



# An efficient methodology to estimate the parameters of a two-dimensional chirp signal model

Rhythm Grover<sup>1</sup> · Debasis Kundu<sup>1</sup> · Amit Mitra<sup>1</sup>

Received: 7 December 2019 / Revised: 12 April 2020 / Accepted: 24 April 2020 / Published online: 8 May 2020  
© Springer Science+Business Media, LLC, part of Springer Nature 2020

## Abstract

In various capacities of statistical signal processing two-dimensional (2-D) chirp models have been considered significantly, particularly in image processing—to model gray-scale and texture images, magnetic resonance imaging, optical imaging etc. In this paper we address the problem of estimation of the unknown parameters of a 2-D chirp model under the assumption that the errors are independently and identically distributed (*i.i.d.*). The key attribute of the proposed estimation procedure is that it is computationally more efficient than the least squares estimation method. Moreover, the proposed estimators are observed to have the same asymptotic properties as the least squares estimators, thus providing computational effectiveness without any compromise on the efficiency of the estimators. We extend the propounded estimation method to provide a sequential procedure to estimate the unknown parameters of a 2-D chirp model with multiple components and under the assumption of *i.i.d.* errors we study the large sample properties of these sequential estimators. Simulation studies and two synthetic data analyses have been performed to show the effectiveness of the proposed estimators.

**Keywords** Parameter estimation · Least squares estimation method · Asymptotic properties · Sequential procedure · Simulations

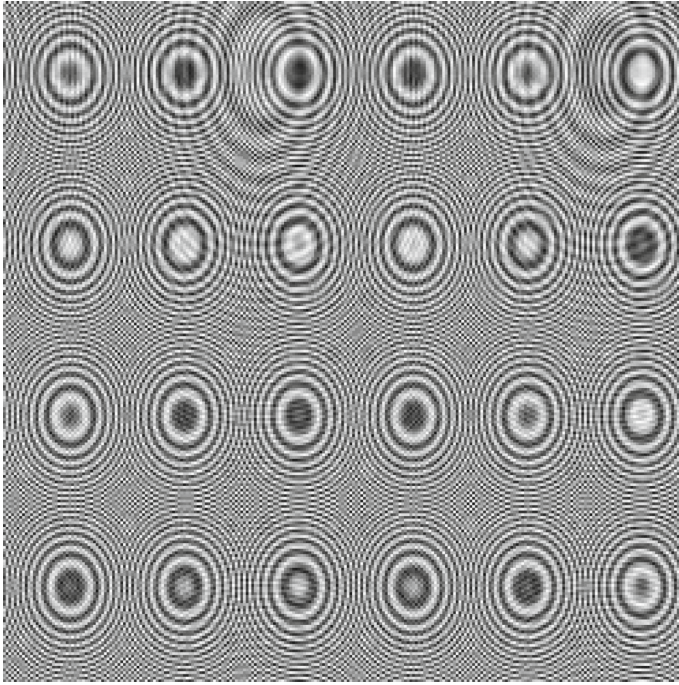
## 1 Introduction

A two-dimensional (2-D) chirp model has the following mathematical expression:

$$y(m, n) = \sum_{k=1}^P \{A_k^0 \cos(\alpha_k^0 m + \beta_k^0 m^2 + \gamma_k^0 n + \delta_k^0 n^2) + B_k^0 \sin(\alpha_k^0 m + \beta_k^0 m^2 + \gamma_k^0 n + \delta_k^0 n^2)\} + X(m, n);$$
$$m = 1, \dots, M; n = 1, \dots, N. \quad (1)$$

✉ Rhythm Grover  
rhythm@iitk.ac.in

<sup>1</sup> Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Kanpur 208016, India



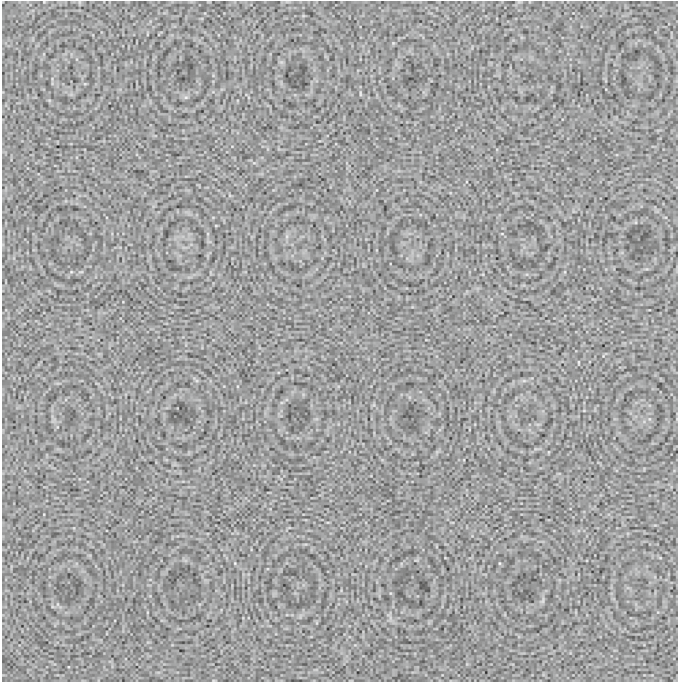
**Fig. 1** Original texture

Here,  $y(m, n)$  is the observed signal data, and the parameters  $A_k^0$ s,  $B_k^0$ s are the amplitudes,  $\alpha_k^0$ s,  $\gamma_k^0$ s are the frequencies and  $\beta_k^0$ s,  $\delta_k^0$ s are the frequency rates. The random component  $X(m, n)$  accounts for the noise component of the observed signal. In this paper, we assume that  $X(m, n)$  is an independently and identically distributed (*i.i.d.*) random field.

It can be seen that the model admits a decomposition of two components—the deterministic component and the random component. The deterministic component represents a gray-scale texture and the random component makes the model more realistic for practical realisation. For illustration, we simulate data with a fixed set of model parameters. Figure 1 represents the gray-scale texture corresponding to the simulated data without the noise component and Fig. 2 represents the contaminated texture image corresponding to the simulated data with the noise component. This clearly suggests that the 2-D chirp signal models can be used effectively in modelling and analysing black and white texture images.

Apart from the applications in image analysis, these signals are commonly observed in mobile telecommunications, surveillance systems, in radars and sonars etc. For more details on the applications, one may see the works of Francos and Friedlander (1998, 1999), Simeunović and Djurović (2016) and Zhang et al. (2008) and the references cited therein.

Parameter estimation of a 2-D chirp signal is an important statistical signal processing problem. Recently Zhang et al. (2008), Lahiri et al. (2013) and Grover et al. (2018) proposed some estimation methods of note. For instance, Zhang et al. (2008) proposed an algorithm based on the product cubic phase function for the estimation of the frequency rates of the 2-D chirp signals under low signal to noise ratio and the assumption of stationary errors. They conducted simulations to verify the performance of the proposed estimation algorithm,



**Fig. 2** Noisy texture

however there was no study of the theoretical properties of the proposed estimators. Lahiri et al. (2015) suggested the least squares estimation method. They observed that the least squares estimators (LSEs) of the unknown parameters of this model are strongly consistent and asymptotically normally distributed under the assumption of stationary additive errors. The rates of convergence of the amplitude estimates were observed to be  $M^{-1/2}N^{-1/2}$ , of the frequencies estimates, they are  $M^{-3/2}N^{-1/2}$  and  $M^{-1/2}N^{-3/2}$  and of the frequency rate estimates, they are  $M^{-5/2}N^{-1/2}$  and  $M^{-1/2}N^{-5/2}$ . Grover et al. (2018) proposed the approximate least squares estimators (ALSEs), obtained by maximising a periodogram-type function and under the same stationary error assumptions, they observed that ALSEs are strongly consistent and asymptotically equivalent to the LSEs.

A chirp signal is a particular case of the polynomial phase signal when the phase is a quadratic polynomial. Although work on parameter estimation of the aforementioned 2-D chirp model is rather limited, several authors have considered the more generalised version of this model—the 2-D polynomial phase signal model. For references, see Djurović et al. (2010), Djurović (2017), Francos and Friedlander (1998, 1999), Friedlander and Francos (1996), Lahiri and Kundu (2017), Simeunović et al. (2014), Simeunović and Djurović (2016) and Djurović and Simeunović (2018).

In this paper, we address the problem of parameter estimation of a one-component 2-D chirp model as well as the more general multiple-component 2-D chirp model. We put forward two methods for this purpose. The key characteristic of the proposed estimation method is that it reduces the foregoing 2-D chirp model into two 1-D chirp models. Thus, instead of fitting a 2-D chirp model, we are required to fit two 1-D chirp models to the given data matrix. For the fitting of 1-D models, one may use any efficient algorithm to estimate the parameters

of the obtained 1-D chirp models. Due to widespread applicability, a lot of work has been done in literature on the parameter estimation of a 1-D chirp model. For references, one may see the works of Abatzoglou (1986), Djuric and Kay (1990), Peleg and Porat (1991), Wood and Barry (1994), Barbarossa (1995), Friedlander and Francos (1995), Xia (2000), Saha and Kay (2002), Wang et al. (2010) and the references cited therein.

The proposed algorithm is numerically more efficient than the usual least squares estimation method proposed by Lahiri (2013). For instance, for a one-component 2-D chirp model, to estimate the parameters using these algorithms, we need to solve two 2-D optimisation problems as opposed to a 4-D optimisation problem in the case of finding the LSEs. This also leads to curtailment of the number of grid points required to find the initial values of the non-linear parameters as the 4-D grid search required in case of the computation of the usual LSEs or ALSEs, reduces to two 2-D grid searches. Therefore, instead of searching along a grid mesh consisting of  $M^3 N^3$  points, we need to search among only  $M^3 + N^3$  points, which is much more feasible to execute computationally. In essence, the contributions of this paper are three-fold:

1. We put forward a computationally efficient algorithm for the estimation of the unknown parameters of 2-D chirp signal models as a practical alternative to the usual least squares estimation method.
2. We examine the asymptotic properties of the proposed estimators under the assumption of *i.i.d.* errors and observe that the proposed estimators are strongly consistent and asymptotically normally distributed. In fact, they are observed to be asymptotically equivalent to the corresponding LSEs. When the errors are assumed to be Gaussian, the asymptotic variance-covariance matrix of the proposed estimators coincides with asymptotic Cramér-Rao lower bound.
3. We conduct simulation experiments and analyse a synthetic texture (see Fig. 2) to assess the effectiveness of the proposed estimators.

The rest of the paper is organised as follows. In the next section, we provide some preliminary results required to study the asymptotic properties of the proposed estimators. In Sect. 2, we consider a one-component 2-D chirp model and state the model assumptions, some notations and present the proposed algorithms along with the asymptotic properties of the proposed estimators. In Sect. 3, we extend the algorithm and develop a sequential procedure to estimate the parameters of a multiple-component 2-D chirp model. We also study the asymptotic properties of the proposed sequential estimators in this section. We perform numerical experiments for different model parameters in Sect. 4 and analyse synthetic data for illustration in Sect. 5. Finally, we conclude the paper in Sect. 6 and we provide the proofs of all the theoretical results in the appendices.

## 2 One-component 2-D chirp model

In this section, we provide the methodology to obtain the proposed estimators for the parameters of a one-component 2-D chirp model, mathematically expressed as follows:

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

Here  $y(m, n)$  is the observed data and the parameters  $A^0, B^0$  are the amplitudes,  $\alpha^0, \gamma^0$  are the frequencies and  $\beta^0, \delta^0$  are the frequency rates of the signal model. As mentioned in the introduction,  $X(m, n)$  accounts for the noise present in the signal.

We will use the following notations:  $\theta = (A, B, \alpha, \beta, \gamma, \delta)$  is the parameter vector,  $\theta^0 = (A^0, B^0, \alpha^0, \beta^0, \gamma^0, \delta^0)$  is the true parameter vector and  $\Theta_1 = [-K, K] \times [-K, K] \times [0, \pi] \times [0, \pi] \times [0, \pi] \times [0, \pi]$  is the parameter space.

### 2.1 Preliminary results

In this section, we provide the asymptotic results obtained for the usual LSEs of the unknown parameters of a 1-D chirp model by Lahiri et al. (2015). These results are later exploited to prove the asymptotic normality of the proposed estimators.

#### One-component 1-D chirp model

Consider a 1-D chirp model with the following mathematical expression:

$$y(n) = A^0 \cos(\alpha^0 n + \beta^0 n^2) + B^0 \sin(\alpha^0 n + \beta^0 n^2) + X(n). \tag{3}$$

Here  $y(n)$  is the observed data at time points  $n = 1, \dots, N$ ,  $A^0, B^0$  are the amplitudes and  $\alpha^0$  is the frequency and  $\beta^0$  is the frequency rate parameter.  $\{X(n)\}_{n=1}^N$  is the sequence of error random variables.

The LSEs of  $\alpha^0$  and  $\beta^0$  can be obtained by minimising the following reduced error sum of squares:

$$R_N(\alpha, \beta) = Q_N(\hat{A}, \hat{B}, \alpha, \beta) = \mathbf{Y}_{N \times 1}^\top (\mathbf{I} - \mathbf{P}_{\mathbf{Z}_N(\alpha, \beta)}) \mathbf{Y}_{N \times 1}$$

where,

$$Q_N(A, B, \alpha, \beta) = (\mathbf{Y}_{N \times 1} - \mathbf{Z}_N(\alpha, \beta)\boldsymbol{\phi})^\top (\mathbf{Y}_{N \times 1} - \mathbf{Z}_N(\alpha, \beta)\boldsymbol{\phi}),$$

is the error sum of squares,  $\mathbf{P}_{\mathbf{Z}_N(\alpha, \beta)} = \mathbf{Z}_N(\alpha, \beta)(\mathbf{Z}_N(\alpha, \beta)^\top \mathbf{Z}_N(\alpha, \beta))^{-1} \mathbf{Z}_N(\alpha, \beta)^\top$  is the projection matrix on the column space of the matrix  $\mathbf{Z}_N(\alpha, \beta)$ ,

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

$\mathbf{Y}_{N \times 1} = [y(1), \dots, y(N)]^\top$  is the observed data vector and  $\boldsymbol{\phi} = [A, B]^\top$  is the vector of linear parameters.

Following are the assumptions, we make on the error component and the parameters of model (3):

**Assumption P1**  $X(n)$  is a sequence of *i.i.d.* random variables with mean zero, variance  $\sigma^2$  and finite fourth order moment.

**Assumption P2**  $(A^0, B^0, \alpha^0, \beta^0)$  is an interior point of the parameter space  $\Theta = [-K, K] \times [-K, K] \times [0, \pi] \times [0, \pi]$ , where  $K$  is a positive real number and  $A^{0^2} + B^{0^2} > 0$ .

**Theorem P1** Let us denote  $\mathbf{R}'_N(\alpha, \beta)$  as the gradient vector and  $\mathbf{R}''_N(\alpha, \beta)$  as the Hessian matrix of the function  $R_N(\alpha, \beta)$ . Then, under the assumptions P1 and P2, we have:

$$-\mathbf{R}'_N(\alpha^0, \beta^0)\boldsymbol{\Delta} \xrightarrow{d} \mathcal{N}_2(\mathbf{0}, 2\sigma^2\boldsymbol{\Sigma}^{-1}) \text{ as } N \rightarrow \infty, \tag{5}$$

$$\Delta \mathbf{R}''_N(\alpha^0, \beta^0) \Delta \xrightarrow{a.s.} \Sigma^{-1} \text{ as } N \rightarrow \infty. \tag{6}$$

Here,  $\Delta = \text{diag}(\frac{1}{N\sqrt{N}}, \frac{1}{N^2\sqrt{N}})$ ,

$$\Sigma = \frac{2}{A^{0^2} + B^{0^2}} \begin{bmatrix} 96 & -90 \\ -90 & 90 \end{bmatrix} \text{ and} \tag{7}$$

$$\Sigma^{-1} = \begin{bmatrix} \frac{A^{0^2} + B^{0^2}}{12} & \frac{A^{0^2} + B^{0^2}}{45} \\ \frac{A^{0^2} + B^{0^2}}{12} & \frac{4(A^{0^2} + B^{0^2})}{45} \end{bmatrix}. \tag{8}$$

The notation  $\mathcal{N}_2(\boldsymbol{\mu}, \mathbf{V})$  means bivariate normally distributed random variable with mean vector  $\boldsymbol{\mu}_{2 \times 1}$  and variance-covariance matrix  $\mathbf{V}_{2 \times 2}$ .

**Proof** This proof follows from Theorem 2 of Lahiri et al. (2015). □

As mentioned in the introduction, any efficient algorithm can be used to estimate the parameters of a 1-D chirp model, once we reduce the dimension of the underlying model. Although the main aim of the paper is not the estimation of parameters of 1-D chirp model, but the efficient estimation of a 2-D chirp model by reducing it into 1-D models to save time. We choose to estimate the parameters of the reduced models using the LSEs as they have optimal convergence rates. For details, one may refer to Lahiri et al. (2015).

### 2.2 Proposed methodology

Let us consider the above-stated 2-D chirp signal model (2) with one-component. Suppose we fix  $n = n_0$ , then (2) can be rewritten as follows:

$$\begin{aligned} y(m, n_0) &= A^0 \cos(\alpha^0 m + \beta^0 m^2 + \gamma^0 n_0 + \delta^0 n_0^2) \\ &\quad + B^0 \sin(\alpha^0 m + \beta^0 m^2 + \gamma^0 n_0 + \delta^0 n_0^2) + X(m, n_0) \\ &= A^0(n_0) \cos(\alpha^0 m + \beta^0 m^2) + B^0(n_0) \sin(\alpha^0 m + \beta^0 m^2) \\ &\quad + X(m, n_0); \quad m = 1, \dots, M, \end{aligned} \tag{9}$$

which represents a 1-D chirp model with  $A^0(n_0), B^0(n_0)$  as the amplitudes,  $\alpha^0$  as the frequency parameter and  $\beta^0$  as the frequency rate parameter. Here,

$$\begin{aligned} A^0(n_0) &= A^0 \cos(\gamma^0 n_0 + \delta^0 n_0^2) + B^0 \sin(\gamma^0 n_0 + \delta^0 n_0^2), \text{ and} \\ B^0(n_0) &= -A^0 \sin(\gamma^0 n_0 + \delta^0 n_0^2) + B^0 \cos(\gamma^0 n_0 + \delta^0 n_0^2). \end{aligned}$$

Thus for each fixed  $n_0 \in \{1, \dots, N\}$ , we have a 1-D chirp model with the same frequency and frequency rate parameters, though different amplitudes. This 1-D model corresponds to a column of the 2-D data matrix.

Our aim is to estimate the non-linear parameters  $\alpha^0$  and  $\beta^0$  from the columns of the data matrix and one of the most reasonable estimators for this purpose are the least squares estimators. Therefore, the estimators of  $\alpha^0$  and  $\beta^0$  can be obtained by minimising the following function:

$$R_M(\alpha, \beta, n_0) = \mathbf{Y}_{n_0}^\top (\mathbf{I} - \mathbf{P}_{\mathbf{Z}_M(\alpha, \beta)}) \mathbf{Y}_{n_0}$$

for each  $n_0$ . Here,  $\mathbf{Y}_{n_0} = [y[1, n_0], \dots, y[M, n_0]]^\top$  is the  $n_0$ th column of the original data matrix,  $\mathbf{P}_{\mathbf{Z}_M(\alpha, \beta)} = \mathbf{Z}_M(\alpha, \beta)(\mathbf{Z}_M(\alpha, \beta)^\top \mathbf{Z}_M(\alpha, \beta))^{-1} \mathbf{Z}_M(\alpha, \beta)^\top$  is the projection matrix on the column space of the matrix  $\mathbf{Z}_M(\alpha, \beta)$  and the matrix  $\mathbf{Z}_M(\alpha, \beta)$  can be obtained by

replacing  $N$  by  $M$  in (4). This process involves minimising  $N$  2-D functions corresponding to the  $N$  columns of the matrix. Thus, for computational efficiency, we propose to minimise the following function instead:

$$R_{MN}^{(1)}(\alpha, \beta) = \sum_{n_0=1}^N R_M(\alpha, \beta, n_0) = \sum_{n_0=1}^N \mathbf{Y}_{n_0}^\top (\mathbf{I} - \mathbf{P}_{\mathbf{Z}_M}(\alpha, \beta)) \mathbf{Y}_{n_0} \tag{10}$$

with respect to  $\alpha$  and  $\beta$  simultaneously and obtain  $\hat{\alpha}$  and  $\hat{\beta}$  which reduces the estimation process to solving only one 2-D optimisation problem. Note that since the errors are assumed to be *i.i.d.* replacing these  $N$  functions by their sum is justifiable.

Similarly, we can obtain the estimates,  $\hat{\gamma}$  and  $\hat{\delta}$ , of  $\gamma^0$  and  $\delta^0$ , by minimising the following criterion function:

$$R_{MN}^{(2)}(\gamma, \delta) = \sum_{m_0=1}^M R_N(\gamma, \delta, m_0) = \sum_{m_0=1}^M \mathbf{Y}_{m_0}^\top (\mathbf{I} - \mathbf{P}_{\mathbf{Z}_N}(\gamma, \delta)) \mathbf{Y}_{m_0} \tag{11}$$

with respect to  $\gamma$  and  $\delta$  simultaneously. The data vector  $\mathbf{Y}_{m_0} = [y[m_0, 1], \dots, y[m_0, N]]^\top$ , is the  $m_0$ th row of the data matrix,  $m_0 = 1, \dots, M$ ,  $\mathbf{P}_{\mathbf{Z}_N}(\gamma, \delta)$  is the projection matrix on the column space of the matrix  $\mathbf{Z}_N(\gamma, \delta)$  and the matrix  $\mathbf{Z}_N(\gamma, \delta)$  can be obtained by replacing  $\alpha$  and  $\beta$  by  $\gamma$  and  $\delta$  respectively in the matrix  $\mathbf{Z}_N(\alpha, \beta)$ , defined in (4).

Once we have the estimates of the non-linear parameters, we estimate the linear parameters by the usual least squares regression technique as proposed by Lahiri et al. (2015):

$$\begin{bmatrix} \hat{A} \\ \hat{B} \end{bmatrix} = [\mathbf{W}(\hat{\alpha}, \hat{\beta}, \hat{\gamma}, \hat{\delta})^\top \mathbf{W}(\hat{\alpha}, \hat{\beta}, \hat{\gamma}, \hat{\delta})]^{-1} \mathbf{W}(\hat{\alpha}, \hat{\beta}, \hat{\gamma}, \hat{\delta})^\top \mathbf{Y}_{MN \times 1}.$$

Here,  $\mathbf{Y}_{MN \times 1} = [y(1, 1), \dots, y(M, 1), \dots, y(1, N), \dots, y(M, N)]^\top$  is the observed data vector, and

$$\mathbf{W}(\alpha, \beta, \gamma, \delta)_{MN \times 2} = \begin{bmatrix} \cos(\alpha + \beta + \gamma + \delta) & \sin(\alpha + \beta + \gamma + \delta) \\ \cos(2\alpha + 4\beta + \gamma + \delta) & \sin(2\alpha + 4\beta + \gamma + \delta) \\ \vdots & \vdots \\ \cos(M\alpha + M^2\beta + \gamma + \delta) & \sin(M\alpha + M^2\beta + \gamma + \delta) \\ \vdots & \vdots \\ \cos(\alpha + \beta + N\gamma + N^2\delta) & \sin(\alpha + \beta + N\gamma + N^2\delta) \\ \cos(2\alpha + 4\beta + N\gamma + N^2\delta) & \sin(2\alpha + 4\beta + N\gamma + N^2\delta) \\ \vdots & \vdots \\ \cos(M\alpha + M^2\beta + N\gamma + N^2\delta) & \sin(M\alpha + M^2\beta + N\gamma + N^2\delta) \end{bmatrix}. \tag{12}$$

We make the following assumptions on the error component and the model parameters before we examine the asymptotic properties of the proposed estimators:

**Assumption 1**  $X(m,n)$  is a double array sequence of *i.i.d.* random variables with mean zero, variance  $\sigma^2$  and finite fourth order moment.

**Assumption 2** The true parameter vector  $\theta^0$  is an interior point of the parametric space  $\Theta_1$ , and  $A^{0^2} + B^{0^2} > 0$ .

## 2.3 Consistency

The results obtained on the consistency of the proposed estimators are presented in the following theorems:

**Theorem 1** *Under Assumptions 1 and 2,  $\hat{\alpha}$  and  $\hat{\beta}$  are strongly consistent estimators of  $\alpha^0$  and  $\beta^0$  respectively, that is,*

$$\begin{aligned}\hat{\alpha} &\xrightarrow{a.s.} \alpha^0 \text{ as } M \rightarrow \infty. \\ \hat{\beta} &\xrightarrow{a.s.} \beta^0 \text{ as } M \rightarrow \infty.\end{aligned}$$

**Proof** See “Appendix A”. □

**Theorem 2** *Under Assumptions 1 and 2,  $\hat{\gamma}$  and  $\hat{\delta}$  are strongly consistent estimators of  $\gamma^0$  and  $\delta^0$  respectively, that is,*

$$\begin{aligned}\hat{\gamma} &\xrightarrow{a.s.} \gamma^0 \text{ as } N \rightarrow \infty. \\ \hat{\delta} &\xrightarrow{a.s.} \delta^0 \text{ as } N \rightarrow \infty.\end{aligned}$$

**Proof** This proof follows along the same lines as the proof of Theorem 1. □

## 2.4 Asymptotic distribution

The following theorems provide the asymptotic distributions of the proposed estimators:

**Theorem 3** *If the Assumptions 1 and 2 are satisfied, then*

$$[(\hat{\alpha} - \alpha^0), (\hat{\beta} - \beta^0)] \mathbf{D}_1^{-1} \xrightarrow{d} \mathcal{N}_2(\mathbf{0}, 2\sigma^2 \mathbf{\Sigma}) \text{ as } M \rightarrow \infty.$$

Here,  $\mathbf{D}_1 = \text{diag}(M^{\frac{-3}{2}} N^{\frac{-1}{2}}, M^{\frac{-5}{2}} N^{\frac{-1}{2}})$  and  $\mathbf{\Sigma}$  is as defined in (7).

**Proof** See “Appendix A”. □

**Theorem 4** *If the Assumptions 1 and 2 are satisfied, then*

$$[(\hat{\gamma} - \gamma^0), (\hat{\delta} - \delta^0)] \mathbf{D}_2^{-1} \xrightarrow{d} \mathcal{N}_2(\mathbf{0}, 2\sigma^2 \mathbf{\Sigma}) \text{ as } N \rightarrow \infty.$$

Here,  $\mathbf{D}_2 = \text{diag}(M^{\frac{-1}{2}} N^{\frac{-3}{2}}, M^{\frac{-1}{2}} N^{\frac{-5}{2}})$  and  $\mathbf{\Sigma}$  is as defined in (7).

**Proof** This proof follows along the same lines as the proof of Theorem 3. □

The asymptotic distributions of  $(\hat{\alpha}, \hat{\beta})$  and  $(\hat{\gamma}, \hat{\delta})$  are observed to be the same as those of the corresponding LSEs. Thus, we get the same efficiency as that of the LSEs without going through the process of actually computing the LSEs.

## 3 Multiple-component 2-D chirp model

In this section, we consider the multipl-component 2-D chirp model with  $p$  number of components, with the mathematical expression of the model as given in (1). Although estimation of  $p$  is an important problem and will be under our future research, in this paper we deal



with the estimation of the other important parameters characterising the observed signal, the amplitudes, the frequencies and the frequency rates, assuming  $p$  to be known. We propose a sequential procedure to estimate these parameters. The main idea supporting the proposed sequential procedure is same as that behind the ones proposed by Prasad et al. (2008) for a sinusoidal model and Lahiri et al. (2015) for a chirp model—the orthogonality of different regressor vectors. Along with the computational efficiency, the sequential method provides estimators with the same rates of convergence as the LSEs.

We use the following notation hereafter: For  $k = 1, \dots, p$ ,  $\theta_k = (A_k, B_k, \alpha_k, \beta_k, \gamma_k, \delta_k)$  is the parameter vector,  $\theta_k^0 = (A_k^0, B_k^0, \alpha_k^0, \beta_k^0, \gamma_k^0, \delta_k^0)$  is the true parameter vector.

### 3.1 Preliminary results

#### Multiple-component 1-D chirp model

Now we consider a 1-D chirp model with multiple components, mathematically expressed as follows:

$$y(n) = \sum_{k=1}^p \{A_k^0 \cos(\alpha_k^0 n + \beta_k^0 n^2) + B_k^0 \sin(\alpha_k^0 n + \beta_k^0 n^2)\} + X(n); \quad n = 1, \dots, N.$$

Here,  $A_k^0$ s,  $B_k^0$ s are the amplitudes,  $\alpha_k^0$ s are the frequencies and  $\beta_k^0$  are the frequency rates, the parameters that characterise the observed signal  $y(n)$  and  $X(n)$  is the random noise component.

Lahiri et al. (2015) suggested a sequential procedure to estimate the unknown parameters of the above model. We discuss in brief, the proposed sequential procedure and then state some of the asymptotic results they established, germane to our work.

**Step 1** The first step of the sequential method is to obtain the estimates,  $\hat{\alpha}_1$  and  $\hat{\beta}_1$ , of the non-linear parameters of the first component,  $\alpha_1^0$  and  $\beta_1^0$ , by minimising the following reduced error sum of squares:

$$R_{1,N}(\alpha_1, \beta_1) = \mathbf{Y}_{N \times 1}^\top (\mathbf{I} - \mathbf{P}_{\mathbf{Z}_N(\alpha_1, \beta_1)}) \mathbf{Y}_{N \times 1}$$

with respect to  $\alpha$  and  $\beta$  simultaneously.

**Step 2** Then the first component linear parameter estimates,  $\hat{A}_1$  and  $\hat{B}_1$  are obtained using the separable linear regression of Richards Richards (1961) as follows:

$$\begin{bmatrix} \hat{A}_1 \\ \hat{B}_1 \end{bmatrix} = [\mathbf{Z}_N(\hat{\alpha}_1, \hat{\beta}_1)^\top \mathbf{Z}_N(\hat{\alpha}_1, \hat{\beta}_1)]^{-1} \mathbf{Z}_N(\hat{\alpha}_1, \hat{\beta}_1)^\top \mathbf{Y}_{N \times 1}.$$

**Step 3** Once we have the estimates of the first component parameters, we take out its effect from the original signal and obtain a new data vector as follows:

$$\mathbf{Y}_1 = \mathbf{Y}_{N \times 1} - \mathbf{Z}_N(\hat{\alpha}_1, \hat{\beta}_1) \begin{bmatrix} \hat{A}_1 \\ \hat{B}_1 \end{bmatrix}.$$

**Step 4** Then the estimates of the second component parameters are obtained by using the new data vector and following the same procedure and the process is repeated  $p$  times.

Under the Assumption P1 on the error random variables and the following assumption on the parameters:

**Assumption P3**  $(A_k^0, B_k^0, \alpha_k^0, \beta_k^0)$  is an interior point of  $\Theta$ , for all  $k = 1, \dots, p$  and the frequencies and the frequency rates are such that  $(\alpha_i^0, \beta_i^0) \neq (\alpha_j^0, \beta_j^0) \forall i \neq j, i, j = 1, \dots, p$ .

**Assumption P4**  $A_k^0$ s and  $B_k^0$ s satisfy the following relationship:

$$K^2 > A_1^0{}^2 + B_1^0{}^2 > A_2^0{}^2 + B_2^0{}^2 > \dots > A_p^0{}^2 + B_p^0{}^2 > 0,$$

we have the following results.

**Theorem P2** Let us denote  $\mathbf{R}'_{k,N}(\alpha, \beta)$  as the gradient vector and  $\mathbf{R}''_{k,N}(\alpha, \beta)$  as the Hessian matrix of the function  $R_{k,N}(\alpha, \beta)$ ,  $k = 1, \dots, p$ . Then, under the Assumptions P1, P3 and P4:

$$-\frac{1}{\sqrt{N}}\mathbf{R}'_{k,N}(\alpha^0, \beta^0)\Delta \xrightarrow{a.s.} 0 \text{ as } N \rightarrow \infty, \tag{13}$$

$$-\mathbf{R}'_{k,N}(\alpha^0, \beta^0)\Delta \xrightarrow{d} \mathcal{N}_2(\mathbf{0}, 2\sigma^2\mathbf{\Sigma}_k^{-1}) \text{ as } N \rightarrow \infty, \tag{14}$$

$$\Delta\mathbf{R}''_{k,N}(\alpha^0, \beta^0)\Delta \xrightarrow{a.s.} \mathbf{\Sigma}_k^{-1} \text{ as } N \rightarrow \infty. \tag{15}$$

Here,  $\Delta$  is as defined in Theorem P1,

$$\mathbf{\Sigma}_k = \frac{2}{A_k^0{}^2 + B_k^0{}^2} \begin{bmatrix} 96 & -90 \\ -90 & 90 \end{bmatrix} \text{ and} \tag{16}$$

$$\mathbf{\Sigma}_k^{-1} = \begin{bmatrix} \frac{A_k^0{}^2 + B_k^0{}^2}{12} & \frac{A_k^0{}^2 + B_k^0{}^2}{45} \\ \frac{A_k^0{}^2 + B_k^0{}^2}{12} & \frac{4(A_k^0{}^2 + B_k^0{}^2)}{45} \end{bmatrix}. \tag{17}$$

**Proof** The proof of (13) follows along the same lines as proof of Lemma 4 of Lahiri et al. (2015) and that of (14) and (15) follows from Theorem 2 of Lahiri et al. (2015). Note that Lahiri et al. (2015) showed that the sequential LSEs have the same asymptotic distribution as the usual LSEs based on a number theory conjecture.  $\square$

### 3.2 Proposed sequential algorithm

The following algorithm is a simple extension of the method proposed to obtain the estimators for a one-component 2-D model in Sect. 2.2:

**Step 1** Compute  $\hat{\alpha}_1$  and  $\hat{\beta}_1$  by minimising the following function:

$$R_{1,MN}^{(1)}(\alpha_1, \beta_1) = \sum_{n_0=1}^N \mathbf{Y}_{n_0}^\top (\mathbf{I} - \mathbf{P}_{\mathbf{Z}_M}(\alpha_1, \beta_1))\mathbf{Y}_{n_0}$$

with respect to  $\alpha_1$  and  $\beta_1$  simultaneously.

**Step 2** Compute  $\hat{\gamma}_1$  and  $\hat{\delta}_1$  by minimising the function:

$$R_{1,MN}^{(2)}(\gamma_1, \delta_1) = \sum_{m_0=1}^M \mathbf{Y}_{m_0}^\top (\mathbf{I} - \mathbf{P}_{\mathbf{Z}_N}(\gamma_1, \delta_1))\mathbf{Y}_{m_0}$$

with respect to  $\gamma_1$  and  $\delta_1$  simultaneously.

**Step 3** Once the nonlinear parameters of the first component of the model are estimated, estimate the linear parameters  $A_1^0$  and  $B_1^0$  by the usual least squares estimation technique:

$$\begin{bmatrix} \hat{A}_1 \\ \hat{B}_1 \end{bmatrix} = [\mathbf{W}(\hat{\alpha}_1, \hat{\beta}_1, \hat{\gamma}_1, \hat{\delta}_1)^\top \mathbf{W}(\hat{\alpha}_1, \hat{\beta}_1, \hat{\gamma}_1, \hat{\delta}_1)]^{-1} \mathbf{W}(\hat{\alpha}_1, \hat{\beta}_1, \hat{\gamma}_1, \hat{\delta}_1)^\top \mathbf{Y}_{MN \times 1}.$$

Here,  $\mathbf{Y}_{MN \times 1} = [y(1, 1), \dots, y(M, 1), \dots, y(1, N), \dots, y(M, N)]^\top$  is the observed data vector, and the matrix  $\mathbf{W}(\hat{\alpha}_1, \hat{\beta}_1, \hat{\gamma}_1, \hat{\delta}_1)$  can be obtained by replacing  $\alpha, \beta, \gamma$  and  $\delta$  by  $\hat{\alpha}_1, \hat{\beta}_1, \hat{\gamma}_1$  and  $\hat{\delta}_1$  respectively in (12).

**Step 4** Eliminate the effect of the first component from the original data and construct new data as follows:

$$\begin{aligned} y_1(m, n) &= y(m, n) - \hat{A}_1 \cos(\hat{\alpha}_1 m + \hat{\beta}_1 m^2 + \hat{\gamma}_1 n + \hat{\delta}_1 n^2) \\ &\quad - \hat{B}_1 \sin(\hat{\alpha}_1 m + \hat{\beta}_1 m^2 + \hat{\gamma}_1 n + \hat{\delta}_1 n^2); \end{aligned} \tag{18}$$

$$m = 1, \dots, M; n = 1, \dots, N.$$

**Step 5:** Using the new data, estimate the parameters of the second component by following the same procedure.

**Step 6:** Continue this process until the parameters of all the  $p$  components are estimated.

In the following sections, we examine the asymptotic properties of the proposed estimators under the Assumptions 1, P4 and the following assumption on the parameters:

**Assumption 3**  $\theta_k^0$  is an interior point of  $\Theta_1$ , for all  $k = 1, \dots, p$  and the frequencies  $\alpha_k^0 s, \gamma_k^0 s$  and the frequency rates  $\beta_k^0 s, \delta_k^0 s$  are such that  $(\alpha_i^0, \beta_i^0) \neq (\alpha_j^0, \beta_j^0)$  and  $(\gamma_i^0, \delta_i^0) \neq (\gamma_j^0, \delta_j^0) \forall i \neq j, i, j = 1, \dots, p$ .

### 3.3 Consistency

Through the following theorems, we proclaim the consistency of the proposed estimators when the number of components,  $p$  is unknown.

**Theorem 5** *If Assumptions 1, 3 and P4 are satisfied, then the following results hold true for  $1 \leq k \leq p$ :*

$$\begin{aligned} \hat{\alpha}_k &\xrightarrow{a.s.} \alpha_k^0 \text{ as } M \rightarrow \infty, \\ \hat{\beta}_k &\xrightarrow{a.s.} \beta_k^0 \text{ as } M \rightarrow \infty. \end{aligned}$$

**Proof** See ‘‘Appendix B’’. □

**Theorem 6** *If Assumptions 1, 3 and P4 are satisfied, then the following results hold true for  $1 \leq k \leq p$ :*

$$\begin{aligned} \hat{\gamma}_k &\xrightarrow{a.s.} \gamma_k^0 \text{ as } N \rightarrow \infty, \\ \hat{\delta}_k &\xrightarrow{a.s.} \delta_k^0 \text{ as } N \rightarrow \infty. \end{aligned}$$

**Proof** This proof can be obtained along the same lines as proof of Theorem 5. □

**Theorem 7** *If the Assumptions 1, 3 and P4 are satisfied, and if  $\hat{A}_k$ ,  $\hat{B}_k$ ,  $\hat{\alpha}_k$ ,  $\hat{\beta}_k$ ,  $\hat{\gamma}_k$  and  $\hat{\delta}_k$  are the estimators obtained at the  $k$ -th step, then for  $1 \leq k \leq p$ ,*

$$\begin{aligned}\hat{A}_k &\xrightarrow{a.s.} A_k^0 \text{ as } \min\{M, N\} \rightarrow \infty \\ \hat{B}_k &\xrightarrow{a.s.} B_k^0 \text{ as } \min\{M, N\} \rightarrow \infty,\end{aligned}$$

and for  $k > p$ ,

$$\begin{aligned}\hat{A}_k &\xrightarrow{a.s.} 0 \text{ as } \min\{M, N\} \rightarrow \infty \\ \hat{B}_k &\xrightarrow{a.s.} 0 \text{ as } \min\{M, N\} \rightarrow \infty.\end{aligned}$$

**Proof** See ‘‘Appendix B’’. □

From the above theorem, it is clear that if the number of components of the fitted model is less than or same as the true number of components,  $p$ , then the amplitude estimators converge to their true values almost surely, else if it is more than  $p$ , then the amplitude estimators upto the  $p$ -th step converge to the true values and past that, they converge to zero almost surely. Thus, this result can be used as a criterion to estimate the number  $p$ . However, this might not work in low signal-to-noise ratio scenarios.

### 3.4 Asymptotic distribution

**Theorem 8** *If Assumptions 1, 3 and P4 are satisfied, then for  $1 \leq k \leq p$  :*

$$[(\hat{\alpha}_k - \alpha_k^0), (\hat{\beta}_k - \beta_k^0)] \mathbf{D}_1^{-1} \xrightarrow{d} \mathcal{N}_2(\mathbf{0}, 2\sigma^2 \mathbf{\Sigma}_k) \text{ as } M \rightarrow \infty.$$

Here  $\mathbf{D}_1$  is as defined in Theorem 3 and  $\mathbf{\Sigma}_k$  is as defined in (16).

**Proof** See ‘‘Appendix B’’. □

**Theorem 9** *If the Assumptions 1, 3 and P4 are satisfied, then*

$$[(\hat{\gamma}_k - \gamma_k^0), (\hat{\delta}_k - \delta_k^0)] \mathbf{D}_2^{-1} \xrightarrow{d} \mathcal{N}_2(\mathbf{0}, 2\sigma^2 \mathbf{\Sigma}_k) \text{ as } N \rightarrow \infty.$$

Here  $\mathbf{D}_2$  is as defined in Theorem 4 and  $\mathbf{\Sigma}_k$  is as defined in (16).

**Proof** This proof follows along the same lines as the proof of Theorem 8. □

## 4 Numerical experiments

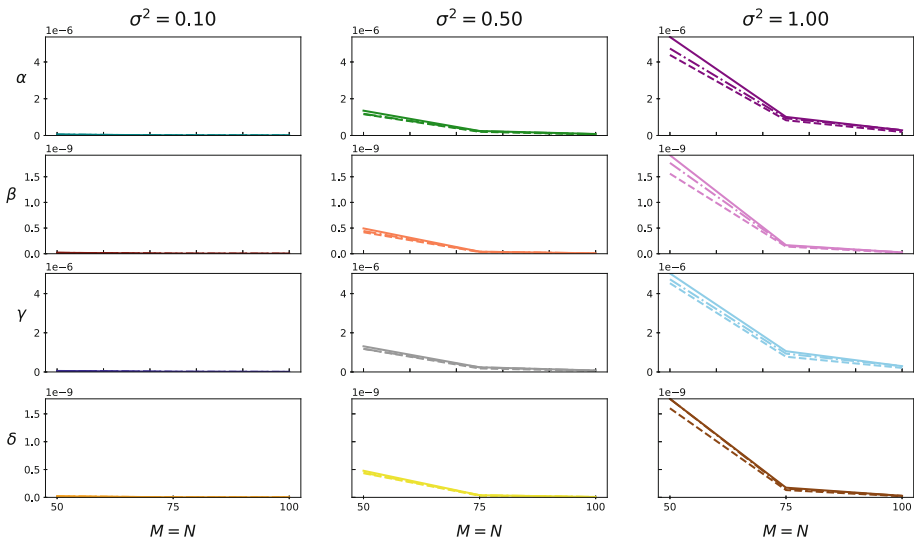
We perform simulations to examine the performance of the proposed estimators. We consider the following two cases:

**Case I** When the data are generated from a one-component model (2), with the following set of parameters:

$$A^0 = 2, B^0 = 3, \alpha^0 = 1.5, \beta^0 = 0.5, \gamma^0 = 2.5 \text{ and } \delta^0 = 0.75.$$

**Case II** When the data are generated from a two components model (1), with the following set of parameters:

$$A_1^0 = 5, B_1^0 = 4, \alpha_1^0 = 2.1, \beta_1^0 = 0.1, \gamma_1^0 = 1.25 \text{ and } \delta_1^0 = 0.25, A_2^0 = 3, B_2^0 = 2, \alpha_2^0 = 1.5, \beta_2^0 = 0.5, \gamma_2^0 = 1.75 \text{ and } \delta_2^0 = 0.75.$$



**Fig. 3** MSEs of the proposed estimates (solid lines) along with the MSEs of the LS estimates (dashed lines) and the corresponding CRLBs (dashed-dotted lines) of the estimators of the parameters of the one-component simulated model versus sample size

The noise used in the simulations is generated from Gaussian distribution with mean 0 and variance  $\sigma^2$ . Also, different values of the error variance,  $\sigma^2$  and sample sizes,  $M$  and  $N$  are considered. We estimate the parameters using the proposed estimation technique as well as the least squares estimation technique for **Case I** and for **Case II**, the proposed sequential technique and the sequential least squares technique proposed by Lahiri et al. (2015) are employed for comparison. For each case, the procedure is replicated 1000 times and the average values of the estimates, the average biases and the mean square errors (MSEs) are reported. The collation of the MSEs and the theoretical asymptotic variances exhibits the efficacy of the proposed estimation method.

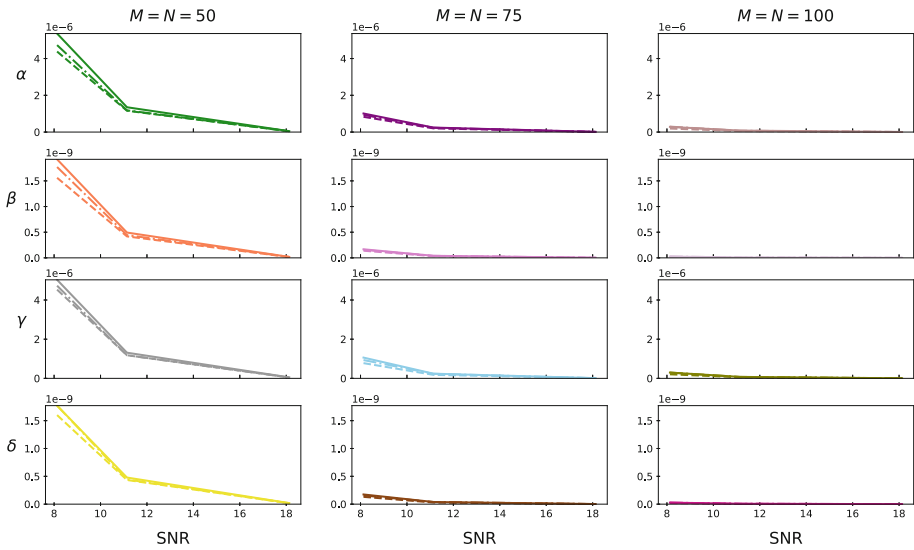
### 4.1 One-component simulation results

#### 4.1.1 Estimation performance versus sample size

In this section, we show how the proposed estimators perform in comparison with the usual LSEs (in this particular case, the usual LSEs are the MLEs as the noise is Gaussian) as well as the asymptotic CRLBs, as the sample size varies. We vary  $M$  and  $N$  from 50 to 100 each. In Fig. 3, we plot the MSEs of both the estimators along with the asymptotic CRLBs. As can be seen from this figure, the MSEs of the propounded estimates as well as the least squares estimates are well-matched with the corresponding theoretical asymptotic variances. This verifies consistency of the proposed estimators.

#### 4.1.2 Estimation performance versus signal-to-noise ratio (SNR)

Here we study the performance of the proposed estimates, again in comparison with the usual LSEs and the corresponding CRLBs, but as the SNR varies. The SNR Stoica et al. (1997) is



**Fig. 4** MSEs of the proposed estimates (solid lines) along with the MSEs of the LS estimates (dashed lines) and the corresponding CRLBs (dashed-dotted lines) of the estimators of the parameters of the one-component simulated model versus SNR

defined as follows:

$$\text{SNR} = 10 \log_{10} \left( \frac{A^{0^2} + B^{0^2}}{\sigma^2} \right)$$

Figure 4 displays this comparison. It can be seen from the figure that for each of the cases, the MSEs of the proposed estimates almost coincides with the corresponding asymptotic CRLBs, particularly for increasing values of SNR and the dimensions of the data matrix.

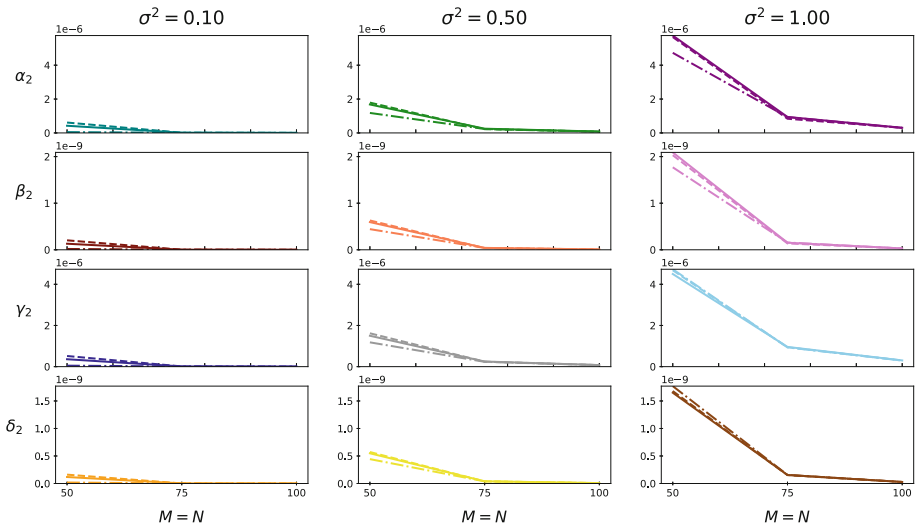
### 4.1.3 Comparison of computational complexity

The efficient method in this paper is proposed to reduce the computational burden involved in finding the usual LSEs of the parameters of model (1). In the above sections, we see that the proposed estimators perform nearly as accurately as the usual LSEs. Here we demonstrate how the computational complexity is reduced significantly through the following example.

We consider the following model:

$$\begin{aligned} y(m, n) &= 2 \cos(1.5m + 0.5m^2 + 2.5n + 0.75n^2) \\ &\quad + 3 \sin(1.5m + 0.5m^2 + 2.5n + 0.75n^2) + X(m, n); \\ m &= 1, \dots, 10, n = 1, \dots, 10. \end{aligned}$$

As explained in the introduction, finding the initial values to compute the proposed estimators involves two 2-D searches which means precisely  $2 \times 9 \times 99 = 1782$  evaluations of the objective function to be minimised. The time taken for the computation of these initial values is 0.81 seconds. On the other hand, to compute the initial guesses for the usual LSEs of the non-linear parameters of the above model, we perform a 4D fine grid search in the parameter space. This requires  $9 \times 99 \times 9 \times 99 = 793881$  computations of the least squares objective function and the time taken for this search is 196.61 s. This gives us a clear idea of the



**Fig. 5** MSEs of the proposed estimates (solid lines) along with the MSEs of the LS estimates (dashed lines) and the corresponding CRLBs (dashed-dotted lines) of the estimators of the parameters of the second component of the two-component simulated model versus sample size

significant time savings if we use the proposed method rather than the least squares estimation method. Moreover, using the usual least squares method increases the computational load exponentially as the dimensions of the data matrix increase. This experiment is performed using Intel® Core™ i7-2600 @ 3.4 GHz, 4 GB RAM machines, the program codes are written in R software (3.3.0).

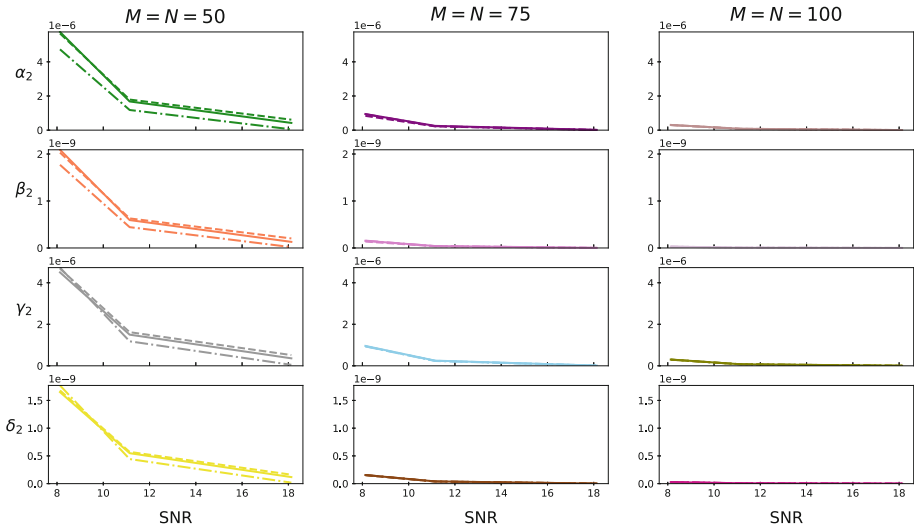
### 4.2 Two component simulation results

In this section, we present the simulation results for **Case II**. In the interest of brevity, we present the results only for the second component of the model. The estimates of the first component behave in the same manner and as they are obtained by minimising the same functions as for the one component model, we omit these results.

#### 4.2.1 Estimation performance versus sample size

We first study the performance of the proposed estimators and compare it with the corresponding sequential LSEs and the asymptotic CRLBS versus the dimensions of the simulated data matrix.

Figure 5 show the results obtained for the parameters of the second component of the underlying two-component model. These results verify consistency of the proposed sequential estimators. It is also observed that the MSEs of the obtained parameter estimates have exactly the same order as the corresponding asymptotic variances.



**Fig. 6** MSEs of the proposed estimates (solid lines) along with the MSEs of the LS estimates (dashed lines) and the corresponding CRLBs (dashed-dotted lines) of the estimators of the parameters of the second component of the two-component simulated model versus SNR

### 4.2.2 Estimation performance versus SNR

Next we study the performance of the proposed estimators as the SNR varies. For the multiple component model, SNR is defined as follows:

$$SNR_k = 10 \log_{10} \left( \frac{A_k^{0^2} + B_k^{0^2}}{\sigma^2} \right)$$

The MSEs of the proposed estimators are plotted along with those of the sequential LSEs and the respective asymptotic CRLBs in Fig. 6. The results show that the MSEs of both types of the estimates practically coincide with the corresponding CRLBs.

It can be concluded that the simulations validate the accuracy of the proposed estimators. Also, as discussed in Sect. 4.1.3, the proposed estimators are much faster to compute than the sequential LSEs. Therefore the proposed estimators achieve computational simplicity along with an excellent estimation accuracy.

## 5 Simulated data analyses

In this section, we assess the performance of the proposed algorithm by analysing two symmetric gray-scale textures, one generated using a one-component 2-D chirp model and the other using a multiple-component 2-D chirp model.



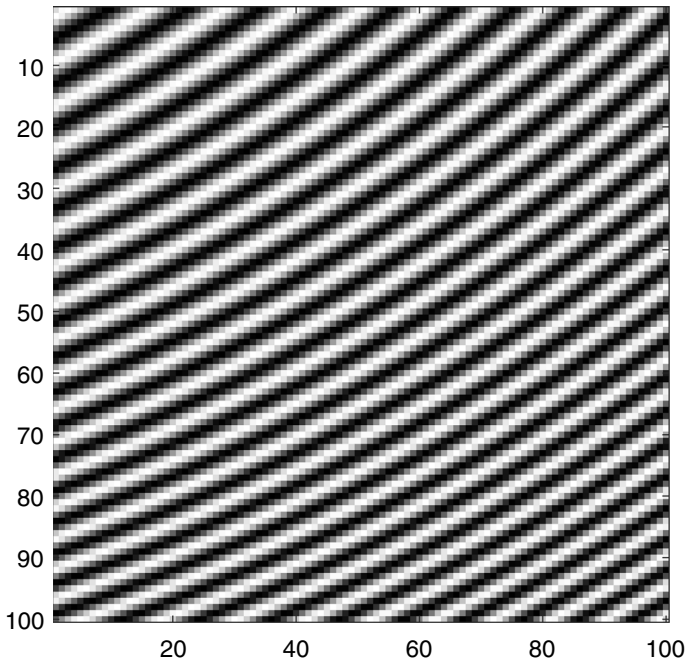


Fig. 7 Original texture corresponding to model (19)

### 5.1 When the data set is generated from a one-component 2-D chirp model

First, we consider a one-component 2-D chirp model defined as follows:

$$\begin{aligned}
 y(m, n) = & \cos(0.45m + 0.0015m^2 + 0.82n + 0.0025n^2) \\
 & + \sin(0.45m + 0.0015m^2 + 0.82n + 0.0025n^2) \\
 & + X(m, n); m = 1, \dots, M, n = 1, \dots, N.
 \end{aligned}
 \tag{19}$$

We begin by generating data from the above defined model structure. Here,  $X(m, n)$  are *i.i.d.* Gaussian errors with mean 0 and variance 1. Figures 7 and 8 represent the true and the noisy image corresponding to the generated data.

We estimate the parameters of the underlying model using the proposed algorithm for the one-component model parameter estimation in Sect. 2.2. Figure 9 shows the estimated gray-scale image. It is evident that the the proposed method provides an accurate estimation of the original texture.

### 5.2 When the data set is generated from a multiple-component 2-D chirp model

We also analyse a synthetic texture data using model (1) to demonstrate how the proposed parameter estimation method works for a 2-D chirp signal with multiple chirp components.

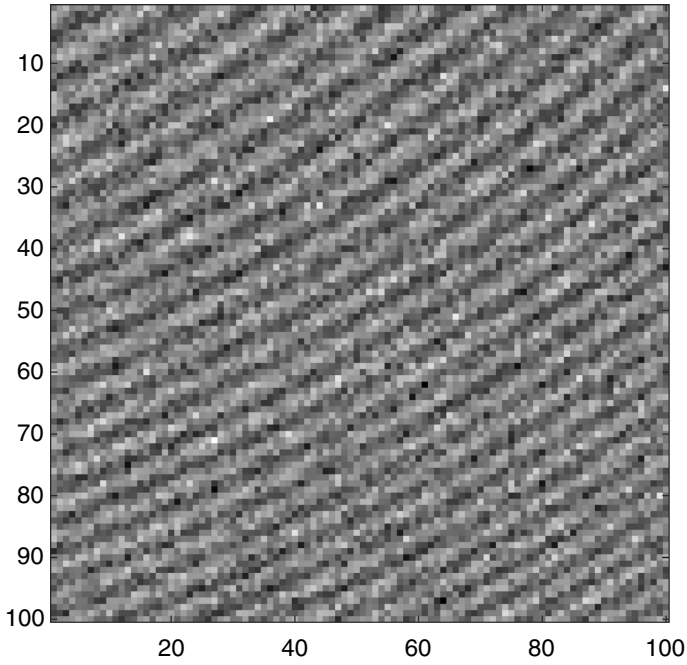


Fig. 8 Noisy texture corresponding to model (19)

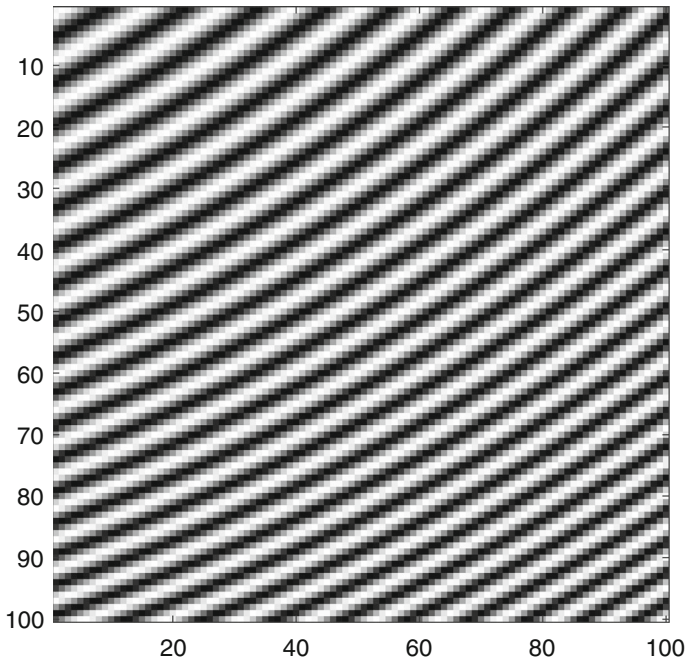
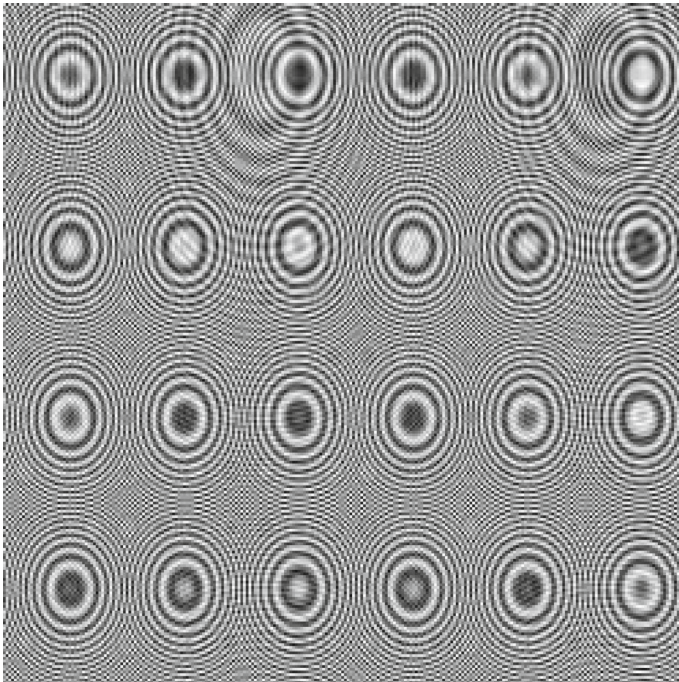


Fig. 9 Estimated texture corresponding to model (19)

**Table 1** True parameters values of the synthetic data

$A_1^0$	6	$B_1^0$	6	$\alpha_1^0$	2.75	$\beta_1^0$	0.05	$\gamma_1^0$	2.5	$\delta_1^0$	0.075
$A_2^0$	2	$B_2^0$	2	$\alpha_2^0$	1.75	$\beta_2^0$	0.01	$\gamma_2^0$	1.5	$\delta_2^0$	0.025
$A_3^0$	1	$B_3^0$	1	$\alpha_3^0$	1.5	$\beta_3^0$	0.15	$\gamma_3^0$	2	$\delta_3^0$	0.25
$A_4^0$	0.5	$B_4^0$	0.5	$\alpha_4^0$	1.75	$\beta_4^0$	0.75	$\gamma_4^0$	2.75	$\delta_4^0$	0.275
$A_5^0$	0.1	$B_5^0$	0.1	$\alpha_5^0$	1.95	$\beta_5^0$	0.95	$\gamma_5^0$	2.95	$\delta_5^0$	0.295



**Fig. 10** Estimated texture for the synthetic data

The signal data are generated using the following model structure and parameters:

$$\begin{aligned}
 y(m, n) = & \sum_{k=1}^5 \{A_k^0 \cos(\alpha_k^0 m + \beta_k^0 m^2 + \gamma_k^0 n + \delta_k^0 n^2) + B_k^0 \sin(\alpha_k^0 m + \beta_k^0 m^2 + \gamma_k^0 n + \delta_k^0 n^2)\} \\
 & + X(m, n)
 \end{aligned}
 \tag{20}$$

The true parameter values are provided in Table 1. The errors  $X(m, n)$ s are *i.i.d.* Gaussian random variables with mean 0 and variance 100. Figure 1 represents the original texture without any contamination and Fig. 2 represents the noisy texture. Our purpose is to extract the original gray-scale texture from the one which is contaminated.

We estimate the parameters of model (20) using the proposed sequential procedure and the estimated texture is plotted in Fig. 10. It is clear that the estimated texture and the original texture look extremely well-matched.

## 6 Concluding remarks

In this paper, we have considered the estimation of unknown parameters of a 2-D chirp model under the assumption of *i.i.d.* additive errors. The main idea is to reduce the computational complexity involved in finding the LSEs of these parameters. The proposed estimators minimise the computations to a great extent and are observed to be strongly consistent and asymptotically equivalent to the LSEs. For a 2-D chirp model with  $p$  number of components, we have proposed a sequential procedure which reduces the problem of estimation of the parameters to solving  $p$  number of 2-D optimisation problems. Moreover, the propounded sequential estimators are observed to be strongly consistent and asymptotically equivalent to the usual LSEs.

The numerical experiments—the simulations and the data analysis, show that the proposed estimation technique provides as accurate results as the least squares estimation method with the additional advantage of being computationally more efficient. Thus to summarise, the proposed estimators seem to be the method of choice as their performance is satisfactory and as efficient as the LSEs, both numerically and analytically.

**Acknowledgements** The authors would like to thank the the Editor and the two unknown reviewers for their positive assessment of the manuscript and their constructive comments.

## Appendix A

Henceforth, we will denote  $\theta(n_0) = (A(n_0), B(n_0), \alpha, \beta)$  as the parameter vector and  $\theta^0(n_0) = (A^0(n_0), B^0(n_0), \alpha^0, \beta^0)$  as the true parameter vector of the 1-D chirp model (9).

To prove Theorem 1, we need the following lemma:

**Lemma 1** Consider the set  $S_c = \{(\alpha, \beta) : |\alpha - \alpha^0| \geq c \text{ or } |\beta - \beta^0| \geq c\}$ . If for any  $c > 0$ ,

$$\liminf \inf_{(\alpha, \beta) \in S_c} \frac{1}{MN} \left[ R_{MN}^{(1)}(\alpha, \beta) - R_{MN}^{(1)}(\alpha^0, \beta^0) \right] > 0 \text{ a.s.} \quad (21)$$

then,  $\hat{\alpha} \rightarrow \alpha^0$  and  $\hat{\beta} \rightarrow \beta^0$  almost surely as  $M \rightarrow \infty$ . Note that  $\liminf$  stands for limit infimum and  $\inf$  stands for the infimum.

**Proof** This proof follows along the same lines as that of Lemma 1 of Wu (1981).  $\square$

**Proof of Theorem 1:** Let us consider the following:

$$\begin{aligned} & \liminf \inf_{(\alpha, \beta) \in S_c} \frac{1}{MN} \left[ R_{MN}^{(1)}(\alpha, \beta) - R_{MN}^{(1)}(\alpha^0, \beta^0) \right] \\ &= \liminf \inf_{(\alpha, \beta) \in S_c} \frac{1}{MN} \left[ \sum_{n_0=1}^N R_M(\alpha, \beta, n_0) - \sum_{n_0=1}^N R_M(\alpha^0, \beta^0, n_0) \right] \\ &= \liminf \inf_{(\alpha, \beta) \in S_c} \frac{1}{MN} \left[ \sum_{n_0=1}^N Q_M(\hat{A}(n_0), \hat{B}(n_0), \alpha, \beta) - \sum_{n_0=1}^N Q_M(\hat{A}(n_0), \hat{B}(n_0), \alpha^0, \beta^0) \right] \\ &\geq \liminf \inf_{(\alpha, \beta) \in S_c} \frac{1}{MN} \left[ \sum_{n_0=1}^N Q_M(\hat{A}(n_0), \hat{B}(n_0), \alpha, \beta) - \sum_{n_0=1}^N Q_M(A^0(n_0), B^0(n_0), \alpha^0, \beta^0) \right] \end{aligned}$$

$$\begin{aligned} &\geq \liminf \inf_{\theta(n_0) \in M_c^{n_0}} \frac{1}{MN} \left[ \sum_{n_0=1}^N Q_M(A(n_0), B(n_0), \alpha, \beta) - \sum_{n_0=1}^N Q_M(A^0(n_0), B^0(n_0), \alpha^0, \beta^0) \right] \\ &\geq \frac{1}{N} \sum_{n_0=1}^N \liminf \inf_{\theta(n_0) \in M_c^{n_0}} \frac{1}{M} \left[ Q_M(\theta(n_0)) - Q_M(\theta^0(n_0)) \right] > 0. \end{aligned}$$

This follows from the proof of Theorem 1 of Kundu and Nandi (2008). Here,  $Q_M(A(n_0), B(n_0), \alpha, \beta) = \mathbf{Y}_{n_0}^\top (\mathbf{I} - \mathbf{Z}_M(\alpha, \beta)(\mathbf{Z}_M(\alpha, \beta)^\top \mathbf{Z}_M(\alpha, \beta))^{-1} \mathbf{Z}_M(\alpha, \beta)^\top) \mathbf{Y}_{n_0}$ . Also note that the set  $M_c^{n_0} = \{\theta(n_0) : |A(n_0) - A^0(n_0)| \geq c \text{ or } |B(n_0) - B^0(n_0)| \geq c \text{ or } |\alpha - \alpha^0| \geq c \text{ or } |\beta - \beta^0| \geq c\}$  which implies  $S_c \subset M_c^{n_0}$ , for all  $n_0 \in \{1, \dots, N\}$ . Thus, using Lemma 1,  $\hat{\alpha} \xrightarrow{a.s.} \alpha^0$  and  $\hat{\beta} \xrightarrow{a.s.} \beta^0$ .  $\square$

**Proof of Theorem 3:** Let us denote  $\xi = (\alpha, \beta)$  and  $\hat{\xi} = (\hat{\alpha}, \hat{\beta})$ , the estimator of  $\xi^0 = (\alpha^0, \beta^0)$  obtained by minimising the function  $R_{MN}^{(1)}(\xi) = R_{MN}^{(1)}(\alpha, \beta)$  defined in (10).

Using multivariate Taylor series, we expand the  $1 \times 2$  gradient vector  $\mathbf{R}_{MN}^{(1)'}(\hat{\xi})$  of the function  $R_{MN}^{(1)}(\xi)$ , around the point  $\xi^0$  as follows:

$$\mathbf{R}_{MN}^{(1)'}(\hat{\xi}) - \mathbf{R}_{MN}^{(1)'}(\xi^0) = (\hat{\xi} - \xi^0) \mathbf{R}_{MN}^{(1)''}(\bar{\xi}),$$

where  $\bar{\xi}$  is a point between  $\hat{\xi}$  and  $\xi^0$  and  $\mathbf{R}_{MN}^{(1)''}(\bar{\xi})$  is the  $2 \times 2$  Hessian matrix of the function  $R_{MN}^{(1)}(\xi)$  at the point  $\bar{\xi}$ . Since  $\hat{\xi}$  minimises the function  $R_{MN}^{(1)}(\xi)$ ,  $\mathbf{R}_{MN}^{(1)'}(\hat{\xi}) = 0$ . Thus, we have

$$(\hat{\xi} - \xi^0) = -\mathbf{R}_{MN}^{(1)'}(\xi^0) [\mathbf{R}_{MN}^{(1)''}(\bar{\xi})]^{-1}.$$

Multiplying both sides by the diagonal matrix  $\mathbf{D}_1^{-1} = \text{diag}(M^{-\frac{3}{2}} N^{-\frac{1}{2}}, M^{-\frac{5}{2}} N^{-\frac{1}{2}})$ , we get:

$$(\hat{\xi} - \xi^0) \mathbf{D}_1^{-1} = -\mathbf{R}_{MN}^{(1)'}(\xi^0) \mathbf{D}_1 [\mathbf{D}_1 \mathbf{R}_{MN}^{(1)''}(\bar{\xi}) \mathbf{D}_1]^{-1}. \tag{22}$$

Consider the vector,

$$\mathbf{R}_{MN}^{(1)'}(\xi^0) \mathbf{D}_1 = \left[ \frac{1}{M^{3/2} N^{1/2}} \frac{\partial R_{MN}^{(1)}(\xi^0)}{\partial \alpha}, \frac{1}{M^{5/2} N^{1/2}} \frac{\partial R_{MN}^{(1)}(\xi^0)}{\partial \beta} \right].$$

On computing the elements of this vector and using preliminary result (5) (see Sect. 2.1) and the definition of the function:

$$R_{MN}^{(1)}(\alpha, \beta) = \sum_{n_0=1}^N R_M(\alpha, \beta, n_0)$$

we obtain the following result:

$$-\mathbf{R}_{MN}^{(1)'}(\xi^0) \mathbf{D}_1 \xrightarrow{d} \mathcal{N}_2(\mathbf{0}, 2\sigma^2 \Sigma) \text{ as } M \rightarrow \infty. \tag{23}$$

Since  $\hat{\xi} \xrightarrow{a.s.} \xi^0$ , and as each element of the matrix  $\mathbf{R}_{MN}^{(1)''}(\xi)$  is a continuous function of  $\xi$ , we have

$$\lim_{M \rightarrow \infty} \mathbf{D}_1 \mathbf{R}_{MN}^{(1)''}(\bar{\xi}) \mathbf{D}_1 = \lim_{M \rightarrow \infty} \mathbf{D}_1 \mathbf{R}_{MN}^{(1)''}(\xi^0) \mathbf{D}_1.$$

Now using preliminary result (6) (see Sect. 2.1), it can be seen that:

$$\lim_{M \rightarrow \infty} \mathbf{D}_1 \mathbf{R}_{MN}^{(1)''}(\xi^0) \mathbf{D}_1 \rightarrow \Sigma^{-1}. \tag{24}$$

On combining (22), (23) and (24), we have the desired result.  $\square$

### Appendix B

To prove Theorem 5, we need the following lemmas:

**Lemma 2** Consider the set  $S_c^1 = \{(\alpha, \beta) : |\alpha - \alpha_1^0| \geq c \text{ or } |\beta - \beta_1^0| \geq c\}$ . If for any  $c > 0$ ,

$$\liminf_{(\alpha, \beta) \in S_c^1} \inf_{\alpha, \beta} \frac{1}{MN} [R_{1,MN}^{(1)}(\alpha, \beta) - R_{1,MN}^{(1)}(\alpha_1^0, \beta_1^0)] > 0 \text{ a.s.} \tag{25}$$

then,  $\hat{\alpha}_1 \rightarrow \alpha_1^0$  and  $\hat{\beta}_1 \rightarrow \beta_1^0$  almost surely as  $M \rightarrow \infty$ .

**Proof** This proof follows along the same lines as proof of Lemma 1. □

**Lemma 3** If Assumptions 1, 3 and P4 are satisfied then:

$$\begin{aligned} M(\hat{\alpha}_1 - \alpha_1^0) &\xrightarrow{a.s.} 0, \\ M^2(\hat{\beta}_1 - \beta_1^0) &\xrightarrow{a.s.} 0. \end{aligned}$$

**Proof** Let us denote  $R_{1,MN}^{(1)'}(\xi)$  as the  $1 \times 2$  gradient vector and  $R_{1,MN}^{(1)''}(\xi)$  as the  $2 \times 2$  Hessian matrix of the function  $R_{1,MN}^{(1)}(\xi)$ . Using multivariate Taylor series expansion, we expand the function  $R_{1,MN}^{(1)'}(\hat{\xi}_1)$  around the point  $\xi_1^0$  as follows:

$$R_{1,MN}^{(1)'}(\hat{\xi}_1) - R_{1,MN}^{(1)'}(\xi_1^0) = (\hat{\xi}_1 - \xi_1^0) R_{1,MN}^{(1)''}(\bar{\xi}_1)$$

where  $\bar{\xi}_1$  is a point between  $\hat{\xi}_1$  and  $\xi_1^0$ . Note that  $R_{1,MN}^{(1)'}(\xi_1^0) = 0$ . Thus, we have:

$$(\hat{\xi}_1 - \xi_1^0) = -R_{1,MN}^{(1)'}(\xi_1^0) [R_{1,MN}^{(1)''}(\bar{\xi}_1)]^{-1}. \tag{26}$$

Multiplying both sides by  $\frac{1}{\sqrt{MN}} D_1^{-1}$ , we get:

$$(\hat{\xi}_1 - \xi_1^0) (\sqrt{MN} D_1)^{-1} = -\frac{1}{\sqrt{MN}} R_{1,MN}^{(1)'}(\xi_1^0) D_1 [D_1 R_{1,MN}^{(1)''}(\bar{\xi}_1) D_1]^{-1}. \tag{27}$$

Since each of the elements of the matrix  $R_{1,MN}^{(1)''}(\xi)$  is a continuous function of  $\xi$ ,

$$\lim_{M \rightarrow \infty} D_1 R_{1,MN}^{(1)''}(\bar{\xi}_1) D_1 = \lim_{M \rightarrow \infty} D_1 R_{1,MN}^{(1)''}(\xi_1^0) D_1.$$

By definition,

$$R_{1,MN}^{(1)}(\xi) = \sum_{n_0=1}^N R_{1,M}(\xi, n_0). \tag{28}$$

Using this and the preliminary result (13) and (15) (see Sect. 3.1), it can be seen that:

$$-\frac{1}{\sqrt{MN}} R_{1,MN}^{(1)'}(\xi_1^0) D_1 \xrightarrow{a.s.} 0 \text{ as } M \rightarrow \infty. \tag{29}$$

$$D_1 R_{1,MN}^{(1)''}(\xi_1^0) D_1 \xrightarrow{a.s.} \Sigma_1^{-1} \text{ as } M \rightarrow \infty. \tag{30}$$

On combining (27), (29) and (30), we have the desired result. □

**Proof of Theorem 5:** Consider the left hand side of (25), that is,

$$\begin{aligned}
 & \liminf_{(\alpha, \beta) \in S_c^!} \inf_{\alpha_1^0, \beta_1^0} \frac{1}{MN} \left[ R_{1,MN}^{(1)}(\alpha, \beta) - R_{1,MN}^{(1)}(\alpha_1^0, \beta_1^0) \right] \\
 &= \liminf_{(\alpha, \beta) \in S_c^!} \inf_{\alpha_1^0, \beta_1^0} \frac{1}{MN} \left[ \sum_{n_0=1}^N Q_{1,M}(\hat{A}_1(n_0), \hat{B}_1(n_0), \alpha, \beta) - \sum_{n_0=1}^N Q_{1,M}(\hat{A}_1(n_0), \hat{B}_1(n_0), \alpha_1^0, \beta_1^0) \right] \\
 &\geq \liminf_{(\alpha, \beta) \in S_c^!} \inf_{\alpha_1^0, \beta_1^0} \frac{1}{MN} \left[ \sum_{n_0=1}^N Q_{1,M}(\hat{A}_1(n_0), \hat{B}_1(n_0), \alpha, \beta) - \sum_{n_0=1}^N Q_{1,M}(A_1^0(n_0), B_1^0(n_0), \alpha_1^0, \beta_1^0) \right] \\
 &\geq \liminf_{\theta_1(n_0) \in M_c^{1,n_0}} \inf_{\theta_1(n_0) \in M_c^{1,n_0}} \frac{1}{MN} \left[ \sum_{n_0=1}^N Q_{1,M}(A_1(n_0), B_1(n_0), \alpha, \beta) - \sum_{n_0=1}^N Q_{1,M}(A_1^0(n_0), B_1^0(n_0), \alpha_1^0, \beta_1^0) \right] \\
 &\geq \frac{1}{N} \sum_{n_0=1}^N \liminf_{\theta_1(n_0) \in M_c^{1,n_0}} \inf_{\theta_1(n_0) \in M_c^{1,n_0}} \frac{1}{M} \left[ Q_{1,M}(\theta_1(n_0)) - Q_{1,M}(\theta_1^0(n_0)) \right] > 0.
 \end{aligned}$$

Here,  $Q_{1,M}(A(n_0), B(n_0), \alpha, \beta) = \mathbf{Y}_{n_0}^\top (\mathbf{I} - \mathbf{Z}_M(\alpha, \beta)(\mathbf{Z}_M(\alpha, \beta)^\top \mathbf{Z}_M(\alpha, \beta))^{-1} \mathbf{Z}_M(\alpha, \beta)^\top) \mathbf{Y}_{n_0}$  and  $M_c^{1,n_0}$  can be obtained by replacing  $\alpha^0$  and  $\beta^0$  by  $\alpha_1^0$  and  $\beta_1^0$  respectively, in the set  $M_c^{n_0}$  defined in Lemma 1. The last step follows from the proof of Theorem 2.4.1 of Lahiri et al. (2015). Thus, using Lemma 2,  $\hat{\alpha}_1 \xrightarrow{a.s.} \alpha_1^0$  and  $\hat{\beta}_1 \xrightarrow{a.s.} \beta_1^0$  as  $M \rightarrow \infty$ .

Following similar arguments, one can obtain the consistency of  $\hat{\gamma}_1$  and  $\hat{\delta}_1$  as  $N \rightarrow \infty$ . Also,

$$\begin{aligned}
 N(\hat{\gamma}_1 - \gamma_1^0) &\xrightarrow{a.s.} 0, \\
 N^2(\hat{\delta}_1 - \delta_1^0) &\xrightarrow{a.s.} 0.
 \end{aligned}$$

The proof of the above equations follows along the same lines as the proof of Lemma 3. From Theorem 7, it follows that as  $\min\{M, N\} \rightarrow \infty$ :

$$\begin{aligned}
 (\hat{A}_1 - A_1^0) &\xrightarrow{a.s.} 0, \\
 (\hat{B}_1 - B_1^0) &\xrightarrow{a.s.} 0.
 \end{aligned}$$

Thus, we have the following relationship between the first component of model (1) and its estimate:

$$\begin{aligned}
 & \hat{A}_1 \cos(\hat{\alpha}_1 m + \hat{\beta}_1 m^2 + \hat{\gamma}_1 n + \hat{\delta}_1 n^2) + \hat{B}_1 \sin(\hat{\alpha}_1 m + \hat{\beta}_1 m^2 + \hat{\gamma}_1 n + \hat{\delta}_1 n^2) \\
 &= A_1^0 \cos(\alpha_1^0 m + \beta_1^0 m^2 + \gamma_1^0 n + \delta_1^0 n^2) + B_1^0 \sin(\alpha_1^0 m + \beta_1^0 m^2 + \gamma_1^0 n + \delta_1^0 n^2) + o(1).
 \end{aligned} \tag{31}$$

Here a function  $g$  is  $o(1)$ , if  $g \rightarrow 0$  almost surely as  $\min\{M, N\} \rightarrow \infty$ .

Using (31) and following the same arguments as above for the consistency of  $\hat{\alpha}_1, \hat{\beta}_1, \hat{\gamma}_1$  and  $\hat{\delta}_1$ , we can show that,  $\hat{\alpha}_2, \hat{\beta}_2, \hat{\gamma}_2$  and  $\hat{\delta}_2$  are strongly consistent estimators of  $\alpha_2^0, \beta_2^0, \gamma_2^0$  and  $\delta_2^0$  respectively. And the same can be extended for  $k \leq p$ . Hence, the result.  $\square$

**Proof of Theorem 7:** We will consider the following two cases that will cover both the scenarios—underestimation as well as overestimation of the number of components:

- **Case 1** When  $k = 1$ :

$$\begin{bmatrix} \hat{A}_1 \\ \hat{B}_1 \end{bmatrix} = [\mathbf{W}(\hat{\alpha}_1, \hat{\beta}_1, \hat{\gamma}_1, \hat{\delta}_1)^\top \mathbf{W}(\hat{\alpha}_1, \hat{\beta}_1, \hat{\gamma}_1, \hat{\delta}_1)]^{-1} \mathbf{W}(\hat{\alpha}_1, \hat{\beta}_1, \hat{\gamma}_1, \hat{\delta}_1)^\top \mathbf{Y}_{MN \times 1} \tag{32}$$

Using Lemma 1 of Lahiri et al. (2015), it can be seen that:

$$\frac{1}{MN} [\mathbf{W}(\hat{\alpha}_1, \hat{\beta}_1, \hat{\gamma}_1, \hat{\delta}_1)^\top \mathbf{W}(\hat{\alpha}_1, \hat{\beta}_1, \hat{\gamma}_1, \hat{\delta}_1)] \rightarrow \frac{1}{2} \mathbf{I}_{2 \times 2} \text{ as } \min\{M, N\} \rightarrow \infty.$$

Substituting this result in (32), we get:

$$\begin{aligned} \begin{bmatrix} \hat{A}_1 \\ \hat{B}_1 \end{bmatrix} &= \frac{2}{MN} \mathbf{W}(\hat{\alpha}_1, \hat{\beta}_1, \hat{\gamma}_1, \hat{\delta}_1)^\top \mathbf{Y}_{MN \times 1} + o(1) \\ &= \begin{bmatrix} \frac{2}{MN} \sum_{n=1}^N \sum_{m=1}^M y(m, n) \cos(\hat{\alpha}_1 m + \hat{\beta}_1 m^2 + \hat{\gamma}_1 n + \hat{\delta}_1 n^2) + o(1) \\ \frac{2}{MN} \sum_{n=1}^N \sum_{m=1}^M y(m, n) \sin(\hat{\alpha}_1 m + \hat{\beta}_1 m^2 + \hat{\gamma}_1 n + \hat{\delta}_1 n^2) + o(1) \end{bmatrix}. \end{aligned}$$

Now consider the estimate  $\hat{A}_1$ . Using multivariate Taylor series, we expand the function  $\cos(\hat{\alpha}_1 m + \hat{\beta}_1 m^2 + \hat{\gamma}_1 n + \hat{\delta}_1 n^2)$  around the point  $(\alpha_1^0, \beta_1^0, \gamma_1^0, \delta_1^0)$  and we obtain:

$$\begin{aligned} \hat{A}_1 &= \frac{2}{MN} \sum_{n=1}^N \sum_{m=1}^M y(m, n) \left\{ \cos(\alpha_1^0 m + \beta_1^0 m^2 + \gamma_1^0 n + \delta_1^0 n^2) \right. \\ &\quad - m(\hat{\alpha}_1 - \alpha_1^0) \sin(\alpha_1^0 m + \beta_1^0 m^2 + \gamma_1^0 n + \delta_1^0 n^2) \\ &\quad + \delta_1^0 n^2 - m^2(\hat{\beta}_1 - \beta_1^0) \sin(\alpha_1^0 m + \beta_1^0 m^2 + \gamma_1^0 n + \delta_1^0 n^2) \\ &\quad - n(\hat{\gamma}_1 - \gamma_1^0) \sin(\alpha_1^0 m + \beta_1^0 m^2 + \gamma_1^0 n + \delta_1^0 n^2) \\ &\quad \left. + \delta_1^0 n^2 - n^2(\hat{\delta}_1 - \delta_1^0) \sin(\alpha_1^0 m + \beta_1^0 m^2 + \gamma_1^0 n + \delta_1^0 n^2) \right\} \\ &\rightarrow 2 \times \frac{A_1^0}{2} = A_1^0 \text{ almost surely as } \min\{M, N\} \rightarrow \infty, \end{aligned}$$

using (1) and Lemma 1 and Lemma 2 of Lahiri et al. (2015). Similarly, it can be shown that  $\hat{B}_1 \rightarrow B_1^0$  almost surely as  $\min\{M, N\} \rightarrow \infty$ .

For the second component linear parameter estimates, consider:

$$\begin{bmatrix} \hat{A}_2 \\ \hat{B}_2 \end{bmatrix} = \begin{bmatrix} \frac{2}{MN} \sum_{n=1}^N \sum_{m=1}^M y_1(m, n) \cos(\hat{\alpha}_2 m + \hat{\beta}_2 m^2 + \hat{\gamma}_2 n + \hat{\delta}_2 n^2) + o(1) \\ \frac{2}{MN} \sum_{n=1}^N \sum_{m=1}^M y_1(m, n) \sin(\hat{\alpha}_2 m + \hat{\beta}_2 m^2 + \hat{\gamma}_2 n + \hat{\delta}_2 n^2) + o(1) \end{bmatrix}.$$

Here,  $y_1(m, n)$  is the data obtained at the second stage after eliminating the effect of the first component from the original data as defined in (18). Using the relationship (31) and following the same procedure as for the consistency of  $\hat{A}_1$ , it can be seen that:

$$\hat{A}_2 \xrightarrow{a.s.} A_2^0 \text{ and } \hat{B}_2 \xrightarrow{a.s.} B_2^0 \text{ as } \min\{M, N\} \rightarrow \infty. \quad (33)$$

It is evident that the result can be easily extended for any  $2 \leq k \leq p$ .

- **Case 2** When  $k = p + 1$ :

$$\begin{bmatrix} \hat{A}_{p+1} \\ \hat{B}_{p+1} \end{bmatrix} = \begin{bmatrix} \frac{2}{MN} \sum_{n=1}^N \sum_{m=1}^M y_p(m, n) \cos(\hat{\alpha}_{p+1} m + \hat{\beta}_{p+1} m^2 + \hat{\gamma}_{p+1} n + \hat{\delta}_{p+1} n^2) + o(1) \\ \frac{2}{MN} \sum_{n=1}^N \sum_{m=1}^M y_p(m, n) \sin(\hat{\alpha}_{p+1} m + \hat{\beta}_{p+1} m^2 + \hat{\gamma}_{p+1} n + \hat{\delta}_{p+1} n^2) + o(1) \end{bmatrix}, \quad (34)$$



where

$$\begin{aligned}
 y_p(m, n) &= y(m, n) - \sum_{j=1}^p \left\{ \hat{A}_j \cos(\hat{\alpha}_j m + \hat{\beta}_j m^2 + \hat{\gamma}_j n + \hat{\delta}_j n^2) \right. \\
 &\quad \left. + \hat{B}_j \sin(\hat{\alpha}_j m + \hat{\beta}_j m^2 + \hat{\gamma}_j n + \hat{\delta}_j n^2) \right\} \\
 &= X(m, n) + o(1), \text{ using (31) and case 1 results.}
 \end{aligned}$$

From here, it is not difficult to see that (34) implies the following result:

$$\hat{A}_{p+1} \xrightarrow{a.s.} 0 \quad \text{and} \quad \hat{B}_{p+1} \xrightarrow{a.s.} 0 \text{ as } \min\{M, N\} \rightarrow \infty.$$

This is obtained using Lemma 2 of Lahiri et al. (2015). It is apparent that the result holds true for any  $k > p$ .

□

**Proof of Theorem 8:** Consider (26) and multiply both sides of the equation with the diagonal matrix,  $\mathbf{D}_1^{-1}$ :

$$(\hat{\xi}_1 - \xi_1^0) \mathbf{D}_1^{-1} = -\mathbf{R}_{1,MN}^{(1)'} (\xi_1^0) \mathbf{D}_1 [\mathbf{D}_1 \mathbf{R}_{1,MN}^{(1)''} (\bar{\xi}_1) \mathbf{D}_1]^{-1}. \tag{35}$$

Computing the elements of the vector  $-\mathbf{R}_{1,MN}^{(1)'} (\xi_1^0) \mathbf{D}_1$  and using definition (28) and the preliminary result (14) (Sect. 3.1), we obtain the following result:

$$-\mathbf{R}_{1,MN}^{(1)'} (\xi_1^0) \mathbf{D}_1 \xrightarrow{d} \mathcal{N}_2(\mathbf{0}, 2\sigma^2 \Sigma_1^{-1}) \text{ as } M \rightarrow \infty. \tag{36}$$

On combining (35), (36) and (30), we have:

$$(\hat{\xi}_1 - \xi_1^0) \mathbf{D}_1^{-1} \xrightarrow{d} \mathcal{N}_2(\mathbf{0}, 2\sigma^2 \Sigma_1)$$

This result can be extended for  $k = 2$  using the relation (31) and following the same argument as above. Similarly, we can continue to extend the result for any  $k \leq p$ . □

## References

Abatzoglou, T. J. (1986). Fast maximum likelihood joint estimation of frequency and frequency rate. *IEEE Transactions on Aerospace and Electronic Systems*, 6, 708–715.

Barbarossa, S. (1995). Analysis of multicomponent LFM signals by a combined Wigner-Hough transform. *IEEE Transactions on Signal Processing*, 43(6), 1511–1515.

Djuric, P. M., & Kay, S. M. (1990). Parameter estimation of chirp signals. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 38(12), 2118–2126.

Djurović, I. (2017). Quasi ML algorithm for 2-D PPS estimation. *Multidimensional Systems and Signal Processing*, 28(2), 371–387.

Djurović, I., & Simeunović, M. (2018). Parameter estimation of 2D polynomial phase signals using NU sampling and 2D CPF. *IET Signal Processing*, 12(9), 1140–1145.

Djurović, I., Wang, P., & Ioana, C. (2010). Parameter estimation of 2-D cubic phase signal using cubic phase function with genetic algorithm. *Signal Processing*, 90(9), 2698–2707.

Francos, J. M., & Friedlander, B. (1998). Two-dimensional polynomial phase signals: Parameter estimation and bounds. *Multidimensional Systems and Signal Processing*, 9(2), 173–205.

Francos, J. M., & Friedlander, B. (1999). Parameter estimation of 2-D random amplitude polynomial-phase signals. *IEEE Transactions on Signal Processing*, 47(7), 1795–1810.

- Friedlander, B., & Francos, J. M. (1995). Estimation of amplitude and phase parameters of multicomponent signals. *IEEE Transactions on Signal Processing*, 43(4), 917–926.
- Friedlander, B., & Francos, J. M. (1996). An estimation algorithm for 2-D polynomial phase signals. *IEEE Transactions on Image Processing*, 5(6), 1084–1087.
- Grover, R., Kundu, D., & Mitra, A. (2018). Approximate least squares estimators of a two-dimensional chirp model and their asymptotic properties. *Journal of Multivariate Analysis*, 168, 211–220.
- Kundu, D., & Nandi, S. (2008). Parameter estimation of chirp signals in presence of stationary noise. *Statistica Sinica*, 18(1), 187–201.
- Lahiri, A. (2013). Estimators of parameters of chirp signals and their properties. PhD thesis, Indian Institute of Technology, Kanpur.
- Lahiri, A., & Kundu, D. (2017). On parameter estimation of two-dimensional polynomial phase signal model. *Statistica Sinica*, 27, 1779–1792.
- Lahiri, A., Kundu, D., & Mitra, A. (2013). Efficient algorithm for estimating the parameters of two dimensional chirp signal. *Sankhya B*, 75(1), 65–89.
- Lahiri, A., Kundu, D., & Mitra, A. (2015). Estimating the parameters of multiple chirp signals. *Journal of Multivariate Analysis*, 139, 189–206.
- Peleg, S., & Porat, B. (1991). Linear FM signal parameter estimation from discrete-time observations. *IEEE Transactions on Aerospace and Electronic Systems*, 27(4), 607–616.
- Prasad, A., Kundu, D., & Mitra, A. (2008). Sequential estimation of the sum of sinusoidal model parameters. *Journal of Statistical Planning and Inference*, 138(5), 1297–1313.
- Richards, F. S. (1961). A method of maximum-likelihood estimation. *Journal of the Royal Statistical Society. Series B*, 23(2), 469–475.
- Saha, S., & Kay, S. M. (2002). Maximum likelihood parameter estimation of superimposed chirps using Monte Carlo importance sampling. *IEEE Transactions on Signal Processing*, 50(2), 224–230.
- Simeunović, M., & Djurović, I. (2016). Parameter estimation of multicomponent 2D polynomial-phase signals using the 2D PHAF-based approach. *IEEE Transactions on Signal Processing*, 64(3), 771–782.
- Simeunović, M., Djurović, I., & Djukanović, S. (2014). A novel refinement technique for 2-D PPS parameter estimation. *Signal Processing*, 94, 251–254.
- Stoica, P., Jakobsson, A., & Li, J. (1997). Cisoid parameter estimation in the colored noise case: asymptotic Cramér-Rao bound, maximum likelihood, and nonlinear least-squares. *IEEE Transactions on Signal Processing*, 45(8), 2048–2059.
- Wang, P., Li, H., Djurovic, I., & Himed, B. (2010). Integrated cubic phase function for linear FM signal analysis. *IEEE Transactions on Aerospace and Electronic Systems*, 46(3), 963–977.
- Wood, J. C., & Barry, D. T. (1994). Radon transformation of time-frequency distributions for analysis of multicomponent signals. *IEEE Transactions on Signal Processing*, 42(11), 3166–3177.
- Wu, C. F. (1981). Asymptotic theory of nonlinear least squares estimation. *The Annals of Statistics*, 9, 501–513.
- Xia, X. G. (2000). Discrete chirp-Fourier transform and its application to chirp rate estimation. *IEEE Transactions on Signal Processing*, 48(11), 3122–3133.
- Zhang, K., Wang, S., & Cao, F. (2008). Product cubic phase function algorithm for estimating the instantaneous frequency rate of multicomponent two-dimensional chirp signals. In *2008 Congress on image and signal processing* (Vol. 5, pp. 498–502).

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Rhythm Grover** received her B.Sc. degree from Delhi University. She obtained her M.Sc. and Ph.D. from Indian Institute of Technology Kanpur. She is currently appointed as a Project Scientist to deliver lectures on Bayesian statistics to a research group in physics department at Indian Institute of Technology Kanpur. Her research interests lie at the intersection of statistics and signal processing.



**Debasis Kundu** received his B.Stat. and M.Stat. degrees from Indian Statistical Institute, his MA in Mathematics from University of Pittsburgh and his Ph.D. in Statistics from Pennsylvania State University. He is a Professor of Statistics in Indian Institute of Technology Kanpur. His primary research interests include statistical signal processing, non-linear regression, survival analysis, generalised exponential distribution and statistical computing.



**Amit Mitra** received his B.Sc. degree in Statistics from University of Calcutta. He obtained his M.Sc. degree and Ph.D. from Indian Institute of Technology Kanpur, both in Statistics. He is a Professor of Statistics in the Indian Institute of Technology Kanpur. His research interests are statistical signal processing, data mining of financial and economic time series and parameter estimation of non-linear time series models.