehat{\boldsymbol{\theta}}_{N,LAD}$ denote the LSE of $\boldsymbol{\theta}$ and LAD estimate of $\boldsymbol{\theta}$ as discussed in Section 2.1.2 and Section 2.1.3, respectively. The 4×4 diagonal matrix $\mathbf{D}$ is as follows:

$$\mathbf{D} = diag\{N^{-1/2}, N^{-1/2}, N^{-3/2}, N^{-5/2}\}.$$

Further, $||\cdot||_2$ and $||\cdot||_1$ denote the Euclidean and L_1 norms, respectively. Note that all these test statistics are based on some normalized values of the distances between the estimates and the parameter value under the null hypothesis. Here, Euclidean and L_1 distances have been chosen, but any other distance function also can be considered. In all these cases, clearly the null hypothesis will be rejected if the values of the test statistics are large.

Now to choose the critical values of the test statistics, the following assumptions are made. Other than the assumptions required on the parameter values defined in Theorem 2.3, the following error assumptions are also required.

Assumption 3 *The i.i.d. error random variables have the positive density function $f(\cdot)$ with finite second moment.*

Assumption 4 *Let F_n be the distribution function of $y(n)$ with the probability density function $f_n(y, \boldsymbol{\theta})$, which is twice continuously differentiable with respect to $\boldsymbol{\theta}$. It is assumed that $E\left[\frac{\partial}{\partial\theta_i} f_n(y, \boldsymbol{\theta})\right]^{2+\delta}_{\boldsymbol{\theta}=\boldsymbol{\theta}^0} < \infty$, for some $\delta > 0$ and $E\left[\frac{\partial^2}{\partial\theta_i\partial\theta_j} f_n(y, \boldsymbol{\theta})\right]^2_{\boldsymbol{\theta}=\boldsymbol{\theta}^0} < \infty$, for all $n = 1, 2, \ldots, N$. Here θ_i and θ_j, for $1 \leq i, j \leq 4$, are the i-th and j-th component of $\boldsymbol{\theta}$.*

Note that Assumptions 3 and 4 are not very unnatural. The Assumption 3 holds for most of the well known probability density functions, e.g. normal, Laplace and Cauchy probability density functions. The smoothness assumptions in Assumption 4 is required to prove the asymptotic normality of the test statistics under contiguous alternatives. Such assumptions are quite common in general across the asymptotic statistics.

Now we are in a position to provide the asymptotic properties of the above test statistics. We use the following notations. Suppose $\boldsymbol{A} = (A_1, A_2, A_3, A_4)^\top$ is a four dimensional Gaussian random vector with mean vector $\mathbf{0}$ and the dispersion matrix $\boldsymbol{\Sigma}_1 = 2\sigma^2\boldsymbol{\Sigma}$, where $\boldsymbol{\Sigma}$ is same as defined in (19). Further, let $\boldsymbol{B} = (B_1, B_2, B_3, B_4)^\top$ be also a four dimensional Gaussian random vector with mean vector $\mathbf{0}$ and the dispersion matrix $\boldsymbol{\Sigma}_2 = \frac{1}{\{f(M)\}^2}\boldsymbol{\Sigma}$. Here, M denotes the median of the distribution function associated with the density function $f(\cdot)$. Then we have the following results.

Theorem 2.11 *Let $c_{1\eta}$ be $(1-\eta)$-th quantile of the distribution of $\sum_{i=1}^{4}\lambda_i Z_i^2$, where λ_is are the eigen values of $\boldsymbol{\Sigma}_1$, as defined above, and Z_is are independent standard normal random variables. If the assumption required on the parameter values defined in Theorem 2.3 and Assumption 3 hold true, then the test based on $T_{N,1}$ will have asymptotic size $= \eta$, when $T_{N,1} \geq c_{1\eta}$. Moreover, under the same set of assumptions $P_{H_1}[T_{N,1} \geq c_{1\eta}] \rightarrow 1$, as $N \rightarrow \infty$, i.e. the test based on $T_{N,1}$ will be a consistent test.*

Theorem 2.12 *Let $c_{2\eta}$ be $(1-\eta)$-th quantile of the distribution of $\sum_{i=1}^{4}\lambda_i^* Z_i^2$, where λ_i^*s are the eigen values of $\boldsymbol{\Sigma}_2$, as defined above, and Z_is are independent standard normal random variables. If the assumption required on the parameter values defined in Theorem 2.3 and Assumption 3 hold true, then the test based on $T_{N,2}$ will have asymptotic size $= \eta$, when $T_{N,2} \geq c_{2\eta}$. Moreover, under the same set of assumptions $P_{H_1}[T_{N,2} \geq c_{2\eta}] \rightarrow 1$, as $N \rightarrow \infty$, i.e. the test based on $T_{N,2}$ will be a consistent test.*

Theorem 2.13 *Let* $c_{3\eta}$ *be* $(1-\eta)$*-th quantile of the distribution of* $\left\{\sum_{i=1}^{4}|A_i|\right\}^2$*, where* A_i*s are same as defined above. If the assumption required on the parameter values defined in Theorem 2.3 and Assumption 3 hold true, then the test based on* $T_{N,3}$ *will have asymptotic size* $=\eta$*, when* $T_{N,3} \geq c_{3\eta}$*. Moreover, under the same set of assumptions* $P_{H_1}[T_{N,3} \geq c_{3\eta}] \to 1$*, as* $N \to \infty$*, i.e. the test based on* $T_{N,3}$ *will be a consistent test.*

Theorem 2.14 *Let* $c_{4\eta}$ *be* $(1-\eta)$*-th quantile of the distribution of* $\left\{\sum_{i=1}^{4}|B_i|\right\}^2$*, where* B_i*s are same as defined above. If the assumption required on the parameter values defined in Theorem 2.3 and Assumption 3 hold true, then the test based on* $T_{N,4}$ *will have asymptotic size* $=\eta$*, when* $T_{N,4} \geq c_{4\eta}$*. Moreover, under the same set of assumptions* $P_{H_1}[T_{N,4} \geq c_{4\eta}] \to 1$*, as* $N \to \infty$*, i.e. the test based on* $T_{N,4}$ *will be a consistent test.*

The proofs of the Theorems 2.11 to 2.14 can be found in Dhar, Kundu and Das [11]. Extensive simulations have been performed by the authors to assess the finite sample performances of the four tests. It is observed that all the four tests are able to maintain the level of significance. In terms of the power, it is observed that $T_{N,1}$ and $T_{N,3}$ (based on least squares methodology) perform well when the data are obtained from a light tailed distribution like normal distribution. On the other hand, the tests based on $T_{N,2}$ and $T_{N,4}$ (based on least absolute deviation methodology) perform better than $T_{N,1}$ and $T_{N,3}$ for heavy tailed distribution like Laplace and t distribution with 5 degrees of freedom. Overall, one can prefer the least absolute deviation methodologies when the data are likely to have influential observations/ outliers.

The authors further obtained the asymptotic distributions of the tests based on local alternatives. The following form of the local alternatives has been considered.

$$H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}^0 \qquad vs. \qquad H_{1,N} : \boldsymbol{\theta} = \boldsymbol{\theta}_N = \boldsymbol{\theta}^0 + \boldsymbol{\delta}_N, \tag{44}$$

where $\boldsymbol{\delta}_N = \left(\frac{\delta_1}{N^{1/2}}, \frac{\delta_2}{N^{1/2}}, \frac{\delta_3}{N^{3/2}}, \frac{\delta_4}{N^{5/2}}\right)$. If the assumption required on the parameter values defined in Theorem 2.3, and Assumptions 3 and 4 hold true, then the authors obtained the asymptotic distributions of $T_{N,1}$, $T_{N,2}$, $T_{N,3}$ and $T_{N,4}$ under the alternative hypothesis. The exact forms of the asymptotic distributions are quite involved and they are not presented here. Interested readers are refereed to the original article of Dhar, Kundu and Das [11] for details.

So far we have discussed about the one component chirp model, now we consider the multicomponent chirp model.

2.2 Multicomponent Chirp Model

The multicomponent chirp model can be written as follows:

$$y(n) = \sum_{j=1}^{p}\{A_j^0 \cos(\alpha_j^0 n + \beta_j^0 n^2) + B_j^0 \sin(\alpha_j^0 n + \beta_j^0 n^2)\} + X(n). \tag{45}$$

Here $y(n)$ is the real valued signal as mentioned before, and it is observed at $n = 1, \ldots, N$. For $j = 1, \ldots, p$, A_j^0, B_j^0 are real valued amplitudes, α_j^0, β_j^0 are frequency and frequency rate, respectively. The problem is to estimate the unknown parameters, $A_j^0, B_j^0, \alpha_j^0, \beta_j^0$, for $j = 1, \ldots, p$ and the number of components p, based on the observed sample. Different methods have been proposed in the literature. We will provide different methods and discuss their theoretical properties. Through out this section, it is assumed that p is known in advance. As in case of single chirp model, it is first assumed that the errors $X(n)$s are i.i.d. normal random variables with mean 0 and variance σ^2. More general forms of errors will be considered later on. Before progressing further, write $\boldsymbol{\alpha}^0 = (\alpha_1^0, \ldots, \alpha_p^0)^\top$, $\boldsymbol{\beta}^0 = (\beta_1^0, \ldots, \beta_p^0)^\top$, $\boldsymbol{A} = (A_1^0, \ldots, A_p^0)^\top$, $\boldsymbol{B} = (B_1^0, \ldots, B_p^0)^\top$ and $\boldsymbol{\Gamma}_j^0 = (A_j^0, B_j^0, \alpha_j^0, \beta_j^0)^\top$, for $j = 1, \ldots, p$.

2.3 Maximum Likelihood Estimators

Proceeding as before, the log-likelihood function without the additive constant can be written as

$$l(\mathbf{\Gamma}_1, \ldots, \mathbf{\Gamma}_p, \sigma^2) = -\frac{N}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{n=1}^{N}\left(y(n) - \sum_{j=1}^{p}\{A_j\cos(\alpha_j n + \beta_j n^2) + B_j\sin(\alpha_j n + \beta_j n^2)\}\right)^2. \tag{46}$$

The MLEs of the unknown parameters can be obtained by maximizing (46) with respect to the unknown parameters $\mathbf{\Gamma}_1, \ldots, \mathbf{\Gamma}_p$ and σ^2. As in the case of simple chirp model, the MLEs of $\mathbf{\Gamma}_j = (A_j, B_j, \alpha_j, \beta_j)^\top$, for $j = 1, \ldots, p$, can be obtained by minimizing

$$Q(\mathbf{\Gamma}_1, \ldots, \mathbf{\Gamma}_p) = \sum_{n=1}^{N}\left(y(n) - \sum_{j=1}^{p}\{A_j\cos(\alpha_j n + \beta_j n^2) + B_j\sin(\alpha_j n + \beta_j n^2)\}\right)^2,$$

with respect to the unknown parameters. If the MLEs of $\mathbf{\Gamma}_j$ is denoted by $\widehat{\mathbf{\Gamma}}_j$, for $j = 1, \ldots, p$, the MLE of σ^2 can be obtained as before;

$$\widehat{\sigma}^2 = \frac{1}{N}\sum_{n=1}^{N}\left(y(n) - \sum_{j=1}^{p}\left\{\widehat{A}_j\cos(\widehat{\alpha}_j n + \widehat{\beta}_j n^2) + \widehat{B}_j\sin(\widehat{\alpha}_j n + \widehat{\beta}_j n^2)\right\}\right)^2.$$

It may be observed that $Q(\mathbf{\Gamma}_1, \ldots, \mathbf{\Gamma}_p)$ can be written as follows:

$$Q(\mathbf{\Gamma}_1, \ldots, \mathbf{\Gamma}_p) = \left[\mathbf{Y} - \sum_{j=1}^{p}\mathbf{W}(\alpha_j, \beta_j)\boldsymbol{\theta}_j\right]^\top\left[\mathbf{Y} - \sum_{j=1}^{p}\mathbf{W}(\alpha_j, \beta_j)\boldsymbol{\theta}_j\right], \tag{47}$$

where the $N \times 2$ matrix $\mathbf{W}(\alpha, \beta)$ is same as defined in (12) and $\boldsymbol{\theta}_j = (A_j, B_j)^\top$ is a 2 $\times$ 1 vector for $j = 1, \ldots, p$. The MLEs of the unknown parameters can be obtained by minimizing (47) with respect to the unknown parameters. Now define the $N \times 2p$ matrix $\widetilde{\mathbf{W}}(\boldsymbol{\alpha}, \boldsymbol{\beta})$ as

$$\widetilde{\mathbf{W}}(\boldsymbol{\alpha}, \boldsymbol{\beta}) = [\mathbf{W}(\alpha_1, \beta_1) : \cdots : \mathbf{W}(\alpha_p, \beta_p)],$$

then, for fixed $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, the MLEs of $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_p$, the linear parameter vectors, can be obtained as

$$[\widehat{\boldsymbol{\theta}}_1^\top(\alpha_1, \beta_1) : \cdots : \widehat{\boldsymbol{\theta}}_p^\top(\alpha_p, \beta_p)]^\top = \left[\widetilde{\mathbf{W}}^\top(\boldsymbol{\alpha}, \boldsymbol{\beta})\widetilde{\mathbf{W}}(\boldsymbol{\alpha}, \boldsymbol{\beta})\right]^{-1}\widetilde{\mathbf{W}}^\top(\boldsymbol{\alpha}, \boldsymbol{\beta})\mathbf{Y}.$$

We note that because $\widetilde{\mathbf{W}}^\top(\boldsymbol{\alpha},\boldsymbol{\beta})\widetilde{\mathbf{W}}(\boldsymbol{\alpha},\boldsymbol{\beta})$ is a diagonal matrix for large N, $\widehat{\boldsymbol{\theta}}_j = \widehat{\boldsymbol{\theta}}_j(\alpha_j,\beta_j)$, the MLE of $\boldsymbol{\theta}_j$, $j = 1,\ldots,p$ can also be expressed as

$$\widehat{\boldsymbol{\theta}}_j(\alpha_j,\beta_j) = \left[\mathbf{W}^\top(\alpha_j,\beta_j)\mathbf{W}(\alpha_j,\beta_j)\right]^{-1}\mathbf{W}^\top(\alpha_j,\beta_j)\mathbf{Y}.$$

Using similar techniques as in Section 2.1.1, the MLEs of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ can be obtained as the argument maximum of

$$\mathbf{Y}^\top\widetilde{\mathbf{W}}(\boldsymbol{\alpha},\boldsymbol{\beta})[\widetilde{\mathbf{W}}^\top(\boldsymbol{\alpha},\boldsymbol{\beta})\widetilde{\mathbf{W}}(\boldsymbol{\alpha},\boldsymbol{\beta})]^{-1}\widetilde{\mathbf{W}}^\top(\boldsymbol{\alpha},\boldsymbol{\beta})\mathbf{Y}. \tag{48}$$

The criterion function, given in (48), is a highly non-linear function of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, therefore, the MLEs of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ cannot be obtained in closed form. Saha and Kay [64] suggested to use the method of Pincus [59] to maximize (48). Alternatively, different other methods can also be used to maximize (48).

It may be easily observed that the MLEs obtained above are same as the LSEs. Kundu and Nandi [38] first derived the consistency and the asymptotic normality properties of the LSEs when the error random variables follow Assumption 1. It has been shown that as the sample size $N \to \infty$, the LSEs are strongly consistent. Kundu and Nandi [38] obtained the following consistency result.

THEOREM 2.15 *Suppose there exists a* K*, such that for* $j = 1,\ldots,p$*,* $0 < |A_j^0| + |B_j^0| < K$*,* $0 < \alpha_j^0, \beta_j^0 < \pi$*,* α_j^0 *are distinct, similarly* β_j^0 *are also distinct and* $\sigma^2 > 0$*. If* $X(n)$ *satisfies Assumption 1, then* $\widehat{\boldsymbol{\Gamma}}_j = (\widehat{A}_j, \widehat{B}_j, \widehat{\alpha}_j, \widehat{\beta}_j)^\top$ *is a strongly consistent estimate of* $\boldsymbol{\Gamma}_j^0 = (A_j^0, B_j^0, \alpha_j^0, \beta_j^0)^\top$*, for* $j = 1,\ldots,p$*.* ∎

Along with the consistency results, the asymptotic normality properties of $\widehat{\boldsymbol{\Gamma}}_j$ have been obtained by Kundu and Nandi [38], but the elements of the asymptotic variance covariance matrix look quite complicated. Later Lahiri, Kundu and Mitra [45] using a powerful number theoretic result establish the following result which simplifies the entries of the asymptotic variance covariance matrix.

RESULT 1: If $(\theta_1, \theta_2) \in (0, \pi) \times (0, \pi)$, and θ_2 is an irrational number, then the followings results are true

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \cos(\theta_1 n + \theta_2 n^2) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \sin(\theta_1 n + \theta_2 n^2) = 0$$

$$\lim_{N\to\infty} \frac{1}{N^{t+1}} \sum_{n=1}^{N} n^t \cos^2(\theta_1 n + \theta_2 n^2) = \lim_{N\to\infty} \frac{1}{N^{t+1}} \sum_{n=1}^{N} n^t \sin^2(\theta_1 n + \theta_2 n^2) = \frac{1}{2(t+1)}$$

$$\lim_{N\to\infty} \frac{1}{N^{t+1}} \sum_{n=1}^{N} n^t \sin(\theta_1 n + \theta_2 n^2) \cos(\theta_1 n + \theta_2 n^2) = 0,$$

where $t = 0, 1, 2, \ldots$. ∎

Based on Result 1, Lahiri, Kundu and Mitra [45] establish the following asymptotic distribution of LSEs of $\mathbf{\Gamma}_1^0, \ldots, \mathbf{\Gamma}_p^0$.

THEOREM 2.16 *Under the same assumptions as in Theorem 2.15, for $j = 1, \ldots, p$,*

$$\left(N^{1/2}(\widehat{A}_j - A_j^0), N^{1/2}(\widehat{B}_j - B^0), N^{3/2}(\widehat{\alpha}_j - \alpha_j^0), N^{5/2}(\widehat{\beta}_j - \beta_j^0)\right)^\top \xrightarrow{d} \mathcal{N}_4(\mathbf{0}, 2c\sigma^2\mathbf{\Sigma}_j), \quad (49)$$

where $\mathbf{\Sigma}_j$ can be obtained from the matrix $\mathbf{\Sigma}$ defined in Theorem 2.6, by replacing A^0 and B^0 with A_j^0 and B_j^0, respectively, and c is same as defined in Theorem 2.4. Moreover, $\widehat{\mathbf{\Gamma}}_j$ and $\widehat{\mathbf{\Gamma}}_k$, for $j \neq k$ are asymptotically independently distributed. ∎

2.4 Sequential Estimation Procedures

It is known that the MLEs are the most efficient estimators, but to compute the MLEs one needs to solve a $2p$ dimensional optimization problem. Hence, for moderate or large p it can be a computationally challenging problem. At the same time, due to highly non-linear nature of the likelihood (least squares) surface, any numerical algorithm often converges to a local maximum (minimum) rather than the global maximum (minimum) unless the initial guesses are very close to the true maximum (minimum).

In order to avoid this problem, Lin and Djurić [46], see also Lahiri, Kundu and Mitra [45], proposed a numerically efficient sequential estimation technique which can produce estimators which are asymptotically equivalent to the LSEs, but it involves solving only p two dimensional optimization problems. Therefore, computationally it is quite easy to implement, and even for large p it can be used quite conveniently. The following algorithm may be used to obtain the sequential estimators of the unknown parameters of the multicomponent chirp model (45).

Algorithm 3:

Step 1: First maximize $\mathbf{Z}(\alpha, \beta)$, given in (16), with respect to (α, β) and take the estimates of α and β as argument maximum of $\mathbf{Z}(\alpha, \beta)$. Obtain the estimates of associated A and B by using the separable least squares method of Richards [61]. Denote the estimates of α, β, A, B as $\widehat{\alpha}_1$, $\widehat{\beta}_1$, $\widehat{A}_1$ and $\widehat{B}_1$, respectively.

Step 2: Now to compute the estimates of α_2, β_2, A_2 and B_2, take out the effect of the first component from the signal. Consider the new data vector as

$$\mathbf{Y}^1 = \mathbf{Y} - \mathbf{W}(\widehat{\alpha}_1, \widehat{\beta}_1) \begin{bmatrix} \widehat{A}_1 \\ \widehat{B}_1 \end{bmatrix},$$

where $\mathbf{W}$ matrix is same as defined in (12).

Step 3: Repeat Step 1 by replacing $\mathbf{Y}$ with $\mathbf{Y}^1$ and obtain the estimates of α_2, β_2, A_2, B_2 as $\widehat{\alpha}_2$, $\widehat{\beta}_2$, $\widehat{A}_2$ and $\widehat{B}_2$, respectively.

Step 4: Continue the process p times and obtain the estimates sequentially.

We note that the above algorithm reduces the computational time significantly. A natural question is: how efficient is the sequential estimators? Although, the theoretical properties of the sequential estimators could not be established, extensive simulation results indicate they behave very similarly as the LSEs in terms of MSEs and biases. Lahiri, Kundu and Mitra [45] established that the sequential estimators are strongly consistent and they have the same asymptotic distribution as the LSEs if the following famous number theoretic conjecture, see Montgomery [50], holds.

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

$$\lim_{N \to \infty} \frac{1}{N^{t+1/2}} \sum_{n=1}^{N} n^t \sin(\theta_1' n + \theta_2' n^2) \cos(\theta_1 n + \theta_2 n^2) = 0; \qquad t = 0, 1, 2.$$

In addition if $\theta_2 \neq \theta_2'$, then

$$\lim_{N \to \infty} \frac{1}{N^{t+1/2}} \sum_{n=1}^{N} n^t \cos(\theta_1 n + \theta_2 n^2) \cos(\theta_1' n + \theta_2' n^2) = 0; \qquad t = 0, 1, 2.$$

and

$$\lim_{N \to \infty} \frac{1}{N^{t+1/2}} \sum_{n=1}^{N} n^t \sin(\theta_1 n + \theta_2 n^2) \sin(\theta_1' n + \theta_2' n^2) = 0; \qquad t = 0, 1, 2.$$

In the same paper, it has also been established that if the sequential procedure is continued beyond p steps, then the corresponding amplitude estimates converge to zero almost surely. Therefore, the sequential procedure in a way provides an estimator of the number of components p which is usually unknown in practice.

Recently, Grover, Kundu and Mitra [24] provided another sequential estimators which is based on the ALSEs as proposed in Section 2.1.5. At each step instead of using the LSEs the authors proposed to use the ALSEs. It has been shown by Grover, Kundu and Mitra [24] that the proposed estimators are strongly consistent and have the same asymptotic distribution as the LSEs. In this case also, it has been established that if the sequential procedure is continued beyond p steps, then the corresponding amplitude estimates converge to zero almost surely. It has been further observed based on extensive simulation results that the computational time is significantly smaller if ALSEs are used instead of LSEs studied by

Lahiri, Kundu and Mitra [45] although the performances are very similar in nature. It should be mentioned here that all the methods proposed for one component chirp model can be used sequentially for multicomponent chirp model also.

2.5 Heavy Tailed Error

So far we have discussed about the estimation of the unknown parameters of the chirp model when the errors have second and higher order moments. Recently, Nandi and Kundu [56] considered the estimation of one component and multicomponent chirp models in presence of heavy tailed error. It is assumed that the error random variables $X(n)$s are i.i.d. random variables with mean zero but it may not have finite variance. The following explicit assumptions are made in this case by Nandi and Kundu [56].

Assumption 5 *The error random variables* $\{X(n)\}$ *is a sequence of i.i.d. random variables with mean zero and* $E|X(n)|^{1+\delta} < \infty$, *for some* $0 < \delta < 1$.

Assumption 6 *The error random variables* $\{X(n)\}$ *is a sequence of i.i.d. random variables with mean zero and distributed as a symmetric* α*-stable distribution with the scale parameter* $\gamma > 0$, *where* $1 + \delta < \alpha < 2$, $0 < \delta < 1$. *It implies that the characteristic function of* $X(n)$ *is of the following form:*

$$E[e^{itX(n)}] = e^{-\gamma^\alpha |t|^\alpha}.$$

Nandi and Kundu [56] consider the LSEs and the ALSEs of the unknown parameters of the one component and multi-component chirp models. In case of one component chirp model (7) the following results have been obtained. The details can be found in Nandi and Kundu [56].

Theorem 2.17 *If there exists K, such that $0 < |A^0|^2 + |B^0|^0 < K$, $0 < \alpha^0, \beta^0 < \pi$, $X(n)$ satisfies Assumption 5, then $\widehat{\mathbf{\Gamma}} = (\widehat{A}, \widehat{B}, \widehat{\alpha}, \widehat{\beta})^\top$, the LSE of $\mathbf{\Gamma}^0$, is strongly consistent.* ■

Theorem 2.18 *Under the same assumptions as Theorem 2.17, and further if $X(n)$ satisfies Assumption 6, then*

$$(N^{\frac{\alpha-1}{\alpha}}(\widehat{A} - A^0), N^{\frac{\alpha-1}{\alpha}}(\widehat{B} - B^0), N^{\frac{2\alpha-1}{\alpha}}(\widehat{\alpha} - \alpha^0), N^{\frac{3\alpha-1}{\alpha}}(\widehat{\beta} - \beta^0))^\top$$

converges to a symmetric α-stable distribution of dimension four. ■

It has been shown that the LSEs and ALSEs are asymptotically equivalent. Moreover, in case of multiple chirp model (45), the sequential procedure can be adopted, and their asymptotic properties along the same line as Theorems 13 and 14, have been obtained. It may be mentioned that in presence of heavy tailed error, the LSEs or the ALSEs may not be robust. In that case M-estimators which are more robust, may be considered. Developing both the theoretical properties and numerically efficient algorithm will be of interest. More work is needed in this direction.

2.6 Polynomial Phase Chirp Model

The real valued single component polynomial phase chirp model can be written as follows;

$$y(n) = A^0 \cos(\alpha_1^0 n + \alpha_2^0 n^2 + \ldots + \alpha_k^0 n^k) + B^0 \sin(\alpha_1^0 n + \alpha_2^0 n^2 + \ldots + \alpha_k^0 n^k) + X(n). \quad (50)$$

Here $y(n)$ is the real valued signal observed at $n = 1, \ldots, N$, A^0 and B^0 are amplitudes, $\alpha_m^0 \in (0, \pi)$ for $m = 1, \ldots, k$ are parameters of polynomial phase. The error component $X(n)$ is from a stationary linear process and satisfies Assumption 1. The problem remains the same, that is, estimate the unknown parameters namely A^0, B^0, $\alpha_1^0, \ldots, \alpha_k^0$ given a sample of size N. The corresponding complex model was originally proposed by Djurić and

Kay [12], and Nandi and Kundu [52] established the consistency and asymptotic normality properties of the LSEs of the unknown parameters for the complex polynomial chirp model. Along similar lines the following consistency and asymptotic normality results of the least squares estimators of the unknown parameters of model (50) can be easily established.

THEOREM 2.19 *Suppose there exists a K, such that for $j = 1, \ldots, p$, $0 < |A^0| + |B^0| < K$, $0 < \alpha_1^0, \ldots, \alpha_k^0 < \pi$ and $\sigma^2 > 0$. If $X(n)$ satisfies Assumption 1, then $\widehat{\mathbf{\Gamma}} = (\widehat{A}, \widehat{B}, \widehat{\alpha}_1, \ldots, \widehat{\alpha}_k)^\top$ is a strongly consistent estimate of $\mathbf{\Gamma}^0 = (A^0, B^0, \alpha_1^0, \ldots, \alpha_k^0)^\top$.* ∎

THEOREM 2.20 *Under the same assumptions as in Theorem 2.19,*

$$\left(N^{1/2}(\widehat{A} - A^0), N^{1/2}(\widehat{B} - B^0), N^{3/2}(\widehat{\alpha}_1 - \alpha_1^0), \ldots, N^{(2k+1)/2}(\widehat{\alpha}_k - \alpha_k^0)\right)^\top \xrightarrow{d} \mathcal{N}_{k+2}(\mathbf{0}, 2c\sigma^2\mathbf{\Sigma}),$$

where

$$\mathbf{\Sigma}^{-1} = \begin{bmatrix} 1 & 0 & \frac{B^0}{2} & \frac{B^0}{3} & \cdots & \frac{B^0}{k+1} \\ 0 & 1 & -\frac{A^0}{2} & -\frac{A^0}{3} & \cdots & -\frac{A^0}{k+1} \\ \frac{B^0}{2} & -\frac{A^0}{2} & \frac{A^{0^2}+B^{0^2}}{3} & \frac{A^{0^2}+B^{0^2}}{4} & \cdots & \frac{A^{0^2}+B^{0^2}}{k+2} \\ \frac{B^0}{3} & -\frac{A^0}{3} & \frac{A^{0^2}+B^{0^2}}{4} & \frac{A^{0^2}+B^{0^2}}{5} & \cdots & \frac{A^{0^2}+B^{0^2}}{k+3} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{B^0}{k+1} & -\frac{A^0}{k+1} & \frac{A^{0^2}+B^{0^2}}{k+2} & \frac{A^{0^2}+B^{0^2}}{k+3} & \cdots & \frac{A^{0^2}+B^{0^2}}{2k+1} \end{bmatrix}.$$

Here σ^2 and c are same as defined in Theorem 2.4. ∎

The theoretical properties of the LSEs of the parameters of model (50) can be obtained, but at the same time, the computation of the LSEs is a challenging problem. Not much attention has been paid in this area. Moreover, the asymptotic distribution of the LSEs, as provided in Theorem 2.20, can be used in constructing confidence intervals and also addressing different testing of hypotheses issues. The following problems will be of interest.

Open Problem 9: Develop numerically efficient algorithm to compute the LSEs of the unknown parameters of model (50).

Open Problem 10: Construct different confidence intervals of the unknown parameters of model (50) under different error assumptions and compare their performances in terms of average confidence lengths and coverage percentages.

Open Problem 11: Consider multicomponent polynomial phase signal and develop the theoretical properties of LSEs.

Some of the other methods which have been proposed in the literature are very specific to the complex chirp model. They are based on complex data and its conjugate, see for example Ikram, Abed-Meraim and Hua [30], Besson, Ghogho and Swami [8], Xinghao, Ran and Siyong [74], Zhang et al. [77], Liu and Yu [47], Wang, Su and Chen [72], and see the references cited therein. There are some related models which have been used in place of the chirp model. We will be discussing those in details later on. Now we consider the two dimensional chirp models in next section.

3 Two Dimensional Chirp Model

A multicomponent two-dimensional (2-D) chirp model can be expressed as follows

$$\begin{aligned} y(m,n) &= \sum_{k=1}^{p} \left(A_k^0 \cos(\alpha_k^0 m + \beta_k^0 m^2 + \gamma_k^0 n + \delta_k^0 n^2) + B_k^0 \sin(\alpha_k^0 m + \beta_k^0 m^2 + \gamma_k^0 n + \delta_k^0 n^2) \right) \\ &\quad + X(m,n); \qquad m = 1, \ldots, M, n = 1, \ldots, N. \end{aligned} \tag{51}$$

Here $y(m,n)$ is the observed value, $A_k^0 s$, $B_k^0 s$ are the amplitudes, α_k^0s, γ_k^0s are the frequencies, and β_k^0s, δ_k^0s are the frequency rates. The error random variables $X(m,n)$s have mean zero and they may have some dependence structure. The details will be mentioned later.

The above model (51) is a natural generalization of the 1-D chirp model. Model (51) and some of its variants have been used quite extensively in modeling and analyzing magnetic resonance imaging (MRI), optical imaging and different texture imaging. It has been used quite extensively in modeling black and white 'gray' images, and to analyze finger print images data. This model has a wide applications in modeling Synthetic Aperture Radar (SAR) data and in particular Interferometric SAR data. See, for example, Pelag and Porat [58], Hedley and Rosenfeld [29], Friedlander and Francos [21], Francos and Friedlander [19, 20], Cao, Wang and Wang [9], Zhang and Liu [78], Zhang, Wang and Cao [79], and see the references cited therein.

In this section also first we discuss different estimation procedures and their properties for single component 2-D chirp model and then we consider the multiple 2-D chirp model.

3.1 Single 2-D Chirp Model

The single 2-D chirp model can be written as follows:

$$\begin{aligned} y(m,n) &= A^0\cos(\alpha^0 m+\beta^0 m^2+\gamma^0 n+\delta^0 n^2)+B^0\sin(\alpha^0 m+\beta^0 m^2+\gamma^0 n+\delta^0 n^2) \\ &\quad +X(m,n); \quad m=1,\ldots,M, n=1,\ldots,N. \end{aligned} \tag{52}$$

Here all the quantities are same as defined for model (51). The problem remains the same as 1-D chirp model, i.e. based on the observed data $y(m,n)$, one needs to estimate the unknown parameters under a suitable error assumption. First assume that $X(m,n)$s are i.i.d. Gaussian random variables with mean zero and variance σ^2, for $m=1,\ldots,M$ and $n=1,\ldots,N$. More general error assumptions will be considered in subsequent sections.

3.1.1 Maximum Likelihood and Least Squares Estimators

The log-likelihood function of the observed data without the additive constant can be written as

$$\begin{aligned} l(\boldsymbol{\Theta}) &= -\frac{MN}{2}\ln\sigma^2-\frac{1}{2\sigma^2}\sum_{m=1}^{M}\sum_{n=1}^{N}\big(y(m,n)-A\cos(\alpha m+\beta m^2+\gamma n+\delta n^2)- \\ &\qquad B\sin(\alpha m+\beta m^2+\gamma n+\delta n^2)\big)^2. \end{aligned}$$

Here, $\boldsymbol{\Theta} = (A, B, \alpha, \beta, \gamma, \delta, \sigma^2)^\top$. The MLEs of the unknown parameters $\boldsymbol{\Gamma} = (A, B, \alpha, \beta, \gamma, \delta)^\top$ can be obtained, similarly as 1-D chirp model, as the argument minimum of

$$Q(\boldsymbol{\Gamma}) = \sum_{m=1}^{M} \sum_{n=1}^{N} \left(y(m,n) - A\cos(\alpha m + \beta m^2 + \gamma n + \delta n^2) - B\sin(\alpha m + \beta m^2 + \gamma n + \delta n^2) \right)^2 . \tag{53}$$

If $\widehat{\boldsymbol{\Gamma}} = (\widehat{A}, \widehat{B}, \widehat{\alpha}, \widehat{\beta}, \widehat{\gamma}, \widehat{\delta})^\top$, is the MLE of $\boldsymbol{\Gamma}$, which minimizes (53), then the MLEs of $\widehat{\sigma}^2$ can be obtained as

$$\widehat{\sigma}^2 = \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} \Big(y(m,n) - \widehat{A}\cos(\widehat{\alpha} m + \widehat{\beta} m^2 + \widehat{\gamma} n + \widehat{\delta} n^2) - \widehat{B}\sin(\widehat{\alpha} m + \widehat{\beta} m^2 + \widehat{\gamma} n + \delta n^2) \Big)^2 .$$

As expected, the MLEs cannot be obtained in explicit forms. One needs to use some numerical techniques to compute the MLEs. Newton-Raphson, Gauss-Newton, Genetic algorithm or simulated annealing method may be used for this purpose. Alternatively, the method suggested by Saha and Kay [64] as described in the section 2.1.1 may be used to find the MLEs of the unknown parameters. The details are avoided.

Lahiri [40], see also Lahiri and Kundu [41] in this respect, established the asymptotic properties of the MLEs. Under a fairly general set of conditions, it has been shown that the MLEs are strongly consistent and they are asymptotically normally distributed. The results are provided in details below.

Theorem 3.1 *If there exists a K, such that $0 < |A^0| + |B^0| < K$, $0 < \alpha^0, \beta^0, \gamma^0, \delta^0 < \pi$, and $\sigma^2 > 0$, then $\widehat{\boldsymbol{\Theta}} = (\widehat{A}, \widehat{B}, \widehat{\alpha}, \widehat{\beta}, \widehat{\gamma}, \widehat{\delta}, \widehat{\sigma}^2)^\top$ is a strongly consistent estimate of $\boldsymbol{\Theta}^0 = (A^0, B^0, \alpha^0, \beta^0, \gamma^0, \delta^0, \sigma^2)^\top$.* ∎

Theorem 3.2 *Under the same assumptions as in Theorem 3.1, if we denote $\mathbf{D}$ as a 6×6 diagonal matrix as*

$$\mathbf{D} = diag\left\{ M^{1/2}N^{1/2}, M^{1/2}N^{1/2}, M^{3/2}N^{1/2}, M^{5/2}N^{1/2}, M^{1/2}N^{3/2}, M^{1/2}N^{5/2} \right\},$$

then

$$\mathbf{D}(\widehat{A}-A^0, \widehat{B}-B^0, \widehat{\alpha}-\alpha^0, \widehat{\beta}-\beta^0, \widehat{\gamma}-\gamma^0, \widehat{\delta}-\delta^0)^\top \xrightarrow{d} \mathcal{N}_6(\mathbf{0}, 2\sigma^2\mathbf{\Sigma}),$$

where

$$\mathbf{\Sigma}^{-1} = \begin{bmatrix} 1 & 0 & \frac{B^0}{2} & \frac{B^0}{3} & \frac{B^0}{2} & \frac{B^0}{3} \\ 0 & 1 & -\frac{A^0}{2} & -\frac{A^0}{3} & -\frac{A^0}{2} & -\frac{A^0}{3} \\ \frac{B^0}{2} & -\frac{A^0}{2} & \frac{A^{0^2}+B^{0^2}}{3} & \frac{A^{0^2}+B^{0^2}}{4} & \frac{A^{0^2}+B^{0^2}}{4} & \frac{A^{0^2}+B^{0^2}}{4.5} \\ \frac{B^0}{3} & -\frac{A^0}{3} & \frac{A^{0^2}+B^{0^2}}{4} & \frac{A^{0^2}+B^{0^2}}{5} & \frac{A^{0^2}+B^{0^2}}{4.5} & \frac{A^{0^2}+B^{0^2}}{5} \\ \frac{B^0}{2} & -\frac{A^0}{2} & \frac{A^{0^2}+B^{0^2}}{4} & \frac{A^{0^2}+B^{0^2}}{4.5} & \frac{A^{0^2}+B^{0^2}}{3} & \frac{A^{0^2}+B^{0^2}}{4} \\ \frac{B^0}{3} & -\frac{A^0}{3} & \frac{A^{0^2}+B^{0^2}}{4.5} & \frac{A^{0^2}+B^{0^2}}{5} & \frac{A^{0^2}+B^{0^2}}{4} & \frac{A^{0^2}+B^{0^2}}{5} \end{bmatrix}.$$

■

Note that the asymptotic distribution of the MLEs can be used to construct confidence intervals of the unknown parameters. They can be used for testing of hypothesis problem also.

Since, the LSEs are MLEs under the assumption of i.i.d. Gaussian errors, the above MLEs are the LSEs also. In fact, the LSEs can be obtained by minimizing (53) for a more general set of error assumptions. Let us make the following assumption on the error component $X(m,n)$.

Assumption 7 *The error component $X(m,n)$ has the following form:*

$$X(m,n) = \sum_{j=-\infty}^{\infty}\sum_{k=-\infty}^{\infty} a(j,k)e(m-j,n-k),$$

with

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

where, $\{e(m,n)\}$ is a double array sequence of i.i.d. random variables with mean zero, variance σ^2, and with finite fourth moment.

Note that Assumption 7 is a natural generalization of Assumption 1 from 1-D to 2-D. The additive error $X(m,n)$ which satisfies Assumption 7, is known as the 2-D linear process. Lahiri [40] established that under Assumption 7, the LSE of $\mathbf{\Theta}$ is strongly consistent under the same assumption as in Theorem 3.1. Moreover, the LSEs are asymptotically normally distributed as provided in the following theorem. Denote $\widehat{\mathbf{\Gamma}}$ as the LSE of $\mathbf{\Gamma}^0$.

Theorem 3.3 *Under the same assumptions as in Theorem 3.1 and Assumption 7,*

$$\mathbf{D}(\widehat{\mathbf{\Gamma}} - \mathbf{\Gamma}^0) \xrightarrow{d} N_6(\mathbf{0}, 2c\sigma^2\mathbf{\Sigma}),$$

where matrices $\mathbf{D}$ *and* $\mathbf{\Sigma}$ *are same as defined in Theorem 3.2 and*

$$c = \sum_{j=-\infty}^{\infty}\sum_{k=-\infty}^{\infty} a^2(j,k).$$

∎

3.1.2 Approximate Least Squares Estimators

Along the same line as the 1-D chirp signal model, Grover, Kundu and Mitra [25] considered a 2-D periodogram type function as follows:

$$I(\alpha,\beta,\gamma,\delta) = \frac{2}{MN}\left\{\left(\sum_{m=1}^{M}\sum_{n=1}^{N} y(m,n)\cos(\alpha m + \beta m^2 + \gamma n + \delta n^2)\right)^2 + \left(\sum_{m=1}^{M}\sum_{n=1}^{N} y(m,n)\sin(\alpha m + \beta m^2 + \gamma n + \delta n^2)\right)^2\right\}. \quad (54)$$

The main idea about the above 2-D periodogram type function has been obtained from the periodogram estimator in case the 2-D sinusoidal model. It has been observed by Kundu and Nandi [37] that the 2-D periodogram type estimators for a 2-D sinusoidal model are consistent and asymptotically equivalent to the corresponding LSEs. The 2-D periodogram type estimators or the ALSEs of the 2-D chirp model can be obtained by the argument maximum of $I(\alpha, \beta, \gamma, \delta)$ given in (54) over the range $(0, \pi) \times (0, \pi) \times (0, \pi) \times (0, \pi)$. The explicit solutions of the argument maximum of $I(\alpha, \beta, \gamma, \delta)$ cannot be obtained analytically. Numerical methods are required to compute the ALSEs. Extensive simulation experiments have been performed by Grover, Kundu and Mitra [25], and it has been observed that the Downhill-Simplex method performs quite well to compute the ALSEs, provided the initial guesses are quite close to the true values. It has been observed in simulation studies that although both LSEs and ALSEs involve solving 4-D optimization problem, computational time of the ALSEs is significantly lower than the LSEs. It has been established that the LSEs and ALSEs are asymptotically equivalent. Therefore, the ALSEs are strongly consistent and the asymptotic distribution of the ALSEs is same as the LSEs.

Since the computation of the LSEs or the ALSEs is a challenging problem, several other computationally efficient methods are available in the literature. But unfortunately in most the cases, either the asymptotic properties are unknown or they may not have the same efficiency as the LSEs or ALSEs. Now we provide two estimators which have the same asymptotic efficiency as the LSEs or ALSEs and at the same time both of them can be computed more efficiently than the LSEs or the ALSEs.

3.1.3 2-D Finite Step Efficient Algorithm

Lahiri, Kundu and Mitra [43] extended the one dimensional efficient algorithm as has been described in Section 2.1.4 for 2-D single chirp signal model. The idea is quite similar. It is observed that if we start with the initial guesses of α^0 and γ^0 having convergence rates $O_p(M^{-1}N^{-1/2})$ and $O_p(N^{-1}M^{-1/2})$, respectively, and β^0 and δ^0 having convergence rates $O_p(M^{-2}N^{-1/2})$ and $O_p(N^{-2}M^{-1/2})$, respectively, then after four iterations the algorithm produces estimates of α^0 and γ^0 having convergence rates $O_p(M^{-3/2}N^{-1/2})$ and

$O_p(N^{-3/2}M^{-1/2})$, respectively, and β^0 and δ^0 having convergence rates $O_p(M^{-5/2}N^{-1/2})$ and $O_p(N^{-5/2}M^{-1/2})$, respectively. Therefore, the efficient algorithm produces estimates which have the same rates of convergence as the LSEs or the ALSEs. Moreover, it is guaranteed that the algorithm stops after four iterations.

Before providing the algorithm in details, we introduce the following notation and some preliminary results similarly as in Section 2.1.4. If $\widetilde{\alpha}$ is an estimator of α^0 such that $\widetilde{\alpha}-\alpha^0 = O_p(M^{(-1-\lambda_{11})}N^{-\lambda_{12}})$, for some $0 < \lambda_{11}, \lambda_{12} \leq 1/2$, and $\widetilde{\beta}$ is an estimator of β^0 such that $\widetilde{\beta} - \beta^0 = O_p(M^{(-2-\lambda_{21})}N^{-\lambda_{22}})$, for some $0 < \lambda_{21}, \lambda_{22} \leq 1/2$, then an improved estimator of α^0 can be obtained as

$$\widetilde{\widetilde{\alpha}} = \widetilde{\alpha} + \frac{48}{M^2}\text{Im}\left(\frac{P^{\alpha}_{MN}}{Q^{\alpha,\beta}_{MN}}\right), \tag{55}$$

with

$$P^{\alpha}_{MN} = \sum_{n=1}^{N}\sum_{m=1}^{M} y(m,n)\left(m - \frac{M}{2}\right)e^{-i(\widetilde{\alpha}m+\widetilde{\beta}m^2)} \tag{56}$$

$$Q^{\alpha,\beta}_{MN} = \sum_{n=1}^{N}\sum_{m=1}^{M} y(m,n)e^{-i(\widetilde{\alpha}m+\widetilde{\beta}m^2)}. \tag{57}$$

Similarly, an improved estimator of β^0 can be obtained as

$$\widetilde{\widetilde{\beta}} = \widetilde{\beta} + \frac{45}{M^4}\text{Im}\left(\frac{P^{\beta}_{MN}}{Q^{\alpha,\beta}_{MN}}\right), \tag{58}$$

with

$$P^{\beta}_{MN} = \sum_{n=1}^{N}\sum_{m=1}^{M} y(m,n)\left(m^2 - \frac{M^2}{3}\right)e^{-i(\widetilde{\alpha}m+\widetilde{\beta}m^2)} \tag{59}$$

and $Q^{\alpha,\beta}_{MN}$ is same as defined above in (57).

The following two results provide the justification for the improved estimators, whose proofs can be obtained in Lahiri, Kundu and Mitra [43].

THEOREM 3.4 *If* $\widetilde{\alpha} - \alpha^0 = O_p(M^{-1-\lambda_{11}}N^{-\lambda_{12}})$ *for* $\lambda_{11}, \lambda_{12} > 0$, *then*

(a) $(\widetilde{\widetilde{\alpha}} - \alpha^0) = O_p(M^{-1-2\lambda_{11}}N^{-\lambda_{12}})$ *if* $\lambda_{11} \leq 1/4$,

(b) $M^{3/2}N^{1/2}(\widetilde{\widetilde{\alpha}} - \alpha^0) \xrightarrow{d} N(0, \sigma_1^2)$ *if* $\lambda_{11} > 1/4, \lambda_{12} = 1/2$, *where* $\sigma_1^2 = \dfrac{384c\sigma^2}{A^{02}+B^{02}}$, *the asymptotic variance of the LSE of* α^0, *and c is same as defined in Theorem 3.3.* ∎

THEOREM 3.5 *If* $\widetilde{\beta} - \beta^0 = O_p(M^{-2-\lambda_{21}}N^{-\lambda_{22}})$ *for* $\lambda_{21}, \lambda_{22} > 0$, *then*

$$\begin{array}{lll} (a)\ (\widetilde{\widetilde{\beta}} - \beta^0) = O_p(M^{-2-\lambda_{21}}N^{-\lambda_{22}}) & if & \lambda_{21} \leq 1/4, \\ (b)\ M^{5/2}N^{1/2}(\widetilde{\widetilde{\beta}} - \beta^0) \xrightarrow{d} N(0, \sigma_2^2) & if & \lambda_{21} > 1/4, \lambda_{22} = 1/2, \end{array}$$

where $\sigma_2^2 = \dfrac{360c\sigma^2}{A^{02}+B^{02}}$, *the asymptotic variance of the LSE of* β^0, *and c is same as defined in the previous theorem.* ∎

In order to find estimators of γ^0 and δ^0, interchange the roles of M and N. If $\widetilde{\gamma}$ is an estimator of γ^0 such that $\widetilde{\gamma} - \gamma^0 = O_p(N^{(-1-\kappa_{11})}N^{-\kappa_{12}})$, for some $0 < \kappa_{11}, \kappa_{12} \leq 1/2$, and $\widetilde{\delta}$ is an estimator of δ^0 such that $\widetilde{\delta} - \delta^0 = O_p(M^{(-2-\kappa_{21})}N^{-\kappa_{22}})$, for some $0 < \kappa_{21}, \kappa_{22} \leq 1/2$, then an improved estimator of γ^0 can be obtained as

$$\widetilde{\widetilde{\gamma}} = \widetilde{\gamma} + \frac{48}{N^2}\mathrm{Im}\left(\frac{P_{MN}^{\gamma}}{Q_{MN}^{\gamma,\delta}}\right), \tag{60}$$

with

$$P_{MN}^{\gamma} = \sum_{n=1}^{N}\sum_{m=1}^{M} y(m,n)\left(n - \frac{N}{2}\right)e^{-i(\widetilde{\gamma}n+\widetilde{\delta}n^2)} \tag{61}$$

$$Q_{MN}^{\gamma,\delta} = \sum_{n=1}^{N}\sum_{m=1}^{M} y(m,n)e^{-i(\widetilde{\gamma}n+\widetilde{\delta}n^2)}, \tag{62}$$

and an improved estimator of δ^0 can be obtained as

$$\widetilde{\widetilde{\delta}} = \widetilde{\delta} + \frac{45}{N^4}\mathrm{Im}\left(\frac{P_{MN}^{\delta}}{Q_{MN}^{\gamma,\delta}}\right), \tag{63}$$

with

$$P_{MN}^{\delta} = \sum_{n=1}^{N}\sum_{m=1}^{M} y(m,n)\left(n^2 - \frac{N^2}{3}\right)e^{-i(\widetilde{\gamma}n+\widetilde{\delta}n^2)} \tag{64}$$

and $Q_{MN}^{\gamma,\delta}$ is same as defined in (62).

In this case, the following two results provide the justification for the improved estimators and the proofs can be obtained in Lahiri, Kundu and Mitra [43].

Theorem 3.6 *If* $\widetilde{\gamma} - \gamma^0 = O_p(N^{-1-\kappa_{11}} M^{-\kappa_{12}})$ *for* $\kappa_{11}, \kappa_{12} > 0$, *then*

$$\begin{array}{lll} (a)(\widetilde{\widetilde{\gamma}} - \gamma^0) = O_p(N^{-1-2\kappa_{11}} M^{-\kappa_{12}}) & if & \kappa_{11} \le 1/4, \\ (b) N^{3/2} M^{1/2} (\widetilde{\widetilde{\gamma}} - \gamma^0) \xrightarrow{d} N(0, \sigma_1^2) & if & \kappa_{11} > 1/4, \kappa_{12} = 1/2. \end{array}$$

Here σ_1^2 *and* c *are same as defined in Theorem 3.4.* ∎

Theorem 3.7 *If* $\widetilde{\delta} - \delta^0 = O_p(N^{-2-\kappa_{21}} M^{-\kappa_{22}})$ *for* $\kappa_{21}, \kappa_{22} > 0$, *then*

$$\begin{array}{lll} (a)(\widetilde{\widetilde{\delta}} - \delta^0) = O_p(N^{-2-\kappa_{21}} M^{-\kappa_{22}}) & if & \kappa_{21} \le 1/4, \\ (b) N^{5/2} M^{1/2} (\widetilde{\widetilde{\delta}} - \delta^0) \xrightarrow{d} N(0, \sigma_2^2) & if & \kappa_{21} > 1/4, \kappa_{22} = 1/2. \end{array}$$

Here σ_2^2 *and* c *are same as defined in Theorem 3.5.* ∎

Now we show that starting from the initial guesses $\widetilde{\alpha}$, $\widetilde{\beta}$, with convergence rates $\widetilde{\alpha} - \alpha^0 = O_p(M^{-1}N^{-1/2})$ and $\widetilde{\beta} - \beta^0 = O_p(M^{-2}N^{-1/2})$, respectively, how the above results can be used to obtain efficient estimators, which have the same rate of convergence as the LSEs. It may be noted that finding initial guesses with the above convergence rates are not difficult. It can be obtained by finding the minimum of $Q_1(\alpha, \beta)$, where

$$Q_1(\alpha, \beta) = \sum_{m=1}^{M} \left(\sum_{n=1}^{N} y(m,n) - \mathcal{A}\cos(\alpha m + \beta m^2) - \mathcal{B}\sin(\alpha m + \beta m^2) \right)^2.$$

Here also the main idea is not to use the whole sample at the beginning. We use part of the sample at the beginning and gradually proceed towards the complete sample. The algorithm is described below. We denote the estimates of α^0 and β^0 obtained at the j-th iteration as $\widetilde{\alpha}^{(j)}$ and $\widetilde{\beta}^{(j)}$, respectively.

Algorithm 4:

Step 1: Choose $M_1 = M^{8/9}$, $N_1 = N$. Therefore,

$$\begin{aligned}\widetilde{\alpha}^{(0)} - \alpha^0 &= O_p(M^{-1}N^{-1/2}) = O_p(M_1^{-1-1/8}N_1^{-1/2}) \quad \text{and} \\ \widetilde{\beta}^{(0)} - \beta^0 &= O_p(M^{-2}N^{-1/2}) = O_p(M_1^{-2-1/4}N_1^{-1/2}).\end{aligned}$$

Perform steps (55) and (58). Therefore, after 1-st iteration, we have

$$\begin{aligned}\widetilde{\alpha}^{(1)} - \alpha^0 &= O_p(M_1^{-1-1/4}N_1^{-1/2}) = O_p(M^{-10/9}N^{-1/2}) \quad \text{and} \\ \widetilde{\beta}^{(1)} - \beta^0 &= O_p(M_1^{-2-1/2}N_1^{-1/2}) = O_p(M^{-20/9}N^{-1/2}).\end{aligned}$$

Step 2: Choose $M_2 = M^{80/81}$, $N_1 = N$. Therefore,

$$\begin{aligned}\widetilde{\alpha}^{(1)} - \alpha^0 &= O_p(M_2^{-1-1/8}N_2^{-1/2}) \quad \text{and} \\ \widetilde{\beta}^{(1)} - \beta^0 &= O_p(M_2^{-2-1/4}N_2^{-1/2}).\end{aligned}$$

Perform steps (55) and (58). Therefore, after 2-nd iteration, we have

$$\begin{aligned}\widetilde{\alpha}^{(2)} - \alpha^0 &= O_p(M_2^{-1-1/4}N_2^{-1/2}) = O_p(M^{-100/81}N^{-1/2}) \quad \text{and} \\ \widetilde{\beta}^{(2)} - \beta^0 &= O_p(M_2^{-2-1/2}N_2^{-1/2}) = O_p(M^{-200/81}N^{-1/2}).\end{aligned}$$

Step 3: Choose $M_3 = M$, $N_3 = N$. Therefore,

$$\begin{aligned}\widetilde{\alpha}^{(2)} - \alpha^0 &= O_p(M_3^{-1-19/81}N_3^{-1/2}) \quad \text{and} \\ \widetilde{\beta}^{(2)} - \beta^0 &= O_p(M_2^{-2-38/81}N_3^{-1/2}).\end{aligned}$$

Again, performing steps (55) and (58), after 3-rd iteration, we have

$$\begin{aligned}\widetilde{\alpha}^{(3)} - \alpha^0 &= O_p(M^{-1-38/81}N^{-1/2}) \quad \text{and} \\ \widetilde{\beta}^{(3)} - \beta^0 &= O_p(M^{-2-76/81}N^{-1/2}).\end{aligned}$$

Step 4: Choose $M_4 = M$, $N_4 = N$, and after performing steps (55) and (58) we obtain the required convergence rates, i.e.

$$\begin{aligned}\widetilde{\alpha}^{(4)} - \alpha^0 &= O_p(M^{-3/2}N^{-1/2}) \quad \text{and} \\ \widetilde{\beta}^{(4)} - \beta^0 &= O_p(M^{-5/2}N^{-1/2}).\end{aligned}$$

Similarly, interchanging the role of M and N, we can get the algorithm corresponding to γ^0 and δ^0. Extensive simulation experiments have been carried out by Lahiri [40], and it is observed that the performance of the finite step algorithm is quite good in terms of MSEs and biases. In initial steps, the part of the sample is selected in such a way that the dependence structure is maintained in the subsample. The MSEs and biases of the finite step algorithm are very similar with the corresponding performance of the LSEs. Therefore, the finite step algorithm can be used quite efficiently in practice. Now we will introduce more general 2-D polynomial phase model and discuss its applications, estimation procedures and their properties.

4 Two-Dimensional Polynomial Phase Signal Model

In previous sections, we have discussed about 1-D and 2-D chirp models in details. But 2-D polynomial phase signal model also has received significant amount of attention in the signal processing literature. Francos and Friedlander [19] first introduced the most general 2-D polynomial (of degree r) phase signal model, and it can be described as follows:

$$\begin{aligned}y(m,n) &= A^0 \cos\left(\sum_{p=1}^{r}\sum_{j=0}^{p} \alpha^0(j,p-j) m^j n^{p-j}\right) + B^0 \sin\left(\sum_{p=1}^{r}\sum_{j=0}^{p} \alpha^0(j,p-j) m^j n^{p-j}\right) \\ &\quad + X(m,n); \quad m = 1,\ldots,M; \quad n = 1,\ldots,N, \qquad (65)\end{aligned}$$

here $X(m,n)$ is the additive error with mean 0, A^0 and B^0 are non zero amplitudes, and for $j = 0,\ldots,p, = 1,\ldots,r$, $\alpha^0(j,p-j)$'s are distinct frequency rates of order $(j,p-j)$, respectively, and lie strictly between 0 and π. Here $\alpha^0(0,1)$ and $\alpha^0(1,0)$ are called frequencies. The explicit assumptions on the error $X(m,n)$ will be provided later.

Different specific forms of model (65) have been used quite extensively in the literature. Friedlander and Francos [21] used 2-D polynomial phase signal model to analyze finger print type data and Djurović et al. [15] used a specific 2-D cubic phase signal model due to its applications in modelling Synthetic Aperture Radar (SAR) data and in particular Interferometric SAR data. Further, 2-D polynomial phase signal model also has been used in modeling and analyzing magnetic resonance imaging (MRI), optical imaging and different texture imaging. For some of the other specific applications of this model, one may refer to Djukanović and Djurović [10], Ikram and Zhou [31], Wang and Zhou [70], Tichavsky and Handel [66], Amar, Leshem and van der Veen [2] and see the references cited therein.

Different estimation procedures have been suggested in the literature based on the assumption that the error components are i.i.d. random variables with mean zero and finite variance. For example, Farquharson, O'Shea and Ledwich [17] provided a computationally efficient estimation procedures of the unknown parameters of a polynomial phase signals. Djurović and Stanković [14] proposed the quasi maximum likelihood estimators based on the normality assumption of the error random variables. Recently, Djurović, Simenunović and Wang [13] considered an efficient estimation method of the polynomial phase signal parameters by using a cubic phase function. Some of the refinement of the parameter estimation of the polynomial phase signals can be obtained in O'Shea [57].

Interestingly, a significant amount of work has been done in developing different estimation procedures to compute parameter estimates of the polynomial phase signal, but not much work has been done in developing the properties of the estimators. In most of the cases, the mean squared errors or the variances of the estimates are compared with the corresponding Cramer-Rao lower bound, without establishing formally that the asymptotic variances of the maximum likelihood estimators attain the lower bound in case the errors are i.i.d. normal random variables. Recently, Lahiri and Kundu [41] established formally the asymptotic properties of the LSEs of parameters in model (65) under a fairly general assumptions on the error random variables. From the results of Lahiri and Kundu [41] it can be easily obtained that when the errors are i.i.d. normally distributed the asymptotic variance of the MLEs attain the Cramer-Rao lower bound.

Lahiri and Kundu [41] established the consistency and the asymptotic normality properties of the LSEs of the parameters of model (65) under a general error assumptions. It is assumed that the errors $X(m,n)$s follow Assumption 7 as defined in Section 3.1.1, and the parameter vector satisfies Assumption 8, given below. For establishing the consistency and asymptotic normality of the LSEs, the following Assumption is required for further development.

Assumption 8 *Denote the true parameters by* $\boldsymbol{\theta}^0 = (A^0, B^0, \alpha^0(j, p-j), j = 0, \ldots, p, p = 1, \ldots, r)^\top$ *and the parameter space by* $\boldsymbol{\Theta} = [-K, K] \times [-K, K] \times [0, \pi]^{r(r+3)/2}$*. Here* $K > 0$ *is an arbitrary constant and* $[0, \pi]^{r(r+3)/2}$ *denotes* $r(r+3)/2$ *fold of* $[0, \pi]$*. It is assumed that* $\boldsymbol{\theta}^0$ *is an interior point of* $\boldsymbol{\Theta}$*.*

Lahiri and Kundu [41] established the following results under the assumption of 2-D stationary error process. The details can be obtained in that paper.

THEOREM 4.1 *If Assumptions 7 and 8 are satisfied, then* $\widehat{\boldsymbol{\theta}}$*, the LSE of* $\boldsymbol{\theta}^0$*, is a strongly consistent estimator of* $\boldsymbol{\theta}^0$*.* ∎

THEOREM 4.2 *Under Assumptions 7 and 8,* $\mathbf{D}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0) \rightarrow \mathcal{N}_d(0, 2c\sigma^2\boldsymbol{\Sigma}^{-1})$ *where the matrix* $\mathbf{D}$ *is a* $(2 + \frac{r(r+3)}{2}) \times (2 + \frac{r(r+3)}{2})$ *diagonal matrix of the form*

$$\mathbf{D} = diag\left(M^{\frac{1}{2}}N^{\frac{1}{2}}, M^{\frac{1}{2}}N^{\frac{1}{2}}, M^{j+\frac{1}{2}}N^{(p-j)+\frac{1}{2}},\ j = 0, \cdots, p,\ p = 1, \cdots, r\right),$$

and

$$\mathbf{\Sigma} = \begin{bmatrix} 1 & 0 & \mathbf{V}_1 \\ 0 & 1 & \mathbf{V}_2 \\ \mathbf{V}_1^T & \mathbf{V}_2^T & \mathbf{W} \end{bmatrix} \tag{66}$$

Here $\mathbf{V}_1 = (\frac{B^0}{(j+1)(p-j+1)},\ j = 0, \cdots, p,\ p = 1, \cdots, r)$, $\mathbf{V}_2 = (-\frac{A^0}{(j+1)(p-j+1)},\ j = 0, \cdots, p,\ p = 1, \cdots, r)$, *are vectors of order* $1 \times \frac{r(r+3)}{2}$, $W = (\frac{{A^0}^2 + {B^0}^2}{(j+k+1)(p+q-j-k+1)}, j = 0, \cdots, p,\ p = 1, \cdots, r,\ k = 0, \cdots, q,\ q = 1, \cdots, r)$, *is a matrix of order* $\frac{r(r+3)}{2} \times \frac{r(r+3)}{2}$, *and* $c = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} a(j,k)^2$. *Further,* $\mathcal{N}_d(0, 2c\sigma^2\mathbf{\Sigma}^{-1})$ *denotes a d-variate normal distribution with the mean vector* 0, *and dispersion matrix* $2c\sigma^2\mathbf{\Sigma}^{-1}$ *with* $d = 2 + \frac{r(r+3)}{2}$. ∎

Theoretical properties of the LSEs have been established by Lahiri and Kundu [41], but finding the LSEs is a computationally challenging problem. No work has been done along this line. It is a very important practical problems. Further work is needed along that direction.

5 Some Other Related Models

In this section, we briefly provide some of the other related models which have been recently introduced in the literature and it has received some attention because of its importance in practice. We will be mainly discussing three related models namely; (a) random amplitude chirp model introduced by Besson, Giannakis and Gini [7], (b) harmonic chirp model of Doweck, Amar and Cohen [16] and (c) chirp like model proposed by Grover, Kundu and Mitra [26].

5.1 Random Amplitude Chirp Model

Besson, Giannakis and Gini [7] first proposed the random amplitude chirp model. It is easier to use the complex domain in this case, hence we provide the complex valued random amplitude chirp model as follows:

$$y(n) = Z(n)e^{i(\phi^0+\alpha^0 n+\beta^0 n^2)} + X(n); \quad n = 1, \ldots, N. \tag{67}$$

Here it is assumed that α^0 and β^0 are same as defined before, and ϕ^0 is the phase. Further, $Z(n)$ is a real valued stationary mixing process with non-zero mean and its covariance matrix is unknown. The complex valued error random variables $X(n)$ is assumed to be a complex circular Gaussian process with mean zero and variance σ^2.

The main problem here is to estimate the unknown parameters α^0 and β^0. Besson, Giannakis and Gini [7] proposed the following estimators of α^0 and β^0;

$$(\widehat{\alpha}, \widehat{\beta}) = \text{arg max}_{\alpha,\beta} \frac{1}{N} \left| \sum_{n=1}^{N} y^2(n) e^{-i2(\alpha n+\beta n^2)} \right|. \tag{68}$$

Besson, Giannakis and Gini [7] provided some heuristic justification about this estimator. Although, they could not provide any theoretical properties of these estimators, it has been observed by extensive simulation experiments that the performances of the estimators (68) are quite satisfactory. Recently, Nandi and Kundu [54] provided the consistency and the asymptotic normality properties of the estimators given in (68) under a quite general condition on $X(n)$ and $Z(n)$. In the same paper, Nandi and Kundu [54] proposed a multicomponent random amplitude chirp model and established the consistency and asymptotic normality properties of the estimators under a similar set of conditions on the error components and on the random amplitudes.

5.2 Harmonic Chirp Model

The harmonic chirp model is a generalization of the harmonic sinusoidal model, where the instantaneous frequency of the harmonics changes linearly as a function of time. The harmonic sinusoidal model, also known as the fundamental frequency model, has received considerable attention in the statistical signal processing literature. The harmonic sinusoidal model can be expressed as follows:

$$y(n) = \sum_{j=1}^{p} \left[A_j^0 \cos(nj\alpha^0) + B_j^0 \sin(nj\alpha^0) \right] + X(n); \qquad n = 1, \ldots, N. \tag{69}$$

Here the error random variables $X(n)$, the amplitudes A_j^0 and B_j^0 are same as defined before. The fundamental frequency $\alpha^0 \in (0, \pi/p)$. The problem here also is to estimate the unknown parameters namely p, α^0 and $(A_1^0, B_1^0), \ldots, (A_p^0, B_p^0)$. Quinn and Thomson [60] first considered this model and proposed an efficient estimation technique of the unknown parameters and obtain the consistency and asymptotic normality properties of the proposed estimators. Some of the results have been extended by Nandi and Kundu [51, 53] and Irizarry [32].

Doweck, Amar and Cohen [16] first considered the harmonic chirp model and it can be written as follows:

$$y(n) = \sum_{j=1}^{p} \left[A_j^0 \cos(j(\alpha^0 n + \beta^0 n^2) + B_j^0 \sin(j(\alpha^0 n + \beta^0 n^2)) \right] + X(n); \qquad n = 1, \ldots, N. \tag{70}$$

In this case the extra parameter $\beta \in (0, \pi/p)$ is known as the fundamental frequency rate. The other parameters are same as defined in (69). The problem remains the same, i.e., based on the observations $y(1), \ldots, y(N)$, and under a suitable error assumptions of $X(n)$, estimate the unknown parameters; p, $A_1^0, \ldots, A_p^0$, $B_1^0, \ldots, B_p^0$, α^0 and β^0. Doweck, Amar and Cohen [16] obtained the MLEs of the unknown parameters namely A_j^0's, B_j^0's α^0 and β^0, for fixed p and when $X(n)$'s are i.i.d. Gaussian random variables with mean zero and finite variance. Note that the MLEs can be obtained by solving a two dimensional optimization problem. Finally they proposed to estimate using some information theoretic criteria like Akaike Information Criterion (AIC) or Bayesian Information Criterion (BIC). Jensen et al. [34] provided a fast algorithm to compute the MLEs of the unknown parameters of model (70) under the same set of error assumptions. Although it has been observed that the performances of the maximum likelihood estimators are quite good, the properties of the estimators have not yet been established. It will be interesting to develop estimation procedures under more general error assumptions and also develop their theoretical properties.

5.3 Chirp Like Model

Recently, Grover, Kundu and Mitra [26] proposed a chirp like model for $n = 1, \ldots, N$, as follows:

$$y(n) = \sum_{j=1}^{p} \left[A_j^0 \cos(\alpha_j^0 n) + B_j^0 \sin(\alpha_j^0 n) \right] + \sum_{j=1}^{q} \left[C_j^0 \cos(\beta_j^0 n^2) + D_j^0 \sin(\beta_j^0 n^2) \right] + X(n). \quad (71)$$

Here the linear parameters A_j^0's, B_j^0's, C_j^0's and D_j^0's can take any real values and $0 < \alpha_j^0, \beta_j^0 < \pi$ as before. The additive error has mean zero and it satisfies Assumption 1.

The main motivation about this proposed model (71) is that it is as flexible as the multicomponent chirp model (45). It has been observed by Grover, Kundu and Mitra [26] based on extensive computer simulations that any multicomponent chirp model (45) can be well approximated by the chirp like model (71) with a proper choice of p and q. But the estimation of the parameters for model (71) becomes computationally less challenging than the multicomponent chirp model (45). It has been observed that the sequential least squares method can be used to compute the estimators of the parameters of model (71) and at each step one needs to solve only one-dimensional optimization problem. The properties of the sequential least squares estimators have been obtained by Grover, Kundu and Mitra [26]. The sequential estimators are strongly consistent estimators of the corresponding parameters and they are asymptotically normally distributed. The rates of convergence of the sequential least squares estimators of the linear, frequency and frequency rate parameters of model (45) are same as the corresponding sequential least squares estimators of the multicomponent chirp model (45). Extensive data analyses have been performed and it has been observed that the proposed chirp like model (71) can be used quite effectively in place of multicomponent chirp model (45). More work is needed in developing different efficient estimation procedures and developing their properties.

6 Conclusions

In this paper, we have provided a brief review of 1-D and 2-D chirp and some related models which have received a considerable amount of attention in recent years in the signal processing literature. These models have been used in analyzing different real life signals or images quite efficiently. It is observed that several sophisticated statistical and computational techniques are needed to analyze these models and developing estimation procedures. We have provided several open problems for future research. Although, these models are in use in the statistical signal processing literature not much attention has been paid in the main stream statistics literature. It is our belief that statisticians can make a significant contributions in these directions. We hope this article will generate some awareness among the statisticians about the challenges and they will be interested in providing effective solutions.

Acknowledgements:

The authors would like to thank the unknown reviewers for their helpful comments which have helped to improve the manuscript significantly.

References

[1] Abatzoglou, T. (1986), "Fast maximum likelihood joint estimation of frequency and frequency rate", *IEEE Transactions on Aerospace and Electronic Systems*, vol. 22, 708 – 715.

[2] Amar, A., Leshem, A. and van der Veen, A. J. (2010), "A low complexity blind estimator of narrowband polynomial phase signals", *IEEE Transactions on Signal Processing*, vol. 58, 4674 – 4683.

[3] Bai, Z.D., Rao, C.R., Chow, M. and Kundu, D. (2003), "An efficient algorithm for estimating the parameters of superimposed exponential signals", *Journal of Statistical Planning and Inference*, vol. 110, 23 – 34.

[4] Bai, Z.D., Rao, C.R., Wu, Y., Zen, M-M., Zhao, L. (1999), "The simultaneous estimation of the number of signals and frequencies of multiple sinusoids when some observations are missing: I. Asymptotic", *Proceedings of National Academy of Sciences*, vol. 96, 11106 V 11110.

[5] Barbarossa, S. and Petrone, V. (1997). Analysis of polynomial-phase signals by the integrated generalized ambiguity function. *IEEE Transactions on Signal Processing*, **45**, 316-327.

[6] Barbarossa, S., Scaglione, A. and Giannakis, G. (1998). Product high-order ambiguity function for multicomponent polynomial-phase signal modeling. *IEEE Transactions on Signal Processing*, **46**, 691-708.

[7] Benson, O., Ghogho, M. and Swami, A. (1999), "Parameter estimation for random amplitude chirp signals", *IEEE Transactions on Signal Processing*, vol. 47, 3208 – 3219.

[8] Besson, O., Giannakis, G.B. and Gini, F. (1999), "Improved estimation of hyperbolic frequency modulated chirp signals", *IEEE Transactions on Signal Processing*, vol. 47, 1384 – 1388.

[9] Cao, F.,Wang, S., Wang, F. (2006). Cross-Spectral Method Based on 2-D Cross Polynomial Transform for 2-D Chirp Signal Parameter Estimation. *ICSP2006 Proceedings*, DOI:10.1109/ICOSP.2006.344475.

[10] Djukanović, S. and Djurović, I. (2012) “Aliasing detection and resolving in the estimation of polynomial-phase signal parameters”, *Signal Processing*, vol. 92, 235 – 239.

[11] Dhar, S.S., Kundu, D. and Das, U. (2019), “On testing parameters of chirp signal model”, *IEEE Transactions on Signal Processing*, vol. 67, 4291 – 4301.

[12] Djurić, P.M. and Kay, S.M. (1990), “Parameter estimation of chirp signals”, *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 38, 2118 – 2126.

[13] Djurović, I., Simeunović, M. and Wang, P. (2017), “Cubic phase function: A simple solution for polynomial phase signal analysis”, *Signal Processing*, vol. 135, 48 – 66.

[14] Djurović, I. and Stanković L. J. (2014), “Quasi maximum likelihood estimators of polynomial phase signals”, *IET Signal Processing*, vol. 13, 347 – 359.

[15] Djurović, I., Wang, P. and Ioana, C. (2010) “Parameter estimation of 2-D cubic phase signal function using genetic algorithm”, *Signal Processing*, vol. 90, 2698 – 2707.

[16] 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*, vol. 63, 1765 – 1778.

[17] Farquharson, M., O’Shea, P. and Ledwich, G. (2005), “A computationally efficient technique for estimating the parameters phase signals from noisy observations”, *IEEE Transactions on Signal Processing*, vol. 53, 3337 – 3342.

[18] Fourier, D., Auger, F., Czarnecki, K. and Meignen, S. (2017), "Chirp rate and instantaneous frequency estimation: application to recursive vertical synchrosqueezing", *IEEE Signal Processing Letters*, 1-1, PP, 99.

[19] Francos, J.M. and Friedlander, B. (1998), "Two-dimensional polynomial phase signals: parameter estimation and bounds", *Multidimensional Systems and Signal Processing*, vol. 9, 173 – 205.

[20] Francos, J.M. and Friedlander, B. (1999), "Parameter estimation of 2-D random amplitude polynomial phase signals", *IEEE Transactions on Signal Processing*, vol. 47, 1795 – 1810.

[21] Friedlander, B. and Francos, J.M. (1996), "An estimation algorithm for 2-D polynomial phase signals", *IEEE Transactions on Image Processing*, vol. 5, 1084 – 1087.

[22] Gabor, D. (1946), "Theory of communication. Part 1: The analysis of information" *Journal of the Institution of Electrical Engineers - Part III: Radio and Communication Engineering*, vol. 93, 429 – 441.

[23] Gini, F., Montanari, M. and Verrazzani, L. (2000), "Estimation of chirp signals in compound Gaussian clutter: a cyclostationary approach", *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 48, 1029 – 1039.

[24] Grover, R., Kundu, D. and Mitra, A. (2018a), "On approximate least squares estimators of parameters of one-dimensional chirp signal", *Statistics*, vol. 52, 1060 – 1085.

[25] Grover, R, Kundu, D. and Mitra, A. (2018b), "Asymptotic of approximate least squares estimators of parameters of two-dimensional chirp signal", *Journal of Multivariate Analysis*, vol. 168, 211 – 220.

[26] Grover, R, Kundu, D. and Mitra, A. (2018c), “A chirp like model to analyze periodic and nearly periodic signals”, *arXiv*.

[27] Grover, R., Kundu, D. and Mitra, A. (2020), “An efficient methodology to estimate the parameters of a two-dimensional chirp signal model”, *Multidimensional Systems and Signal Processing*, (to appear).

[28] Guo, J., Zou, H., Yang, X. and Liu, G. (2011), “Parameter estimation of multicomponent chirp signals via sparse representation”, *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, 2261 – 2268.

[29] Hedley, M. and Rosenfeld, D. (1992). A new two-dimensional phase unwrapping algorithm for MRI images. *Magnetic Resonance in Medicine*, **24**, 177-181.

[30] Ikram, M.Z., Abed-Meraim, K., Hua, Y. (1998), “Estimating the parameters of chirp signals: an iterative approach”, *IEEE Transactions on Signal Processing*, vol. 46, 3436 –3441.

[31] Ikram, M.Z. and Zhou, G.T. (2001), “Estimation of multicomponent phase signals of mixed orders”, *Signal Processing*, vol. 81, 2293 – 2308.

[32] Irizarry, R. A. (2000), “Asymptotic distribution of estimates for a time-varying parameter in a harmonic model with multiple fundamentals”, *Statistica Sinica*, vol. 10, 1041 – 1067.

[33] Jennrich, R.I. (1969), “Asymptotic properties of the nonlinear least squares estimators”, *Annals of Mathematical Statistics*, vol. 40, 633 – 643.

[34] 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*, vol. 65, 5137 – 5152.

[35] Kennedy, W.J. Jr. and Gentle, J.E. (1980), *Statistical Computing*, Marcel Dekker, Inc., New York.

[36] Kundu, D. (2020), "Professor C.R. Rao's contribution in Statistical Signal Processing and its Longterm Implications", *Proceedings of Indian Academy of Sciences (Mathematical Sciences)*, vol. 130, Article No. 43.

[37] Kundu, D. and Nandi, S. (2003), "Determination of discrete spectrum in a random field", *Statistica Neerlandica*, vol. 57, 258 – 283.

[38] Kundu, D. and Nandi, S. (2008), "Parameter estimation of chirp signals in presence of stationary Christensen, noise", *Statistica Sinica*, vol. 18, 187 – 201.

[39] Kundu, D. and Nandi, S. (2012), *Statistical Signal Processing: Frequency Estimation*, Springer, New Delhi.

[40] Lahiri, A. (2011), *Estimators of parameters of chirp signals and their properties*, Ph.D. thesis, Department of Mathematics and Statistics, Indian Institute of Technology, Kanpur, India.

[41] Lahiri, A. and Kundu, D. (2017), "On parameter estimation of two-dimensional polynomial phase signal model", *Statistica Sinica*, vol. 27, 1779 – 1792.

[42] Lahiri, A., Kundu, D. and Mitra, A. (2012), "Efficient algorithm for estimating the parameters of chirp signal", *Journal of Multivariate Analysis*, vol. 108, 15 – 27.

[43] Lahiri, A., Kundu, D. and Mitra, A. (2013), "Efficient algorithm for estimating the parameters of two dimensional chirp signal", *Sankhya, Ser. B*, vol. 75, 65 – 89.

[44] Lahiri, A., Kundu, D. and Mitra, A. (2014), "On least absolute deviation estimator of one dimensional chirp model", *Statistics*, vol. 48, 405 – 420.

[45] Lahiri, A., Kundu, D. and Mitra, A. (2015), "Estimating the parameters of multiple chirp signals", *Journal of Multivariate Analysis*, vol. 139, 189 – 205.

[46] Lin, C-C. and Djurić, P.M. (2000), "Estimation of chirp signals by MCMC", *ICASSP-1998*, vol. 1, 265 – 268.

[47] Liu, X. and Yu, H. (2013), "Time-domain joint parameter estimation of chirp signal based on SVR", *Mathematical Problems in Engineering*, Article ID 952743, 1 – 9.

[48] Lu, Y., Demirli, R., Cardoso, G. and Saniie, J. (2006), "A successive parameter estimation algorithm for chirplet signal decomposition", *IEEE Transactions on Ultrasonic, Ferroelectrics and Frequency Control*, vol. 53,

[49] Mazumder, S. (2017), "Single-step and multiple-step forecasting in one-dimensional chirp signal using MCMC-based Bayesian analysis", *Communications in Statistics - Simulation and Computation*, vol. 46, 2529 – 2547.

[50] Montgomery, H.L. (1990), *Ten Lectures on the Interface Between Analytic Number Theory and Harmonic Analysis*, American Mathematical Society, 196.

[51] Nandi, S. and Kundu, D. (2003), "Estimating the fundamental frequency of a periodic function", *Statistical Methods and Applications*, vol. 12, 341 – 360.

[52] 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*, vol. 56, 529 – 544.

[53] Nandi, S. and Kundu, D. (2006), "Analyzing non-stationary signals using a cluster type model", *Journal of Statistical Planning and Inference*, vol. 136, 3871 – 3903.

[54] Nandi, S. and Kundu, D. (2020), "Random amplitude chirp model", *Signal Processing*, vol. 168, Art. 107328.

[55] Nandi, S. and Kundu, D. (2020), *Statistical Signal Processing*, 2nd edition, Springer, London.

[56] Nandi, S., Kundu, D. and Grover, R. (2019), "Estimation of Parameters of Multiple Chirp Signal in presence of Heavy Tailed Errors", submitted for publication.

[57] O'Shea, P. (2010), "On refining polynomial phase signal parameter estimates", *IEEE Transactions on Aerospace, Electronic Systems*, vol. 4, 978 – 987.

[58] Pelag, S. and Porat, B. (1991), "Estimation and classification of polynomial phase signals", *IEEE Transactions on Information Theory*, vol. 37, 422 – 430.

[59] Pincus, M. (1968), "A closed form solution of certain programming problems", *Operation Research*, vol. 16, 690 – 694.

[60] Quinn, B. G. and Thomson, P. J. (1991), "Estimating the frequency of a periodic function", *Biometrika*, vol. 78, 65 – 74.

[61] Richards, F.S.G. (1961), "A method of maximum likelihood estimation", *Journal of the Royal Statistical Society, Ser. B*, vol. 23, 469 - 475.

[62] Rihaczek, A.W. (1969), *Principles of high resolution radar*, McGraw-Hill, New York.

[63] Robertson, S.D., Gray, H.L. and Woodward, W.A. (2010), "The generalized linear chirp process", *Journal of Statistical Planning and Inference*, vol. 140, 3676 – 3687.

[64] 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, 224 – 230.

[65] Seber, G. A. F. and Wild, C.J. (1989), *Nonlinear Regression*, John Wiley and Sons, New York.

[66] Ticahvsky, P. and Handel, P. (1999), "Multicomponent polynomial phase signal analysis using a tracking algorithm", *IEEE Transactions on Signal Processing*, vol. 47, 1390 – 1395.

[67] Vinogradov, I. M. (1954), "The method of trigonometrical sums in the theory of numbers", *Interscience*, Translated from Russian, Revised and annoted by K.F. Roth and A. Devenport. Reprint of the 1954 translation. Dover publications, Inc., Mineola, New York, U.S.A., 2004.

[68] Volcker, B. and Ottersten, B. (2001), "Chirp parameter estimation from a sample covariance matrix", *IEEE Transactions on Signal Processing*, vol. 49, 603 – 612.

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

[70] Wang, Y. and Zhou, Y.G.T. (1998), "On the use of high-order ambiguity function for multi-component polynomial phase signals", *Signal Processing*, vol. 5, 283 – 296.

[71] Wu, C.F.J. (1981), "Asymptotic theory of the nonlinear least squares estimation", *Annals of Statistics*, vol. 9, 501-513.

[72] Wang, J. Z., Su, S. Y., Chen, Z. . (2015), "Parameter estimation of chirp signal under low SNR", *Science China: Information Sciences*, vol. 58, 020307:1 – 020307:13.

[73] Wu, Y., So, H.C. and Liu, H. (2008). Subspace-Based Algorithm for Parameter Estimation of Polynomial Phase Signals. *IEEE Transactions on Signal Processing*, **56**, 4977 - 4983.

[74] Xinghao, Z., Ran, T., Siyong, Z. (2003), "A novel sequential estimation algorithm for chirp signal parameters", *IEEE Conference in Neural Networks and & Signal Processing*, Nanjing, China, December 14-17, 2003, pp 628 – 631.

[75] Yang, P., Liu, Z. and Jiang, W-L. (2015), "parameter estimation of multicomponent chirp signals based on discrete chirp Fourier transform and population Monte Carlo", *Signal, Image and Video Processing*, vol. 9, 1137 – 1149.

[76] Yaron, D., Alon, A. and Israel, C. (2015), "Joint model order selection and parameter estimation of chirps with harmonic components", *IEEE Transactions on Signal Processing*, vol. 63, 1765 – 1778.

[77] Zhang, H., Liu, H., Shen, S., Zhang, Y., Wang, X. (2013), "Parameter estimation of chirp signals based on fractional Fourier transform", *The Journal of China Universities of Posts and Telecommunications*, vol. 20, (Suppl. 2), 95 – 100.

[78] Zhang, H. and Liu, Q. (2006). Estimation of instantaneous frequency rate for multicomponent polynomial phase signals. *ICSP2006 Proceedings*, 498- 502, DOI: 10.1109/ICOSP.2006.344448.

[79] Zhang, K., Wang, S., Cao, F. (2008). Product Cubic Phase Function Algorithm for Estimating the Instantaneous Frequency Rate of Multicomponent Two-dimensional Chirp Signals. *2008 Congress on Image and Signal Processing*, DOI: 10.1109/CISP.2008.352.