Robust Parameter Estimation of Two Dimensional Chirp Model

Debasis Kundu *

Abstract

The two dimensional chirp model has received a significant amount of attention in the statistical signal processing literature. Among different estimators the least squares estimators are the most efficient estimators. In presence of additive Gaussian errors the asymptotic variances of the least squares estimators attend the Cramer-Rao lower bound. Even if they are the most efficient estimators, it is well known that the least squares estimators are not robust. In presence of even few outliers the performances of the least squares estimators affect significantly. Some of the standard robust estimators like least absolute deviation estimators and Huber-M estimators are difficult to implement in practice. Moreover, to establish the consistency and asymptotic normality properties of these robust estimators, one needs stronger assumptions than what are needed for the least squares estimators. In this paper we have proposed weighted least squares estimators which can be used quite effectively as robust estimators and as alternative to the least absolute deviation estimators or Huber-M estimators. The implementation of the proposed estimators is quite simple compared to the other existing robust estimators. Moreover, the asymptotic properties of the weighted least squares estimators can be established under the same set of assumptions as the least squares estimators. Extensive simulations have been performed to show the performances of the weighted least squares estimators and one synthetic data set has been analyzed to show how it can be implemented in practice. Finally we have demonstrated how this method can be extended for a more general model also.

Keywords: Sinusoidal model; chirp model; chirp like model; least squares estimators, asymptotic distribution; Cramer-Rao bound; consistent estimators; sequential procedures.

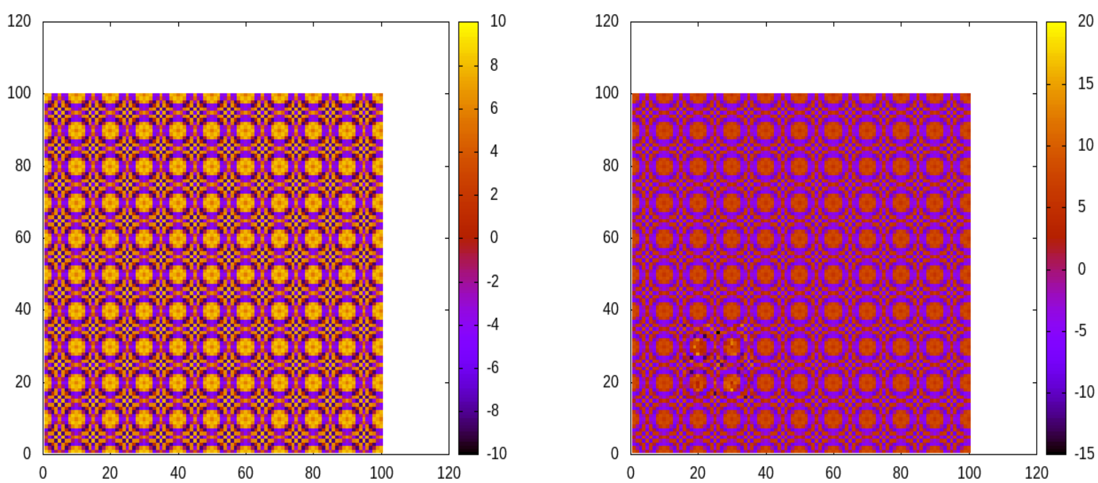
*Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Pin 208016, India, e-mail: kundu@iitk.ac.in. Corresponding author.

1 Introduction

In this paper we consider the following two-dimensional (2-D) chirp model;

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

Here $y(m,n)$ is the observed signal at the point (m,n), A^0, B^0 are the amplitude parameters, α^0, γ^0 are the frequencies, β^0, δ^0 are the frequency rates. The random component $X(m,n)$ accounts for the noise component of the observed signal. It is assumed that $X(m,n)$'s are independent and identically distributed (i.i.d.) random field. The explicit assumptions of $X(m,n)$ will be mentioned later. Based on the observed sample $\{y(m,n); m = 1, \ldots, M; , n = 1, \ldots, N\}$, the problem is to estimate the unknown parameters $\boldsymbol{\theta}^0 = (A^0, B^0, \alpha^0, \beta^0, \gamma^0, \delta^0)$. Typical two-dimensional image plots of the data obtained from the model (1) with and without noise are provided in Figure 1. The problem is to extract the original image as in Figure 1(a) from the noisy image as obtained in Figure 1(b).



(a) (b)

Figure 1: The image plot of the data obtained from the model (1): (a) Without noise (b) With noise.

It can be seen that the model admits a decomposition of two components: the deterministic component and a random component. The 2-D chirp model (1) has received a significant amount of attention in the signal processing literature due to its various applications. The 2-D chirp models find applications in diverse fields like image processing, biomedical spectral analysis, and telecommunications. They are particularly useful for modeling non-homogeneous patterns, including gray-scale and texture images, as well as in applications like magnetic resonance imaging (MRI) and interferometric synthetic aperture radar (INSAR), see for example Francos and Friedlander [8, 9]. They have been applied to analyze non-homogeneous gray-scale texture images. The 2-D chirp model can be used to identify and analyze specific features or objects within an image, see for example Super and Bovik [47], Grover et al. [15], Shukla et al. [44] and see the references cited therein. Applications of these models are not limited to image processing only, but also in studying interferometric synthetic aperture radar (INSAR), where ground elevations and terrain maps can be obtained using INSAR data see for example Francos and Friedlander [9], Friedlander and Francos [11], Wang and Jiang [49] and some of their references. The 2-D chirp model can be used to analyze and process MRI data, particularly in areas where tissues or material variations are present. Similar to MRI, these models can be used to analyze and process optical images, particularly in situations where there are spatially varying refractive indices, see for example Mondal and Achilefum [35]. The above model also can be used to analyze and interpret biomedical spectral data. The 2-D chirp models have been used as a spreading function in digital watermarking which is also helpful in monitoring data, medical safety, fingerprinting and observing content manipulations, see Zhang et al. [51].

Since the inception of the model, a significant amount of work has been done in developing different estimation procedures of the unknown parameters. The model (1) can be seen as a typical 2-D non-linear regression model. It is an extension of the 2-D sinusoidal model, see for example Prasad et al. [40]. Two dimensional Fourier transformation has been used for the estimation of the chirp parameters, see for example Tekavec et al. [48], although, they are not efficient. The least squares estimators (LSEs) are the most efficient estimators. It can be shown that if the errors are Gaussian, then the asymptotic variances of the maximum likelihood estimators (MLEs), which are equivalent to the LSEs, attend the Cramer-Rao lower bound, see for example Lahiri and Kundu [30]. The least squares surface is a highly non-linear surface. Hence, finding the LSEs is a numerically challenging problem. Due to this reason several alternative approaches have been proposed in the literature to estimate the unknown parameters of the model (1). Polynomial phase differencing (PD) operator was introduced by Francos and Friedlander [7] as an extension of the polynomial phase transformation proposed in Pelag and Porat [39]. Friedlander and Francos [11] and Francos and Friedlander [8, 9] utilized PD operator to develop computationally efficient algorithms for estimating similar polynomial phase signals. The cubic phase function (CPF) originally proposed by Oshea [38] was extended by Zhang et al. [52] for the 2-D chirp model. Djuroví et al. [4] proposed genetic algorithm for efficient estimation of the 2-D chirp parameters. Lahiri et al. [31] proposed an efficient algorithm to estimate the parameters of the 2-D chirp model. It converges in finite number of steps and the asymptotic variances of the efficient estimators proposed by Lahiri et al. [31] are same as the LSEs. Grover et al. [14] proposed approximate LSEs which also have the same asymptotic properties as the LSEs. By reducing 2-D problem to 1-D problem Grover et al. [15] proposed a numerically efficient algorithm to estimate the unknown parameters of the 2-D chirp model, which also have the same asymptotic properties as the LSEs, see also Shukla et al. [45] in this respect.

A special case of the model (1) is the 2-D sinusoidal model and it has the following representation;

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

The model (2) was first introduced by Lang and McClellan [33] and it was proposed by Francos et al. [10] as a united texture model, since then it has received a significant amount of attention in the statistical signal processing literature. Rao et al. [42] developed the asymptotic properties of the MLEs under the assumptions of Gaussian errors. Gupta and Kundu [19], Zhang and Manderakar [50], Kundu and Nandi [27] provided the properties of the LSEs under the more general error assumptions. Different efficient estimation techniques using 2-D Prony's method have been proposed by different authors, see for example Barbieri and Barone [1], Hua [20], Rao et al. [41], Cabrera and Bose [2], Chun and Bose [3], Prasad et al. [40], Nandi et al. [37], Grover et al. [16], Kundu and Grover [26] and the references cited therein.

Another special case of the model (1) is the following;

$$y(m,n) = A^0 \cos(\beta^0 m^2 + \delta^0 n^2) + X(m,n); \ m = 1,\ldots,M; \quad n = 1,\ldots,N. \tag{3}$$

The model (3) has been used quite extensively in analyzing Newton's rings. Estimating the parameters of Newton's rings is an important problems in optics. An extensive amount of work has been done in developing different efficient estimation procedures of the unknown parameters of a Newton's rings based on noisy image, see for example Fan and Liu [5], Guo and Li [17], Gorthi and Rastogi [12, 13], Guo et al. [18], Kundu [25] and see the references cited therein.

The main aim of this paper is to provide efficient estimation procedure of the unknown parameters of the model (1), when there are outliers in the data set. The standard robust estimators like least absolute deviation estimators (LADEs) or Huber-M estimators (HMEs) may be used for this purpose. But the implementation of these robust estimators for a non-linear regression model is a numerically challenging problem. Moreover, establishing the asymptotic properties, like strong consistency and asymptotic normality, of these robust estimators require stronger assumptions that what are needed for the LSEs. Further, the performances of these robust estimators are not very satisfactory if there are no outliers in the data set. Recently, the weighted least squares estimators (WLSEs) have been proposed in the literature as an alternative to the well known LADEs or HMEs, see for example Kundu and Grover [26], Kundu [23, 24], Kundu et al. [29] and Mittal et al. [34]. In this paper we develop WLSEs in case of 2-D chirp model (1). The consistency and asymptotic normality properties of the proposed WLSEs have been developed under the same set of assumptions as are needed for the LSEs. Extensive simulations have been performed, and it is observed that when the outliers are present the performances of the WLSEs are at par with the LADEs and HMEs and they are better than the LSEs. Moreover, when there are no outliers, the performances of WLSEs are at par with the LSEs, and they are better than the LADEs and HMEs. One synthetic data set has been analyzed to show how the proposed method can be used in practice. Further we have extended our results to multicomponent 2-D chirp model also.

The rest of the paper is organized as follows. In Section 2 first we briefly mention about the existing robust estimators like LADEs and HMEs, and then provide the WLSEs for polynomial weight function. The asymptotic properties of the WLSEs have been provided in Section 3. Extensive simulation results have been presented in Section 4. It is observed that the performances of the WLSEs depend very much on the weight function, in Section 5, we recommend how to choose the proper weight function and one synthetic data set has been analyzed in Section 6. In Section 7 we have shown how the method can be extended for the multicomponent 2-D chirp model and finally in Section 8 we have concluded the paper and also provided some open problems.

2 Robust Estimators

In this section we first describe two most well known robust estimators namely LADEs and HMEs which can be used for the model (1). We will mention about the merits and demerits of both the methods. Then we will propose the WLSEs as an alternative to both these estimators. From now on we will be using the following notations:

$$\mu_{mn}(\boldsymbol{\theta}) = A\cos(\alpha m + \beta m^2 + \gamma n + \delta n^2) + B\sin(\alpha m + \beta m^2 + \gamma n + \delta n^2), \tag{4}$$

$$= Ac_{mn}(\boldsymbol{\phi}) + Bs_{mn}(\boldsymbol{\phi}), \tag{5}$$

here $\boldsymbol{\phi} = (\alpha, \beta, \gamma, \delta)$, $\boldsymbol{\theta} = (A, B, \boldsymbol{\phi})$, $c_{mn}(\boldsymbol{\phi}) = \cos(\alpha m + \beta m^2 + \gamma n + \delta n^2)$ and $s_{mn}(\boldsymbol{\phi}) = \sin(\alpha m + \beta m^2 + \gamma n + \delta n^2)$

2.1 Least Absolute Deviation Estimators

It is well known that the LSEs are most efficient estimators particularly in presence of additive errors and when the errors are Gaussian. The LSEs of the unknown parameters of the model (1) can be obtained by minimizing

$$Q_S(\boldsymbol{\theta}) = \sum_{m=1}^{M}\sum_{n=1}^{N}\left(y(m,n) - \mu_{mn}(\boldsymbol{\theta})\right)^2, \tag{6}$$

with respect to the unknown parameter vector $\boldsymbol{\theta} = (A, B, \alpha, \beta, \gamma, \delta)$. Here A and B are linear parameters and $\alpha, \beta, \gamma, \delta$ are non-linear parameters. It can be shown by using the separable regression technique of Richards [43] that the LSEs can be obtained by solving a four-dimensional optimization problem. It is known, see for example Lahiri and Kundu [30], that the LSEs are consistent and they are the most efficient estimators. In presence of i.i.d. Gaussian errors the asymptotic variances of the LSEs attend the Cramer-Rao lower bound. Although, the LSEs are the most efficient estimators, they are not robust. Even if there are few outliers present in the data set, they can affect the performances of the LSEs significantly. The main reason is that the LSEs put significant weight on the residuals where outliers are present. The LADEs try to take care that deficiency. The LADEs of the unknown parameters of the model (1) can be obtained by minimizing

$$Q_A(\boldsymbol{\theta}) = \sum_{m=1}^{M}\sum_{n=1}^{N}\left|y(m,n) - \mu_{mn}(\boldsymbol{\theta})\right|, \tag{7}$$

with respect to $\boldsymbol{\theta}$. The LADEs can be obtained by solving a six-dimensional optimization problem. Hence, the implementation of this method is computationally more demanding than the least squares method. It may be mentioned that Lahiri et al. [32] established the asymptotic properties of the LADEs for one dimensional chirp model. To establish these properties it needs stronger error assumptions than the LSEs. No such results are available for two-dimensional chirp model. It is expected that even in 2-D case stronger assumptions will be needed to establish the asymptotic properties of the LADEs.

2.2 Huber-M Estimator

The Huber M-estimator (HME) is a robust regression method designed to minimize the impact of outliers in regression models. It is a generalization of the ordinary least squares estimator, but it uses a different error function that gives less weight to large errors, making it less susceptible to outliers. The HMEs of of the unknown parameters of the model (1) can be obtained by minimizing

$$Q_H(\boldsymbol{\theta}) = \sum_{m=1}^{M}\sum_{n=1}^{N} \rho\left(y(m,n) - \mu_{mn}(\boldsymbol{\theta})\right), \tag{8}$$

with respect to the unknown parameter vector $\boldsymbol{\theta}$. The classical Huber estimator is defined with the following Huber's loss function

$$\rho(\delta) = \begin{cases} \frac{\delta^2}{2} & \text{if} \quad |\delta| \leq \tau \\ \tau|\delta| - \frac{\tau^2}{2}, & \text{if} \quad |\delta| > \tau, \end{cases} \tag{9}$$

where $\tau > 0$ is a user defined tuning parameter. The tuning parameter τ is typically set equal to 1.345, see for example Zhang et al. [53]. To compute the HMEs, one needs to solve a six-dimensional optimization problem. No theoretical properties are available about the asymptotic properties of the HMEs for two-dimensional chirp model. It may be mentioned that both the LADEs and HMEs perform better than the LSEs in presence of outliers, but if there are no outliers, the LSEs perform better than both these estimators.

2.3 Weighted Least Squares Estimators

Recently, the weighted least squares estimators (WLSEs) have been introduced in the literature as an alternative to the LADEs and HMEs. The method has been applied for different signal processing models, see for example Kundu et al. [28, 29], Kundu and Grover [26], Kundu [23, 24], Mittal et al. [34]. The basic idea about the WLSEs is same as the LADEs and HMEs, i.e. it tries to put less weight to large errors compared to the LSEs. Hence, they become more robust compared to the LSEs in presence of outliers. We need to introduce the following to define the WLSEs.

Suppose the weight function $w(s,t)$ is a polynomial of degree k, defined for $0 \leq s, t \leq 1$, and it can be written as follows:

$$w(s,t) = c_{00} + c_{10}s + c_{01}t + c_{20}s^2 + c_{11}st + c_{02}t^2 + \ldots + c_{k0}s^k + \ldots + c_{0k}t^k. \tag{10}$$

The coefficients of the polynomial (10) are such that $w(s,t) \geq \eta > 0$, for $0 \leq s, t \leq 1$, for some η. The weighted residual sums of squares can be defined as follows;

$$Q_W(\boldsymbol{\theta}) = \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} w\left(\frac{m}{M}, \frac{n}{N}\right) \left(y(m,n) - \mu_{mn}(\boldsymbol{\theta})\right)^2. \tag{11}$$

The WLSEs of $\boldsymbol{\theta}$ can be obtained by minimizing (11) with respect to the unknown parameter vector $\boldsymbol{\theta}$. We use the separable regression technique of Richards [43] to compute the WLSEs of the unknown parameters. It reduces the computation in the sense, instead of solving six-dimensional optimization problem, they can be obtained by solving a four-dimensional optimization problem. For fixed $(\alpha, \beta, \gamma, \delta)$, let us denote the WLSEs of A and B as $\widehat{A}(\alpha, \beta, \gamma, \delta)$ and $\widehat{B}(\alpha, \beta, \gamma, \delta)$, respectively. Hence,

$$\begin{aligned}
\widehat{A}(\alpha, \beta, \gamma, \delta) &= \frac{1}{a_{11}a_{22} - (a_{12})^2} \left[a_{22}d_1 - a_{12}d_2\right] \\
\widehat{B}(\alpha, \beta, \gamma, \delta) &= \frac{1}{a_{11}a_{22} - (a_{12})^2} \left[a_{11}d_2 - a_{12}d_1\right],
\end{aligned}$$

where

$$
\begin{aligned}
a_{11} &= \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right)\cos^2(\alpha m+\beta m^2+\gamma n+\delta n^2) \\
a_{22} &= \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right)\sin^2(\alpha m+\beta m^2+\gamma n+\delta n^2) \\
a_{12}=a_{21} &= \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right)\cos(\alpha m+\beta m^2+\gamma n+\delta n^2)\sin(\alpha m+\beta m^2+\gamma n+\delta n^2),
\end{aligned}
$$

$$
\begin{aligned}
d_1 &= \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right) y(m,n)\cos(\alpha m+\beta m^2+\gamma n+\delta n^2) \\
d_2 &= \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right) y(m,n)\sin(\alpha m+\beta m^2+\gamma n+\delta n^2).
\end{aligned}
$$

Therefore, the WLSEs of the frequency and frequency rate parameters namely α, β, γ and δ can be obtained by minimizing

$$Q_W(\widehat{A}(\alpha,\beta,\gamma,\delta),\widehat{B}(\alpha,\beta,\gamma,\delta),\alpha,\beta,\gamma,\delta),$$

with respect to $\alpha,\beta,\gamma,\delta$. Hence, it becomes a four-dimensional optimization problem. Once, $\widehat{\alpha},\widehat{\beta},\widehat{\gamma},\widehat{\delta}$ are obtained as the WLSEs of $\alpha,\beta,\gamma,\delta$, respectively, then the WLSEs of A and B can be obtained as $\widehat{A}(\widehat{\alpha},\widehat{\beta},\widehat{\gamma},\widehat{\delta})$ and $\widehat{B}(\widehat{\alpha},\widehat{\beta},\widehat{\gamma},\widehat{\delta})$, respectively. Using Proposition 1 of Lahiri and Kundu [30], it can be seen that

$$\lim_{M,N\to\infty} a_{11} = \lim_{M,N\to\infty} a_{11} = \frac{1}{2}\int_0^1\int_0^1 w(s,t)dsdt \quad \text{and} \quad \lim_{M,N\to\infty} a_{12} = \lim_{M,N\to\infty} a_{21} = 0.$$

Hence, for large M and N, $\widehat{A}(\alpha,\beta,\gamma,\delta)$ and $\widehat{B}(\alpha,\beta,\gamma,\delta)$ can be written as

$$\widehat{A}(\alpha,\beta,\gamma,\delta) = \frac{2d_1}{\int_0^1\int_0^1 w(s,t)dsdt} \quad \text{and} \quad \widehat{B}(\alpha,\beta,\gamma,\delta) = \frac{2d_2}{\int_0^1\int_0^1 w(s,t)dsdt}. \tag{12}$$

Now before going to provide the theoretical properties of the WLSEs, first we would like to provide some intuitive reason why the WLSEs can be used as robust estimators compared to the LSEs. In the computation of the LSEs, the residual squares have been given equal weights at each (m, n). Hence, if there are outliers in the data set at certain (m, n)'s it is expected that the residual squares will be high at those points, and they influence significantly in the evaluation of the LSEs. Due to this reason, the LSEs are not robust in presence of outliers. On the other hand, in case of WLSEs, if the weight function is chosen in such a manner that it puts less weight at those points (m, n)'s where outliers are present, then it is expected that it will produce more robust estimators compared to the LSEs. Hence, one key issue is to choose the weight function properly. It is definitely an important problem for practical implementation of the proposed method. We will discuss this in details in Sections 5 and 6.

3 Asymptotic Properties

In this section we provide the consistency and asymptotic normality properties of the WLSEs. We need the following assumptions to establish the asymptotic properties of the WLSEs.

Assumption 1: It is assumed that there exists a M, such that $0 < |A^0|, |B^0| < M$. Moreover $0 < \alpha^0, \beta^0, \gamma^0, \delta^0 < \pi$. Here, $\boldsymbol{\Theta} = \{(A, B, \alpha, \beta, \gamma, \delta) : |A| < M, |B| < M, 0 < \alpha, \beta, \gamma, \delta < \pi\}$ is the parameter space.

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

Assumption 3: (i) $\inf_{0 \le s,t \le 1} w(s, t) \ge \eta > 0$, and (ii) $\sup_{0 \le s,t \le 1} w(s, t) \le 1$.

Theorem 1: Under Assumptions 1,2 and 3, the WLSEs of $A^0, B^0, \alpha^0, \beta^0, \gamma^0, \delta^0$ are strongly consistent.

Proof: See in the Appendix. ∎

Now we would like to provide the asymptotic distribution of the WLSEs. We need the following notations for that purpose. For $p, q = 0, 1, 2, \ldots$, let us define

$$u_{pq} = \int_0^1 \int_0^1 s^p t^q w(s,t) ds dt \quad \text{and} \quad v_{pq} = \int_0^1 \int_0^1 s^p t^q w^2(s,t) ds dt. \tag{13}$$

The 6×6 diagonal matrix $\boldsymbol{D}$ is as follows;

$$\boldsymbol{D} = \text{diag}\left\{M^{1/2}N^{1/2}.M^{1/2}N^{1/2}, M^{3/2}N^{1/2}, M^{5/2}N^{1/2}, M^{1/2}N^{3/2}, M^{1/2}N^{5/2}\right\}. \tag{14}$$

The 6×6 matrix $\boldsymbol{\Sigma}$ and $\boldsymbol{\Gamma}$ are as

$$\boldsymbol{\Sigma} = \begin{bmatrix} v_{00} & 0 & B^0 v_{10} & B^0 v_{20} & B^0 v_{01} & B^0 v_{02} \\ 0 & v_{00} & -A^0 v_{10} & -A^0 v_{20} & -A^0 v_{01} & -A^0 v_{02} \\ B^0 v_{10} & -A^0 v_{10} & (A^{0^2} + B^{0^2}) v_{20} & (A^{0^2} + B^{0^2}) v_{30} & (A^{0^2} + B^{0^2}) v_{11} & (A^{0^2} + B^{0^2}) v_{12} \\ B^0 v_{20} & -A^0 v_{20} & (A^{0^2} + B^{0^2}) v_{30} & (A^{0^2} + B^{0^2}) v_{40} & (A^{0^2} + B^{0^2}) v_{21} & (A^{0^2} + B^{0^2}) v_{22} \\ B^0 v_{01} & -A^0 v_{01} & (A^{0^2} + B^{0^2}) v_{11} & (A^{0^2} + B^{0^2}) v_{12} & (A^{0^2} + B^{0^2}) v_{02} & (A^{0^2} + B^{0^2}) v_{03} \\ B^0 v_{02} & -A^0 v_{02} & (A^{0^2} + B^{0^2}) v_{12} & (A^{0^2} + B^{0^2}) v_{22} & (A^{0^2} + B^{0^2}) v_{03} & (A^{0^2} + B^{0^2}) v_{04} \end{bmatrix}$$

and

$$\boldsymbol{\Gamma} = \begin{bmatrix} u_{00} & 0 & B^0 u_{10} & B^0 u_{20} & B^0 u_{01} & B^0 u_{02} \\ 0 & u_{00} & -A^0 u_{10} & -A^0 u_{20} & -A^0 u_{01} & -A^0 u_{02} \\ B^0 u_{10} & -A^0 u_{10} & (A^{0^2} + B^{0^2}) u_{20} & (A^{0^2} + B^{0^2}) u_{30} & (A^{0^2} + B^{0^2}) u_{11} & (A^{0^2} + B^{0^2}) u_{12} \\ B^0 u_{20} & -A^0 u_{20} & (A^{0^2} + B^{0^2}) u_{30} & (A^{0^2} + B^{0^2}) u_{40} & (A^{0^2} + B^{0^2}) u_{21} & (A^{0^2} + B^{0^2}) u_{22} \\ B^0 u_{01} & -A^0 u_{01} & (A^{0^2} + B^{0^2}) u_{11} & (A^{0^2} + B^{0^2}) u_{12} & (A^{0^2} + B^{0^2}) u_{02} & (A^{0^2} + B^{0^2}) u_{03} \\ B^0 u_{02} & -A^0 u_{02} & (A^{0^2} + B^{0^2}) u_{12} & (A^{0^2} + B^{0^2}) u_{22} & (A^{0^2} + B^{0^2}) u_{03} & (A^{0^2} + B^{0^2}) u_{04} \end{bmatrix},$$

respectively.

Theorem 2: Under Assumptions 1,2 and 3,

$$\boldsymbol{D}\left(\widehat{A} - A^0, \widehat{B} - B^0, \widehat{\alpha} - \alpha^0, \widehat{\beta} - \beta^0, \widehat{\gamma} - \gamma^0, \widehat{\delta} - \delta^0\right)^\top \xrightarrow{d} N_6\left(\mathbf{0}, 2\sigma^2 \boldsymbol{\Gamma}^{-1} \boldsymbol{\Sigma} \boldsymbol{\Gamma}^{-1}\right).$$

Here, $\xrightarrow{d}$ means converges in distribution and $N_6(\mathbf{0}, \boldsymbol{C})$, means 6-variate normal distribution with the mean vector $\mathbf{0}$ and the dispersion matrix $\boldsymbol{C}$.

Proof: See in the Appendix.

Comments: It may be mentioned that although we have proved Theorems 1 and 2 for 2-D polynomial $w(s,t)$ as defined in (10) which satisfies Assumption 3, the results can be extended for any continuous function $w(s,t)$, which satisfies Assumption 3. The results mainly follows from the fact that for any continuous function $w(s,t)$ and for any $\epsilon > 0$, there exists a 2-D polynomial $p_\epsilon(s,t)$, such that $|w(s,t) - p_\epsilon(s,t)| \leq \epsilon$, for all $0 \leq s,t \leq 1$. The details are avoided.

4 Simulation Results

In this Section we have presented some simulation results to show how the proposed methods perform in presence of outliers for different sample sizes and for different error variances. The error random variables $X(m,n)$ are assumed to be i.i.d. Gaussian with mean zero and variance σ_0^2. It is assumed that the outliers are present in the middle portion of the data set and they are also Gaussian with mean 0 and variance σ_1^2. The following models have been considered:

Model 1: $A = B = 1, \alpha = 1.25\pi = 3.92699, \beta = 0.75\pi = 2.35619,$
$\gamma = 1.75\pi = 5.49779, \delta = 0.35\pi = 1.09956, \sigma_0^2 = 1.0, \sigma_1^2 = 6.25.$

Model 2: $A = B = 5, \alpha = 1.25\pi = 3.92699, \beta = 0.75\pi = 2.35619,$
$\gamma = 1.75\pi = 5.49779, \delta = 0.35\pi = 1.09956, \sigma_0^2 = 1.0, \sigma_1^2 = 6.25.$

For both the models we have kept all the parameters except the amplitude parameters remain same. We have considered different sample sizes and different proportion of outliers. We have computed the LSEs which can be obtained by minimizing

$$Q_{LSE}(\boldsymbol{\theta}) = \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}(y(m,n) - \mu_{mn}(\boldsymbol{\theta}))^2.$$

The LADEs, which have been obtained by minimizing

$$Q_{LAD}(\boldsymbol{\theta}) = \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\left|y(m,n) - \mu_{mn}(\boldsymbol{\theta})\right|,$$

and also HMEs and they have been obtained by minimizing

$$Q_{HME}(\boldsymbol{\theta}) = \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\rho\left(y(m,n) - \mu_{mn}(\boldsymbol{\theta})\right),$$

where

$$\rho(a) = \begin{cases} \frac{a^2}{2} & \text{if} \quad |a| \le \delta \\ \delta\left(|a| - \frac{\delta}{2}\right) & \text{if} \quad |a| > \delta. \end{cases}$$

Since the outliers are present in the middle portion of the data set, we have considered the two different weight functions which put less weight in the middle portion of the data set and obtain the corresponding WLSEs. The two different weight functions are as follows:

Weight Function 1 (W1) : $\quad w(s,t) = 0.5 + (s-0.5)^2 + (t-0.5)^2; \quad 0 \le s, t \le 1,$

Weight Function 2 (W2) : $\quad w(s,t) = 0.5 + |s-0.5| + |t-0.5|; \quad 0 \le s, t \le 1.$

Here, W1 is a polynomial loss function whereas W2 is a general continuous loss function. Both the loss functions put less weight in the middle portion of the data set. For each data set we have computed LSE, two different WLSEs, LADEs and HMEs and we report the average estimates and the square root of the mean squared errors (SMSE) of all the six estimates over 1000 replications. The results are reported in Table 1 (no outliers), Table 2 (10% outliers) and Table 3 (20% outliers) for Model 1 and the corresponding results for Model 2 are reported in Tables 4, 5 and 6, respectively.

Some of the points are quite clear from these experimental results. In all these cases as sample size increases the biases and the SMSEs decrease. It verifies the consistency properties of all the estimators. In Tables 1 and 4 it is observed that the SMSEs of the LSEs are smaller than the other estimators. It shows that the performance of the LSEs are the best when there are no outliers in the data set. It is also observed in Tables 1 and 4 that the performances of the WLSEs are better than the LADEs and HMEs in terms of lower SMSE. From Tables 2, 3,5 and 6 it is observed that the performances of the LADEs, HMEs and WLSEs are better than the LSEs. The performances of the two WLSEs are at par with the LADEs and HMEs and many times the WLSEs perform better than the LADEs or HMEs. Since, we know about the theoretical properties of the WLSEs and computationally they are

n	PAR	LSE	LADE	HME	W1E	W2E
25	A	0.9996	1.0002	0.9999	1.0005	0.9996
		(8.345E-3)	(1.131E-2)	(1.190E-2)	(1.109E-2)	(1.026E-2)
	B	0.9983	1.0008	1.0002	1.0019	0.9992
		(3.530E-2)	(4.391E-2)	(4.631E-2)	(3.607E-2)	(4.081E-2)
	α	3.9272	3.9273	3.9268	3.9273	3.9268
		(1.489E-2)	1.857E-2	(1.624E-2)	(1.556E-2)	(1.511E-2)
	β	2.3562	2.3562	2.3562	2.3562	2.3562
		(6.165E-4)	(7.697E-4)	(6.743E-4)	(6.376E-4)	(6.137E-4)
	γ	5.4979	5.4978	5.4980	5.4973	5.4985
		(1.577E-2)	(1.989E-2)	(1.678E-2)	(1.615E-2)	(1.614E-2)
	δ	1.0996	1.0996	1.0995	1.0996	1.0995
		(6.453E-4)	(8.230E-4)	(6.814E-4)	(6.662E-4)	(6.649E-4)
50	A	0.9999	0.9999	0.9999	1.0000	0.9999
		(9.190E-4)	(1.321E-3)	(1.152E-3)	(8.378E-4)	(8.755E-4)
	B	0.9997	0.9997	0.9997	0.9998	0.9996
		(3.420E-3)	(5.052E-3)	(4.426E-3)	(3.094E-3)	(3.221E-3)
	α	3.9269	3.9270	3.9270	3.9270	3.9269
		(2.175E-3)	(2.672E-3)	(2.452E-3)	(2.054E-3)	(2.200E-3)
	β	2.3562	2.3562	2.3562	2.3562	2.3562
		(5.100E-5)	(6.551E-5)	(5.776E-5)	(4.984E-5)	(5.172E-5)
	γ	5.4978	5.4978	5.4978	5.4977	5.5978
		(2.378E-3)	(3.014E-3)	(2.690E-3)	(2.376E-3)	(2.322E-3)
	δ	1.0996	1.0996	1.0996	1.0996	1.0996
		(5.604E-5)	(7.090E-5)	(6.115E-5)	(5.781E-5)	(5.618E-5)
75	A	1.0000	1.0000	1.0000	1.0000	1.0000
		(2.171E-4)	(2.745E-4)	(2.575E-4)	(1.874E-4)	(1.755E-4)
	B	0.9999	0.9999	0.9999	0.9998	0.9999
		(7.668E-4)	(9.445E-4)	(9.199E-4)	(7.775E-4)	(6.821E-4)
	α	3.9270	3.9270	3.9270	3.9269	3.9270
		(7.238E-4)	(8.804E-4)	(7.533E-4)	(6.607E-4)	(7.130E-4)
	β	2.3562	2.3562	2.3562	2.3562	2.3562
		(1.239E-5)	(1.497E-5)	(1.326E-5)	(1.197E-5)	(1.212E-5)
	γ	5.4978	5.4978	5.4977	5.4977	5.4978
		(8.984E-4)	(1.064E-3)	(9.065E-4)	(9.022E-4)	(9.052E-4)
	δ	1.0996	1.0996	1.0996	1.0995	1.0996
		(1.468E-5)	(1.748E-5)	(1.527E-5)	(1.502E-5)	(1.511E-5)

Table 1: The average estimates and the square root of the MSEs of the different estimates when there are no outliers. The results are for Model 1.

n	PAR	LSE	LADE	HME	W1E	W2E
25	A	0.9998	0.9999	1.0003	1.0006	1.0000
		(1.241E-2)	(1.100E-2)	(1.148E-2)	(1.146E-2)	(1.066E-2)
	B	0.9994	0.9990	1.0008	1.0016	1.0005
		(4.553E-2)	(4.391E-2)	(4.546E-2)	(4.392E-2)	(4.343E-2)
	α	3.9278	3.9277	3.9268	3.9272	3.9265
		(1.988E-2)	1.890E-2	(1.696E-2)	(1.594E-2)	(1.514E-2)
	β	2.3562	2.3562	2.3562	2.3562	2.3562
		(7.841E-4)	(7.795E-4)	(7.059E-4)	(6.578E-4)	(6.186E-4)
	γ	5.4978	5.4976	5.4980	5.4974	5.4988
		(2.239E-2)	(2.086E-2)	(1.739E-2)	(1.655E-2)	(1.681E-2)
	δ	1.0996	1.0996	1.0995	1.0996	1.0995
		(8.750E-4)	(8.606E-4)	(7.092E-4)	(6.895E-4)	(6.962E-4)
50	A	1.0000	0.9999	0.9999	1.0000	0.9999
		(1.246E-3)	(1.192E-3)	(1.362E-3)	(1.120E-3)	(1.198E-3)
	B	0.9999	0.9997	0.9996	0.9999	0.9997
		(4.356E-3)	(4.266E-3)	(5.133E-3)	(4.005E-3)	(4.172E-3)
	α	3.9269	3.9270	3.9270	3.9269	3.9271
		(2.978E-3)	(2.886E-3)	(2.556E-3)	(2.402E-3)	(2.432E-3)
	β	2.3562	2.3562	2.3562	2.3562	2.3562
		(6.948E-5)	(6.931E-5)	(6.006E-5)	(5.578E-5)	(5.650E-5)
	γ	5.4978	5.4978	5.4978	5.4978	5.4977
		(3.230E-3)	(3.174E-3)	(2.786E-3)	(2.756E-3)	(2.653E-3)
	δ	1.0996	1.0996	1.0996	1.0996	1.0996
		(7.551E-5)	(7.408E-5)	(6.468E-5)	(6.425E-5)	(6.302E-5)
75	A	1.0000	1.0000	1.0000	1.0000	1.0000
		(2.436E-4)	(1.947E-4)	(2.336E-4)	(2.218E-4)	(2.213E-4)
	B	0.9999	0.9999	0.9999	0.9998	0.9999
		(8.413E-4)	(7.359E-4)	(8.166E-4)	(7.975E-4)	(7.987E-4)
	α	3.9269	3.9269	3.9270	3.9269	3.9270
		(8.940E-4)	(8.767E-4)	(8.117E-4)	(8.045E-4)	(7.981E-4)
	β	2.3562	2.3562	2.3562	2.3562	2.3562
		(1.558E-5)	(1.514E-5)	(1.366E-5)	(1.339E-5)	(1.348E-5)
	γ	5.4978	5.4978	5.4977	5.4977	5.4978
		(1.168E-3)	(1.029E-3)	(1.000E-3)	(9.932E-4)	(9.963E-4)
	δ	1.0996	1.0996	1.0996	1.0996	1.0996
		(1.768E-5)	(1.695E-5)	(1.630E-5)	(1.604E-5)	(1.575E-5)

Table 2: The average estimates and the square root of the MSEs of the different estimates when there are 10% outliers. The results are for Model 1.

n	PAR	LSE	LADE	HME	W1E	W2E
25	A	0.9996	1.0007	1.0000	1.0001	1.0000
		(1.395E-2)	(1.194E-2)	(1.137E-2)	(1.110E-2)	(1.124E-2)
	B	0.9984	1.0022	1.0007	1.0005	0.9998
		(5.476E-2)	(4.419E-2)	(4.568E-2)	(4.386E-2)	(4.524E-2)
	α	3.9277	3.9277	3.9266	3.9277	3.9268
		(2.651E-2)	1.893E-2	(1.703E-2)	(1.639E-2)	(1.586E-2)
	β	2.3562	2.3562	2.3562	2.3562	2.3562
		(8.546E-3)	(7.835E-3)	(7.256E-4)	(6.750E-4)	(6.514E-4)
	γ	5.4976	5.4979	5.4979	5.4969	5.4988
		(2.451E-2)	(2.088E-2)	(1.731E-2)	(1.730E-2)	(1.729E-2)
	δ	1.0996	1.0996	1.0995	1.0996	1.0995
		(9.138E-4)	(8.634E-4)	(7.108E-4)	(7.220E-4)	(7.168E-4)
50	A	1.0000	0.9999	0.9999	1.0000	1.0000
		(1.421E-3)	(1.235E-3)	(1.440E-3)	(1.244E-3)	(1.211E-3)
	B	0.9998	0.9997	0.9995	0.9999	0.9999
		(5.619E-3)	(4.701E-3)	(5.491E-3)	(4.561E-3)	(4.347E-3)
	α	3.9269	3.9269	3.9270	3.9269	3.9269
		(3.887E-3)	(2.723E-3)	(2.645E-3)	(2.503E-3)	(2.480E-3)
	β	2.3562	2.3562	2.3562	2.3562	2.3562
		(7.751E-5)	(6.696E-5)	(6.116E-5)	(5.576E-5)	(5.856E-5)
	γ	5.4978	5.4979	5.4978	5.4977	5.4978
		(4.336E-3)	(3.133E-3)	(2.925E-3)	(2.768E-3)	(2.743E-3)
	δ	1.0996	1.0996	1.0996	1.0996	1.0996
		(8.761E-5)	(7.307E-5)	(6.708E-5)	(6.534E-5)	(6.381E-5)
75	A	1.0000	1.0000	1.0000	1.0000	1.0000
		(2.926E-4)	(2.696E-4)	(2.428E-4)	(2.413E-4)	(2.629E-4)
	B	0.9999	0.9999	0.9999	0.9999	0.9999
		(1.001E-3)	(9.389E-4)	(8.178E-4)	(8.222E-4)	(9.087E-4)
	α	3.9269	3.9270	3.9270	3.9269	3.9270
		(9.913E-4)	(8.695E-4)	(8.331E-4)	(8.205E-4)	(8.163E-4)
	β	2.3562	2.3562	2.3562	2.3562	2.3562
		(1.765E-5)	(1.546E-5)	(1.478E-5)	(1.360E-5)	(1.341E-5)
	γ	5.4978	5.4977	5.4978	5.4978	5.4978
		(2.341E-3)	(1.052E-3)	(1.112E-3)	(1.042E-3)	(9.957E-4)
	δ	1.0996	1.0996	1.0996	1.0996	1.0996
		(2.651E-5)	(1.752E-5)	(1.701E-5)	(1.681E-5)	(1.574E-5)

Table 3: The average estimates and the square root of the MSEs of the different estimates when there are 20% outliers. The results are for Model 1.

n	PAR	LSE	LADE	HME	W1E	W2E
25	A	5.0000	4.9991	4.9991	4.9996	4.9999
		(8.833E-3)	(1.247E-2)	(1.084E-2)	(9.349E-3)	(9.338E-2)
	B	4.9998	4.9977	4.9977	4.9987	4.9996
		(2.632E-2)	(3.217E-2)	(2.868E-2)	(2.712E-2)	(2.754E-2)
	α	3.9271	3.9269	3.9270	3.9270	3.9270
		(2.781E-3)	3.653E-3	(3.124E-3)	(2.778E-3)	(2.588E-3)
	β	2.3562	2.3562	2.3562	2.3562	2.3562
		(1.160E-4)	(1.486E-4)	(1.286E-4)	(1.152E-4)	(1.075E-4)
	γ	5.4979	5.4978	5.4978	5.4977	5.4979
		(2.8448-3)	(3.733E-3)	(2.952E-3)	(2.824E-3)	(2.808E-3)
	δ	1.0996	1.0996	1.0996	1.0996	1.0996
		(1.180E-4)	(1.523E-4)	(1.233E-4)	(1.186E-4)	(1.186E-4)
50	A	4.9997	4.9999	4.9998	4.9997	4.9996
		(1.079E-3)	(1.433E-3)	(1.164E-3)	(1.113E-3)	(1.153E-3)
	B	4.9992	4.9994	4.9993	4.9992	4.9992
		(2.893E-3)	(3.510E-3)	(3.675E-3)	(3.176E-3)	(3.108E-3)
	α	3.9269	3.9270	3.9269	3.9270	3.9270
		(3.991E-4)	(5.788E-4)	(4.833E-4)	(4.106E-4)	(4.108E-4)
	β	2.3562	2.3562	2.3562	2.3562	2.3562
		(9.516E-6)	(1.370E-5)	(1.148E-5)	(9.624E-6)	(9.531E-6)
	γ	5.4978	5.4978	5.4978	5.4978	5.5978
		(4.781E-4)	(6.088E-4)	(5.313E-4)	(4.597E-4)	(4.712E-4)
	δ	1.0996	1.0996	1.0996	1.0996	1.0996
		(1.119E-5)	(1.407E-5)	(1.230E-5)	(1.106E-5)	(1.114E-5)
75	A	4.9998	4.9998	4.9998	4.9998	4.9998
		(4.164E-4)	(6.014E-4)	(4.451E-4)	(3.824E-4)	(3.914E-4)
	B	4.9995	4.9993	4.9993	4.9994	4.9995
		(1.164E-4)	(1.519E-3)	(1.361E-3)	(1.054E-3)	(1.111E-3)
	α	3.9270	3.9270	3.9270	3.9270	3.9270
		(1.711E-4)	(2.116E-4)	(1.792E-4)	(1.412E-4)	(1.486E-4)
	β	2.3562	2.3562	2.3562	2.3562	2.3562
		(2.766E-6)	(3.343E-6)	(2.814E-6)	(2.504E-6)	(2.515E-6)
	γ	5.4978	5.4978	5.4978	5.4977	5.4978
		(2.025E-4)	(2.568E-4)	(2.236E-4)	(1.813E-4)	(1.942E-4)
	δ	1.0996	1.0996	1.0996	1.0995	1.0996
		(3.264E-6)	(3.994E-6)	(3.524E-6)	(3.005E-6)	(3.171E-6)

Table 4: The average estimates and the square root of the MSEs of the different estimates when there are no outliers. The results are for Model 2.

n	PAR	LSE	LADE	HME	W1E	W2E
25	A	4.9998	4.9997	4.9998	4.9995	4.9998
		(1.456E-2)	(1.162E-2)	(1.078E-2)	(9.215E-3)	(9.117E-3)
	B	4.9992	4.9992	4.9994	4.9984	4.9993
		(2.808E-2)	(2.741E-2)	(2.751E-2)	(2.709E-2)	(2.707E-2)
	α	3.9270	3.9269	3.9269	3.9272	3.9269
		(3.758E-3)	(3.604E-3)	(2.987E-3)	(2.723E-3)	(2.671E-3)
	β	2.3562	2.3562	2.3562	2.3562	2.3562
		(1.648E-4)	(1.468E-4)	(1.233E-4)	(1.135E-4)	(1.102E-4)
	γ	5.4978	5.4979	5.4978	5.4977	5.4980
		(3.927E-3)	(3.795E-3)	(3.027E-3)	(2.787E-3)	(2.880E-3)
	δ	1.0996	1.0996	1.0995	1.0996	1.0996
		(1.713E-4)	(1.541E-4)	(1.249E-4)	(1.173E-4)	(1.213E-4)
50	A	4.9997	4.9999	4.9998	4.9997	4.9997
		(1.179E-3)	(9.321E-4)	(8.825E-4)	(1.001E-3)	(1.126E-3)
	B	4.9992	4.9996	4.9995	4.9992	4.9994
		(3.830E-3)	(3.327E-3)	(3.129E-3)	(2.857E-3)	(3.023E-3)
	α	3.9269	3.9270	3.9269	3.9269	3.9270
		(6.335E-4)	(5.560E-4)	(5.127E-4)	(4.142E-4)	(4.061E-4)
	β	2.3562	2.3562	2.3562	2.3562	2.3562
		(1.671E-5)	(1.360E-5)	(1.187E-5)	(9.738E-6)	(9.534E-6)
	γ	5.4978	5.4978	5.4978	5.4978	5.4978
		(6.750E-4)	(5.878E-4)	(5.309E-4)	(4.753E-4)	(4.727E-4)
	δ	1.0996	1.0996	1.0996	1.0996	1.0996
		(1.679E-5)	(1.407E-5)	(1.210E-5)	(1.125E-5)	(1.125E-5)
75	A	4.9998	4.9997	4.9998	4.9998	4.9998
		(7.376E-4)	(6.907E-4)	(4.360E-4)	(3.994E-4)	(3.909E-4)
	B	4.9995	4.9994	4.9993	4.9995	4.9994
		(1.920E-3)	(1.747E-3)	(1.331E-3)	(1.097E-3)	(1.109E-3)
	α	3.9270	3.9270	3.9270	3.9270	3.9270
		(2.751E-4)	(2.204E-4)	(1.790E-4)	(1.463E-4)	(1.469E-4)
	β	2.3562	2.3562	2.3562	2.3562	2.3562
		(4.804E-6)	(3.475E-6)	(2.813E-6)	(2.599E-6)	(2.576E-6)
	γ	5.4978	5.4978	5.4978	5.4978	5.4978
		(2.672E-4)	(2.477E-4)	(2.158E-4)	(1.852E-4)	(1.983E-4)
	δ	1.0996	1.0996	1.0996	1.0996	1.0996
		(4.138E-6)	(3.917E-6)	(3.262E-6)	(3.074E-6)	(3.216E-6)

Table 5: The average estimates and the square root of the MSEs of the different estimates when there are 10% outliers. The results are for Model 2.

n	PAR	LSE	LADE	HME	W1E	W2E
25	A	4.9996	4.9993	4.9993	4.9996	5.0000
		(1.617E-2)	(1.173E-2)	(1.075E-2)	(9.719E-3)	(1.013E-2)
	B	4.9987	4.9981	4.9980	4.9985	5.0001
		(4.838E-2)	(3.060E-2)	(2.853E-2)	(2.848E-2)	(2.999E-2)
	α	3.9271	3.9268	3.9269	3.9271	3.9269
		(4.866E-3)	3.519E-3	(3.023E-3)	(2.769E-3)	(2.781E-3)
	β	2.3562	2.3562	2.3562	2.3562	2.3562
		(2.197E-4)	(1.452E-4)	(1.254E-4)	(1.155E-4)	(1.141E-4)
	γ	5.4979	5.4979	5.4979	5.4977	5.4980
		(4.985E-3)	(3.617E-3)	(3.002E-3)	(2.916E-3)	(3.032E-3)
	δ	1.0996	1.0996	1.0996	1.0996	1.0995
		(2.241E-4)	(1.494E-4)	(1.248E-4)	(1.233E-4)	(1.272E-4)
50	A	4.9997	4.9998	4.9998	4.9997	4.9998
		(1.238E-3)	(1.013E-3)	(9.499E-4)	(1.101E-3)	(1.184E-3)
	B	4.9992	4.9994	4.9994	4.9993	4.9994
		(4.332E-3)	(3.707E-3)	(3.329E-3)	(3.294E-3)	(3.172E-3)
	α	3.9270	3.9270	3.9270	3.9270	3.9270
		(6.801E-4)	(6.015E-4)	(4.994E-4)	(4.747E-4)	(4.281E-4)
	β	2.3562	2.3562	2.3562	2.3562	2.3562
		(1.703E-5)	(1.414E-5)	(1.167E-5)	(1.081E-5)	(1.010E-5)
	γ	5.4978	5.4978	5.4978	5.4978	5.4978
		(7.350E-4)	(6.512E-4)	(5.708E-4)	(5.200E-4)	(4.9672-4)
	δ	1.0996	1.0996	1.0996	1.0996	1.0996
		(1.918E-5)	(1.505E-5)	(1.296E-5)	(1.210E-5)	(1.170E-5)
75	A	4.9998	4.9997	4.9998	4.9998	4.9998
		(7.814E-4)	(5.915E-4)	(4.225E-4)	(4.104E-4)	(4.011E-4)
	B	4.9995	4.9993	4.9994	4.9995	4.9995
		(1.947E-3)	(1.610E-3)	(1.241E-3)	(1.127E-3)	(1.125E-3)
	α	3.9270	3.9270	3.9270	3.9270	3.9270
		(3.823E-4)	(2.106E-4)	(1.835E-4)	(1.592E-4)	(1.558E-4)
	β	2.3562	2.3562	2.3562	2.3562	2.3562
		(4.925E-6)	(3.342E-6)	(2.918E-6)	(2.671E-6)	(2.596E-6)
	γ	5.4978	5.4978	5.4978	5.4978	5.4978
		(2.984E-4)	(2.481E-4)	(2.163E-4)	(2.072E-4)	(2.208E-4)
	δ	1.0996	1.0996	1.0996	1.0996	1.0996
		(4.316E-6)	(3.882E-6)	(3.346E-6)	(2.274E-6)	(3.436E-6)

Table 6: The average estimates and the square root of the MSEs of the different estimates when there are 20% outliers. The results are for Model 2.

less demanding than the LADEs or HMEs, we propose to use the WLSEs for all practical purposes when there are outliers in the data set.

5 Choice of Weight Function

It has been observed in our extensive simulation experiments that the WLSEs perform quite well when there are outliers in the data set. They in general perform better than the ordinary LSEs and their performances are at par with the other robust estimators like LADEs and HMEs, although computationally they are less demanding. On the other hand the performances of the WLSEs are better than LADEs and HMEs when there are no outliers in the data set. But one important question is how to choose the proper weight function? It is very clear from our experimental results that the weight function should be chosen in such a manner that it should put less weight where the outliers are present. The performances of the WLSEs depend significantly on the choice of the weight function.

In practice often it is not known where the outliers are present. Hence, we suggest the following. We should choose a class of weight functions which are continuous in $(0,1)\times(0,1)$ and satisfy Assumption 3, and put different weights at different locations. For example some of them might be putting less weights in the middle, whereas some of them might be putting less weights in other location of the data set. We compute the WLSEs for all the weight functions and choose that particular WLSEs for which the residual sum of squares (RSS) is the minimum. In the class of weight functions we should keep the weight function $w(s,t) = 1$ also. It corresponds to LSEs. Therefore, when there are no outliers in the data set, the proposed method chooses the LSEs, which are the most efficient estimators under this set up. We explain the method in the next section in details.

6 Synthetic Data Analysis

In this section we explain the above method when we do not know whether there is any outliers in the data set or not. Moreover, even if there are outliers in the data set it is not known the position of the outliers. We would like to use the proposed method to estimate the model parameters of the noisy image as provided in Figure 1(b) to extract the original image as in Figure 1(a). We have used WLSEs with different weight functions. Since it is not known whether there is any outliers present or not and even if they are present their location is not known, we have used different weight functions which put different weights at different places. We have used the following weight functions:

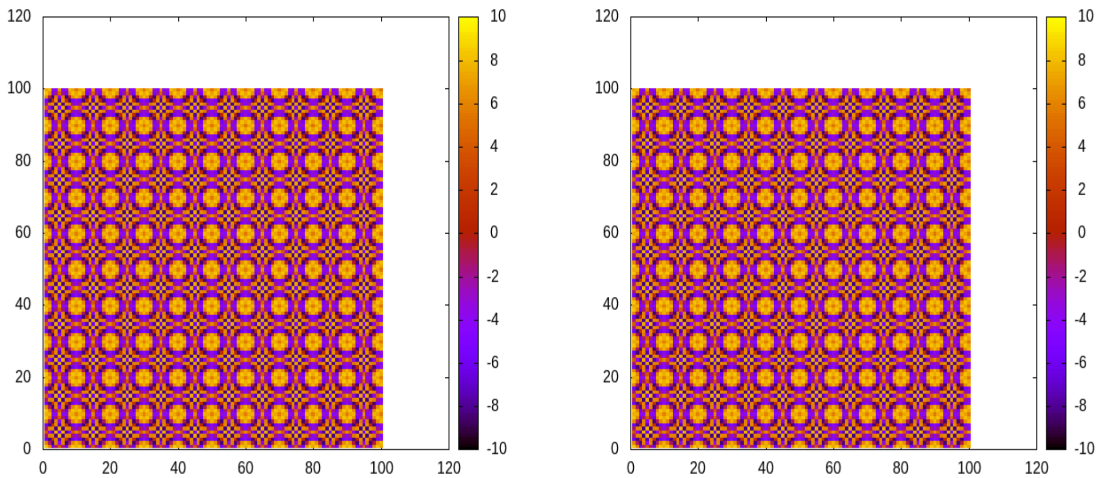
$$
\begin{aligned}
&W-1: && w(t,s) = 1\\
&W-2: && w(t,s) = 1 + (s-0.5)^2 + (t-0.5)^2\\
&W-3: && w(t,s) = 1 + |s-0.5)| + |t-0.5|\\
&W-4: && w(t,s) = 1 + (s-0.5)^2 + (t-0.5)^2 + |(s-0.5)\times(t-0.5)|\\
&W-5: && w(t,s) = 1 + (s-0.75)^2 + (t-0.75)^2\\
&W-6: && w(t,s) = 1 + (s-0.75)^2 + (t-0.75)^2 + |(s-0.75)\times(t-0.75)|\\
&W-7: && w(t,s) = 1 + |s-0.75| + |t-0.75|\\
&W-8: && w(t,s) = 1 + (s-0.25)^2 + (t-0.25)^2\\
&W-9: && w(t,s) = 1 + |s-0.25| + |t-0.25|\\
&W-10 && w(t,s) = 1 + |s-0.25| + |t-0.25| + |(s-0.25)(t-0.25)|.
\end{aligned}
$$

Apart from these WLSEs we have also used LADEs and HMEs for comparison purposes. We have reported the RSS for all the cases in Table 7. It is observed that $W-9$ provides the best fit among different WLSEs in terms of lowest RSS. It also provides better fit than LSEs, LADEs and HMEs. We have presented the estimated image plots based on $W-9$

estimators and also based on HMEs in Figure 2. Both of them estimate the true image as in Figure 1 (a) quite well. Hence, for all practical purposes we can use WLSEs as they are easy to implement in practice than HMEs and their theoretical properties are also known.

Methods	LSE	$W-2$	$W-3$	$W-4$	$W-5$	$W-6$
RSS	1960.93	3621.31	1741.14	1675.00	2391.94	1618.16
Methods	$W-7$	$W-8$	$W-9$	$W-10$	LSE	HME
RSS	1632.52	1624.97	1614.02	1707.53	1658.60	1632.45

Table 7: The residual sum of squares based on different estimators.



(a) (b)

Figure 2: The estimated image plot based on: (a) WLSEs ($W-9$) and (b) HMEs.

7 Multicomponent 2-D Chirp Model

In this section we have indicated how the proposed method can be used for the multicomponent 2-D chirp model. The multicomponent 2-D chirp model has the following form:

$$y(m,n) = \sum_{k=1}^{p} \left\{ A_k^0 \cos(\alpha_k^0 m + \beta_k^0 m^2 + \gamma_k^0 n + \delta_k^0 n^2) + B_k^0 \sin(\alpha_k^0 m + \beta_k^0 m^2 + \gamma_k^0 n + \delta_k^0 n^2) \right\} + X(m,n); \quad m = 1, \ldots, M; \quad n = 1, \ldots, N. \tag{15}$$

The model (15) is an extension of the model (1). Here also for $k = 1, \ldots, p$, $\{A_k^0, B_k^0\}$ are the amplitude parameters and $\{\alpha_k^0, \beta_k^0\}$ are the frequencies and $\{\gamma_k^0, \delta_k^0\}$ are frequency rates. The error random variables $X(m, n)$'s are same as defined before. The problem remains the same, i.e. to estimate the unknown parameters based on $\{y(m, n); m = 1, \ldots, M;\ n = 1, \ldots, N\}$.

The multicomponent 2-D chirp model has received a significant amount of attention in the signal processing literature. An extensive amount of work has been done in developing different estimation procedures and establishing their theoretical properties, see for example Francos and Friedlander [8, 9], Simeunovíc and Djurovíc [46], Grover et al. [14] and the references cited therein. Among the different estimators, the most efficient one are the LSEs, and it can be shown that under the assumption of i.i.d. Gaussian errors the asymptotic variances of the LSEs attend the Cramer-Rao lower bound. Although, the LSEs are most efficient estimators, they are not robust. In this case also we use WLSEs as robust estimators, in presence of outliers. The WLSEs can be obtained as follows. Let $\boldsymbol{\theta}_k = (A_k, B_k \alpha_k, \beta_k, \delta_k, \gamma_k)$, for $k = 1, \ldots, p$. The WLSEs of the unknown parameters can be obtained by minimizing

$$Q_{MW}(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_p) = \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} w\left(\frac{m}{M}, \frac{n}{N}\right) \left(y(m, n) - \sum_{k=1}^{p} \mu_{mn}(\boldsymbol{\theta}_k) \right)^2. \tag{16}$$

Here, $\mu_{mn}(\boldsymbol{\theta})$ is same as defined in (4). We need the following assumption to establish the asymptotic properties of the WLSEs of the model (15).

Assumption 4: It is assumed that there exists a M, such that $0 < |A_k^0|, |B_k^0| < M$, $0 < \alpha_k^0, \beta_k^0, \gamma_k^0, \delta_k^0 < \pi$, for $k = 1, \ldots, p$. Here, $\boldsymbol{\Theta}^p$, where $\boldsymbol{\Theta} = \{(A, B, \alpha, \beta, \gamma, \delta) : |A| < M, |B| < M, 0 < \alpha, \beta, \gamma, \delta < \pi\}$ is the parameter space. Further, $(\alpha_k^0, \beta_k^0, \gamma_k^0, \delta_k^0) \neq (\alpha_j^0, \beta_j^0, \gamma_j^0, \delta_j^0)$, for $1 \leq k \neq j \leq p$.

The following asymptotic results can be obtained along the same line as Theorems 1 and 2. The details are avoided.

Theorem 3: Under Assumptions 2,3 and 4, the WLSEs of $A_k^0, B_k^0, \alpha_k^0, \beta_k^0, \gamma_k^0, \delta_k^0$ for $k =$

$1, \ldots, p$, are strongly consistent.

Theorem 4: Under Assumptions 2,3 and 4, for $k = 1, \ldots, p$

$$\boldsymbol{D}\left(\widehat{A}_k - A_k^0, \widehat{B}_k - B_k^0, \widehat{\alpha}_k - \alpha_k^0, \widehat{\beta}_k - \beta_k^0, \widehat{\gamma}_k - \gamma_k^0, \widehat{\delta}_k - \delta_k^0\right)^\top \xrightarrow{d} N_6\left(\mathbf{0}, 2\sigma^2 \boldsymbol{\Gamma}_k^{-1} \boldsymbol{\Sigma}_k \boldsymbol{\Gamma}_k^{-1}\right).$$

Here, $\boldsymbol{\Gamma}_k$ ($\boldsymbol{\Sigma}_k$) can be obtained from $\boldsymbol{\Gamma}$ ($\boldsymbol{\Sigma}$) of Section 3, by replacing A^0 and B^0 with A_k^0 and B_k^0, respectively. Further, for $k \neq j$, they are asymptotically independent.

7.1 Sequential Least Squares Estimators

It may be observed that to compute the WLSEs of the model (15) one needs to solve a 4p dimensional optimization problem. The problem becomes computationally challenging if p is large. To avoid that one may use the sequential method as suggested by Grover et al. [14] for which one needs to solve p separate 4-D optimization problem. The following algorithm can be used for that purpose. Let us assume without loss of generality that

$$|A_1^0| + |B_1^0| > |A_2^0| + |B_2^0| > \ldots > |A_p^0| + |B_p^0|.$$

Algorithm

Step 1: First minimize with respect to $\boldsymbol{\theta}_1 = (A_1, B_1, \alpha_1, \beta_1, \delta_1, \gamma_1)$

$$Q_{MW}(\boldsymbol{\theta}_1) = \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} w\left(\frac{m}{M}, \frac{n}{N}\right) \left(y(m,n) - \mu_{mn}(\boldsymbol{\theta}_1)\right)^2,$$

and obtain the estimate of $\boldsymbol{\theta}_1^0$ as $\widetilde{\boldsymbol{\theta}}_1 = (\widetilde{A}_1, \widetilde{B}_1, \widetilde{\alpha}_1, \widetilde{\beta}_1, \widetilde{\delta}_1, \widetilde{\gamma}_1)$.

Step 2: Take out the effect of the first component $\boldsymbol{\theta}_1^0$, i.e. compute

$$y_1(m,n) = y(m,n) - \mu_{mn}(\widetilde{\boldsymbol{\theta}}_1); \quad m = 1, \ldots, M; n = 1, \ldots, N.$$

Step 3: Minimize with respect to $\boldsymbol{\theta}_2 = (A_2, B_2, \alpha_2, \beta_2, \delta_2, \gamma_2)$

$$Q_{MW}(\boldsymbol{\theta}_2) = \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} w\left(\frac{m}{M}, \frac{n}{N}\right) \left(y_1(m,n) - \mu_{mn}(\boldsymbol{\theta}_2)\right)^2,$$

and obtain the estimate of $\boldsymbol{\theta}_2^0$ as $\widetilde{\boldsymbol{\theta}}_2 = (\widetilde{A}_2, \widetilde{B}_2, \widetilde{\alpha}_2, \widetilde{\beta}_2, \widetilde{\delta}_2, \widetilde{\gamma}_2)$.

Step 4: Continue the process p times and obtain sequentially the estimates of $\boldsymbol{\theta}_1^0, \ldots, \boldsymbol{\theta}_p^0$, respectively.

It is computationally more tractable. Further, it can be shown along the same line as in Grover et al. [14] that the sequential WLSEs have the same asymptotic properties as the WLSEs.

8 Conclusions and Open Problems

In this paper we have proposed weighted least squares method to estimate unknown parameters of 2-D chirp model as robust estimators. The proposed method is quite easy to implement in practice and they can be used as an alternative to other robust estimators like LADEs or HMEs. Theoretical properties of the WLSEs have been established under the same set of assumptions as the LSEs. Extensive simulations indicate that the proposed estimators can be used quite effectively in practice as an alternative to other robust estimators. One synthetic data set has been analyzed to show how the method can be implemented in practice.

In this paper we have developed the WLSEs of 2-D chirp model when the amplitude parameters are fixed. The random amplitude 2-D chirp model also has received a considerable amount of attention in the signal processing literature, see Nandi et al. [36] and the references cited therein. In a random amplitude 2-D chirp model, say for example in model (1), A^0 and B^0 are not constants, they may be random variables, similar to the random amplitude 2-D sinusoidal model, see for example Kligler and Francos [21, 22]. The random amplitude 2-D chirp model has significant applications in image analysis and telecommunications. It will be interesting to develop WLSEs for random amplitude 2-D chirp model and to establish their properties. More work is needed in that direction.

Data Availability and Conflict of Interest Statement

No real data has been used in this paper. The author does not have any conflict of interest in the preparation of this manuscript.

Appendix A: Proof of Asymptotic Properties

We need the following lemmas to prove the necessary results.

Lemma 1: If $X(m,n)$ satisfies Assumption 2, then as $\min\{M,N\} \to \infty$, for $s,t = 0,1,\ldots,$

$$\sup_{0<\alpha,\beta,\gamma,\delta<\pi} \left| \frac{1}{M^{s+1}N^{t+1}} \sum_{m=1}^{M}\sum_{n=1}^{N} m^s n^t X(m,n)\cos(\alpha m + \beta m^2 + \gamma n + \delta n^2) \right| \to 0 \quad a.s.$$

and

$$\sup_{0<\alpha,\beta,\gamma,\delta<\pi} \left| \frac{1}{M^{s+1}N^{t+1}} \sum_{m=1}^{M}\sum_{n=1}^{N} m^s n^t X(m,n)\sin(\alpha m + \beta m^2 + \gamma n + \delta n^2) \right| \to 0 \quad a.s.$$

Proof: The proof follows from Lemma 2 of Lahiri and Kundu [30]. ∎

Lemma 2: If $X(m,n)$ satisfies Assumption 2 and $w(s,t)$ satisfies Assumption 3, then as $\min\{M,N\} \to \infty$,

$$\sup_{0<\alpha,\beta,\gamma,\delta<\pi} \left| \frac{1}{MN} \sum_{m=1}^{M}\sum_{n=1}^{N} X(m,n) w\left(\frac{m}{M},\frac{n}{N}\right)\cos(\alpha m + \beta m^2 + \gamma n + \delta n^2) \right| \to 0 \quad a.s.$$

and

$$\sup_{0<\alpha,\beta,\gamma,\delta<\pi} \left| \frac{1}{MN} \sum_{m=1}^{M}\sum_{n=1}^{N} X(m,n) w\left(\frac{m}{M},\frac{n}{N}\right)\sin(\alpha m + \beta m^2 + \gamma n + \delta n^2) \right| \to 0 \quad a.s.$$

Proof: The proof follows from Lemma 1. ∎

Lemma 3: If $X(m,n)$ satisfies Assumption 2, $w(s,t)$ satisfies Assumption 3 and $A,B,\alpha,\beta,\gamma,\delta$ satisfy Assumption 1, then as $\min\{M,N\} \to \infty$,

$$\sup_{0<\alpha,\beta,\gamma,\delta<\pi, |A|\le M, |B|\le M} \left| \frac{1}{MN} \sum_{m=1}^{M}\sum_{n=1}^{N} X(m,n) w\left(\frac{m}{M},\frac{n}{N}\right)\mu_{mn}(A,B,\alpha,\beta,\gamma,\delta) \right| \to 0 \quad a.s.$$

Here $\mu_{mn}(A,B,\alpha,\beta,\gamma,\delta)$ is defined in (4).

Proof: The proof follows from Lemma 2. ∎

Lemma 4: For any $\xi > 0$,

$$\liminf_{M,N\to\infty} \inf_{|\theta^0-\theta|>\xi} \left\{Q_W(\boldsymbol{\theta}) - Q_W(\boldsymbol{\theta}^0\right\} > 0.$$

Proof: Consider

$$\begin{aligned}
Q_W(\boldsymbol{\theta}) &= \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right)(y(m,n) - \mu_{mn}(\boldsymbol{\theta})^2 \\
&= \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right)(\mu_{mn}(\boldsymbol{\theta}^0) + X(m,n) - \mu_{mn}(\boldsymbol{\theta})^2 \\
&= \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right)(\mu_{mn}(\boldsymbol{\theta}^0) - \mu_{mn}(\boldsymbol{\theta})^2 \\
&\quad + \frac{2}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right)(\mu_{mn}(\boldsymbol{\theta}^0) - \mu_{mn}(\boldsymbol{\theta})X(m,n) \\
&\quad + \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right)X^2(m,n).
\end{aligned}$$

For any given $\xi > 0$, using Lemma 3 and Assumption 3, it follows that

$$\begin{aligned}
&\liminf_{M,N\to\infty} \inf_{|\theta^0-\theta|>\xi} \left\{Q_W(\boldsymbol{\theta}) - Q_W(\boldsymbol{\theta}^0\right\} = \\
&\quad \liminf_{M,N\to\infty} \inf_{|\theta^0-\theta|>\xi} \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right)(\mu_{mn}(\boldsymbol{\theta}^0) - \mu_{mn}(\boldsymbol{\theta})^2 \geq \\
&\quad \liminf_{M,N\to\infty} \inf_{|\theta^0-\theta|>\xi} \frac{\eta}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}(\mu_{mn}(\boldsymbol{\theta}^0) - \mu_{mn}(\boldsymbol{\theta})^2.
\end{aligned}$$

Consider the following sets:

$$\begin{aligned}
\Gamma_1 &= \{\boldsymbol{\theta} : \boldsymbol{\theta} \in \boldsymbol{\Theta}, |A - A^0| > \xi/2\}, \quad \Gamma_2 = \{\boldsymbol{\theta} : \boldsymbol{\theta} \in \boldsymbol{\Theta}, |B - B^0| > \xi/2\}, \\
\Gamma_3 &= \{\boldsymbol{\theta} : \boldsymbol{\theta} \in \boldsymbol{\Theta}, |\alpha - \alpha^0| > \xi/2\}, \quad \Gamma_4 = \{\boldsymbol{\theta} : \boldsymbol{\theta} \in \boldsymbol{\Theta}, |\beta - \beta^0| > \xi/2\}, \\
\Gamma_5 &= \{\boldsymbol{\theta} : \boldsymbol{\theta} \in \boldsymbol{\Theta}, |\gamma - \gamma^0| > \xi/2\}, \quad \Gamma_6 = \{\boldsymbol{\theta} : \boldsymbol{\theta} \in \boldsymbol{\Theta}, |\delta - \delta^0| > \xi/2\}.
\end{aligned}$$

Note that

$$\{\boldsymbol{\theta} : \boldsymbol{\theta} \in \boldsymbol{\Theta}, |\boldsymbol{\theta}^0 - \boldsymbol{\theta}| > \xi\} \subset \Gamma_1 \cup \Gamma_2 \cup \Gamma_3 \cup \Gamma_4 \cup \Gamma_5 \cup \Gamma_6.$$

Further

$$\liminf_{M,N\to\infty} \inf_{\theta\in\Gamma_1} \frac{\eta}{MN} \sum_{m=1}^{M}\sum_{n=1}^{N} (\mu_{mn}(\boldsymbol{\theta}^0) - \mu_{mn}(\boldsymbol{\theta})^2$$
$$\liminf_{M,N\to\infty} \frac{\eta|A-A^0|^2}{MN} \sum_{m=1}^{M}\sum_{n=1}^{N} \cos^2(\alpha^0 m + \beta^0 m^2 + \gamma^0 n + \delta^0 n^2) = \frac{\eta|A-A^0|^2}{2} > 0.$$

Similarly, it can be shown that

$$\liminf_{M,N\to\infty} \inf_{\theta\in\Gamma_j} \frac{\eta}{MN} \sum_{m=1}^{M}\sum_{n=1}^{N} (\mu_{mn}(\boldsymbol{\theta}^0) - \mu_{mn}(\boldsymbol{\theta})^2 > 0,$$

for $j = 2, \ldots, 6$. Since

$$\liminf_{M,N\to\infty} \inf_{|\theta^0-\theta|>\xi} \left\{Q_W(\boldsymbol{\theta}) - Q_W(\boldsymbol{\theta}^0\right\} \geq \liminf_{M,N\to\infty} \inf_{\theta\in\Gamma_1\cup\ldots\cup\Gamma_6} \left\{Q_W(\boldsymbol{\theta}) - Q_W(\boldsymbol{\theta}^0\right\},$$

the result follows. ∎

Proof of Theorem 1: Let us denote $\widehat{\boldsymbol{\theta}} = (\widehat{A}, \widehat{B}, \widehat{\alpha}, \widehat{\beta}, \widehat{\gamma}, \widehat{\delta})$ as the LSE of $\boldsymbol{\theta}$. In this proof only, we will denote $\widehat{\boldsymbol{\theta}}$ as $\widehat{\boldsymbol{\theta}}_{MN}$ just to emphasize that it is a function of M and N. Clearly,

$$Q(\widehat{\boldsymbol{\theta}}_{MN}) - Q(\boldsymbol{\theta}^0) < 0. \tag{17}$$

Suppose, $\widehat{\boldsymbol{\theta}}_{MN}$ is not a consistent estimator of $\boldsymbol{\theta}^0$. Since $\widehat{\boldsymbol{\theta}}_{MN} \in \boldsymbol{\Theta}$, a compact subset in $\mathbb{R}^6$, there exists a subsequence $\{M_k N_k\}$, such that $\widehat{\boldsymbol{\theta}}_{M_k N_k} \to \boldsymbol{\theta}' \neq \boldsymbol{\theta}^0$. Now due to (17),

$$Q(\boldsymbol{\theta}') - Q(\boldsymbol{\theta}) \leq 0,$$

this contradicts Lemma 4. Hence, $\widehat{\boldsymbol{\theta}}_{MN}$ is a consistent estimator of $\boldsymbol{\theta}^0$. ∎

Proof of Theorem 2: To prove this theorem, we need the multivariate Taylor series expansion of $(\boldsymbol{\theta} = MNQ_W(\boldsymbol{\theta})$ around the true value $\boldsymbol{\theta}^0$,

$$S(\boldsymbol{\theta}) = MNQ_W(\boldsymbol{\theta}) = \sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M}, \frac{n}{N}\right) \left(y(m,n) - \mu_{mn}(\boldsymbol{\theta})\right)^2.$$

It becomes

$$S'(\widehat{\boldsymbol{\theta}}) - S'(\boldsymbol{\theta}^0) = (\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0)S''(\overline{\boldsymbol{\theta}}), \tag{18}$$

where $S'(\boldsymbol{\theta})$ is a 1×6 gradient vector, i.e.

$$S'(\boldsymbol{\theta}) = \left(\frac{\partial S(\boldsymbol{\theta})}{\partial A}, \frac{\partial S(\boldsymbol{\theta})}{\partial B}, \frac{\partial S(\boldsymbol{\theta})}{\partial \alpha}, \frac{\partial S(\boldsymbol{\theta})}{\partial \beta}, \frac{\partial S(\boldsymbol{\theta})}{\partial \gamma}, \frac{\partial S(\boldsymbol{\theta})}{\partial \delta}\right),$$

$S''(\boldsymbol{\theta})$ is the 6×6 Hessian matrix, i.e.

$$S''(\boldsymbol{\theta}) = \begin{bmatrix} \frac{\partial^2 S(\boldsymbol{\theta})}{\partial A^2} & \frac{\partial^2 S(\boldsymbol{\theta})}{\partial A \partial B} & \frac{\partial^2 S(\boldsymbol{\theta})}{\partial A \partial \alpha} & \frac{\partial^2 S(\boldsymbol{\theta})}{\partial A \partial \beta} & \frac{\partial^2 S(\boldsymbol{\theta})}{\partial A \partial \gamma} & \frac{\partial^2 S(\boldsymbol{\theta})}{\partial A \partial \delta} \\ \frac{\partial^2 S(\boldsymbol{\theta})}{\partial B \partial A} & \frac{\partial^2 S(\boldsymbol{\theta})}{\partial B^2} & \frac{\partial^2 S(\boldsymbol{\theta})}{\partial B \partial \alpha} & \frac{\partial^2 S(\boldsymbol{\theta})}{\partial B \partial \beta} & \frac{\partial^2 S(\boldsymbol{\theta})}{\partial B \partial \gamma} & \frac{\partial^2 S(\boldsymbol{\theta})}{\partial B \partial \delta} \\ \frac{\partial^2 S(\boldsymbol{\theta})}{\partial \alpha \partial A} & \frac{\partial^2 S(\boldsymbol{\theta})}{\partial \alpha \partial B} & \frac{\partial^2 S(\boldsymbol{\theta})}{\partial \alpha^2} & \frac{\partial^2 S(\boldsymbol{\theta})}{\partial \alpha \partial \beta} & \frac{\partial^2 S(\boldsymbol{\theta})}{\partial \alpha \partial \gamma} & \frac{\partial^2 S(\boldsymbol{\theta})}{\partial \alpha \partial \delta} \\ \frac{\partial^2 S(\boldsymbol{\theta})}{\partial \beta \partial A} & \frac{\partial^2 S(\boldsymbol{\theta})}{\partial \beta \partial B} & \frac{\partial^2 S(\boldsymbol{\theta})}{\partial \beta \partial \alpha} & \frac{\partial^2 S(\boldsymbol{\theta})}{\partial \beta^2} & \frac{\partial^2 S(\boldsymbol{\theta})}{\partial \beta \partial \gamma} & \frac{\partial^2 S(\boldsymbol{\theta})}{\partial \beta \partial \delta} \\ \frac{\partial^2 S(\boldsymbol{\theta})}{\partial \gamma \partial A} & \frac{\partial^2 S(\boldsymbol{\theta})}{\partial \gamma \partial B} & \frac{\partial^2 S(\boldsymbol{\theta})}{\partial \gamma \partial \alpha} & \frac{\partial^2 S(\boldsymbol{\theta})}{\partial \gamma \partial \beta} & \frac{\partial^2 S(\boldsymbol{\theta})}{\partial \gamma^2} & \frac{\partial^2 S(\boldsymbol{\theta})}{\partial \gamma \partial \delta} \\ \frac{\partial^2 S(\boldsymbol{\theta})}{\partial \delta \partial A} & \frac{\partial^2 S(\boldsymbol{\theta})}{\partial \delta \partial B} & \frac{\partial^2 S(\boldsymbol{\theta})}{\partial \delta \partial \alpha} & \frac{\partial^2 S(\boldsymbol{\theta})}{\partial \delta \partial \beta} & \frac{\partial^2 S(\boldsymbol{\theta})}{\partial \delta \partial \gamma} & \frac{\partial^2 S(\boldsymbol{\theta})}{\partial \delta^2} \end{bmatrix}$$

and $\overline{\boldsymbol{\theta}}$ is a point on the line joining $\widehat{\boldsymbol{\theta}}$ and $\boldsymbol{\theta}^0$. Since, $S'(\widehat{\boldsymbol{\theta}}) = 0$, we can obtain the following

$$-S'(\boldsymbol{\theta}^0)\boldsymbol{D}^{-1} = (\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0)\boldsymbol{D}\boldsymbol{D}^{-1}S''(\overline{\boldsymbol{\theta}})\boldsymbol{D}^{-1} \tag{19}$$

from (18). Now using (4) and (5), we can write

$$\begin{aligned}
\frac{\partial S(\boldsymbol{\theta}^0)}{\partial A} &= -2\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M}, \frac{n}{N}\right) X(m,n) c_{mn}(\boldsymbol{\phi}^0) \\
\frac{\partial S(\boldsymbol{\theta}^0)}{\partial B} &= -2\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M}, \frac{n}{N}\right) X(m,n) s_{mn}(\boldsymbol{\phi}^0) \\
\frac{\partial S(\boldsymbol{\theta}^0)}{\partial \alpha} &= -2\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M}, \frac{n}{N}\right) mX(m,n)\left(-A^0 s_{mn}(\boldsymbol{\phi}^0) + B^0 c_{mn}(\boldsymbol{\phi}^0)\right) \\
\frac{\partial S(\boldsymbol{\theta}^0)}{\partial \beta} &= -2\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M}, \frac{n}{N}\right) m^2X(m,n)\left(-A^0 s_{mn}(\boldsymbol{\phi}^0) + B^0 c_{mn}(\boldsymbol{\phi}^0)\right) \\
\frac{\partial S(\boldsymbol{\theta}^0)}{\partial \gamma} &= -2\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M}, \frac{n}{N}\right) nX(m,n)\left(-A^0 s_{mn}(\boldsymbol{\phi}^0) + B^0 \phi_{mn}(\boldsymbol{\phi}^0)\right) \\
\frac{\partial S(\boldsymbol{\theta}^0)}{\partial \delta} &= -2\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M}, \frac{n}{N}\right) n^2X(m,n)\left(-A^0 c_{mn}(\boldsymbol{\phi}^0) + B^0 s_{mn}(\boldsymbol{\phi}^0)\right) \\
\frac{\partial^2 S(\boldsymbol{\theta}^0)}{\partial A^2} &= 2\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M}, \frac{n}{N}\right) c_{mn}^2(\boldsymbol{\phi}^0)
\end{aligned}$$

$$
\begin{aligned}
\frac{\partial^2 S(\boldsymbol{\theta}^0)}{\partial B^2} &= 2\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right) s_{mn}^2(\boldsymbol{\phi}^0) \\
\frac{\partial^2 S(\boldsymbol{\theta}^0)}{\partial \alpha^2} &= 2\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right) \Big\{X(m,n)(m^2A^0c_{mn}(\boldsymbol{\phi}^0) + m^2B^0s_{mn}(\boldsymbol{\phi}^0) \\
&\qquad\qquad + \left(mA^0s_{mn}(\boldsymbol{\phi}^0) - mB^0c_{mn}(\boldsymbol{\phi}^0)\right)^2\Big\} \\
\frac{\partial^2 S(\boldsymbol{\theta}^0)}{\partial \beta^2} &= 2\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right) \Big\{X(m,n)(m^4A^0c_{mn}(\boldsymbol{\phi}^0) + m^4B^0s_{mn}(\boldsymbol{\phi}^0) \\
&\qquad\qquad + \left(m^2A^0s_{mn}(\boldsymbol{\phi}^0) - m^2B^0c_{mn}(\boldsymbol{\phi}^0)\right)^2\Big\} \\
\frac{\partial^2 S(\boldsymbol{\theta}^0)}{\partial \gamma^2} &= 2\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right) \Big\{X(m,n)(n^2A^0c_{mn}(\boldsymbol{\phi}^0) + n^2B^0s_{mn}(\boldsymbol{\phi}^0)) \\
&\qquad\qquad + \left(nA^0s_{mn}(\boldsymbol{\phi}^0) - nB^0c_{mn}(\boldsymbol{\phi}^0)\right)^2\Big\} \\
\frac{\partial^2 S(\boldsymbol{\theta}^0)}{\partial \delta^2} &= 2\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right) \{X(m,n)(n^4A^0c_{mn}(\boldsymbol{\phi}^0) + n^4B^0s_{mn}(\boldsymbol{\phi}^0) \\
&\qquad\qquad + \left(n^2A^0s_{mn}(\boldsymbol{\phi}^0) - n^2B^0c_{mn}(\boldsymbol{\phi}^0)\right)^2\} \\
\frac{\partial^2 S(\boldsymbol{\theta}^0)}{\partial A\partial B} &= 2\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right) c_{mn}(\boldsymbol{\phi}^0)s_{mn}(\boldsymbol{\phi}^0) \\
\frac{\partial^2 S(\boldsymbol{\theta}^0)}{\partial A\partial \alpha} &= 2\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right) \left\{mX(m,n)s_{mn}(\boldsymbol{\phi}^0) + mB^0c_{mn}^2(\boldsymbol{\phi}^0) - mA^0c_{mn}(\boldsymbol{\phi}^0)s_{mn}(\boldsymbol{\phi}^0)\right\} \\
\frac{\partial^2 S(\boldsymbol{\theta}^0)}{\partial A\partial \beta} &= 2\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right) \left\{m^2X(m,n)s_{mn}(\boldsymbol{\phi}^0) + m^2B^0c_{mn}^2(\boldsymbol{\phi}^0) - m^2A^0c_{mn}(\boldsymbol{\phi}^0)s_{mn}(\boldsymbol{\phi}^0)\right\} \\
\frac{\partial^2 S(\boldsymbol{\theta}^0)}{\partial A\partial \gamma} &= 2\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right) \left\{nX(m,n)s_{mn}(\boldsymbol{\phi}^0) + nB^0c_{mn}^2(\boldsymbol{\phi}^0) - nA^0c_{mn}(\boldsymbol{\phi}^0)s_{mn}(\boldsymbol{\phi}^0)\right\} \\
\frac{\partial^2 S(\boldsymbol{\theta}^0)}{\partial A\partial \delta} &= 2\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right) \left\{n^2X(m,n)s_{mn}(\boldsymbol{\phi}^0) + n^2B^0c_{mn}^2(\boldsymbol{\phi}^0) - n^2A^0c_{mn}(\boldsymbol{\phi}^0)s_{mn}(\boldsymbol{\phi}^0)\right\} \\
\frac{\partial^2 S(\boldsymbol{\theta}^0)}{\partial B\partial \alpha} &= 2\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right) \left\{-mX(m,n)c_{mn}(\boldsymbol{\phi}^0) + mB^0c_{mn}(\boldsymbol{\phi}^0)s_{mn}(\boldsymbol{\phi}^0) - mA^0s_{mn}^2(\boldsymbol{\phi}^0)\right\}
\end{aligned}
$$

$$\frac{\partial^2 S(\boldsymbol{\theta}^0)}{\partial B \partial \beta} = 2\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right)\left\{-m^2 X(m,n)c_{mn}(\boldsymbol{\phi}^0) + m^2 B^0 c_{mn}(\boldsymbol{\phi}^0)s_{mn}(\boldsymbol{\phi}^0) - m^2 A^0 s_{mn}^2(\boldsymbol{\phi}^0)\right\}$$

$$\frac{\partial^2 S(\boldsymbol{\theta}^0)}{\partial B \partial \gamma} = 2\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right)\left\{-n X(m,n)c_{mn}(\boldsymbol{\phi}^0) + n B^0 c_{mn}(\boldsymbol{\phi}^0)s_{mn}(\boldsymbol{\phi}^0) - n A^0 s_{mn}^2(\boldsymbol{\phi}^0)\right\}$$

$$\frac{\partial^2 S(\boldsymbol{\theta}^0)}{\partial B \partial \delta} = 2\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right)\left\{-n^2 X(m,n)c_{mn}(\boldsymbol{\phi}^0) + n^2 B^0 c_{mn}(\boldsymbol{\phi}^0)s_{mn}(\boldsymbol{\phi}^0) - n^2 A^0 s_{mn}^2(\boldsymbol{\phi}^0)\right\}$$

$$\frac{\partial^2 S(\boldsymbol{\theta}^0)}{\partial \alpha \partial \beta} = 2\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right)\left\{(Ams_{mn}(\boldsymbol{\phi}^0) - Bmc_{mn}(\boldsymbol{\phi}^0))^2 + m^3 X(m,n)\mu(\boldsymbol{\theta}^0)\right\}$$

$$\frac{\partial^2 S(\boldsymbol{\theta}^0)}{\partial \alpha \partial \gamma} = 2\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right)\left\{mn(As_{mn}(\boldsymbol{\phi}^0) - Bc_{mn}(\boldsymbol{\phi}^0))^2 + mnX(m,n)\mu(\boldsymbol{\theta}^0)\right\}$$

$$\frac{\partial^2 S(\boldsymbol{\theta}^0)}{\partial \alpha \partial \delta} = 2\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right)\left\{mn^2(As_{mn}(\boldsymbol{\phi}^0) - Bc_{mn}(\boldsymbol{\phi}^0))^2 + mn^2X(m,n)\mu(\boldsymbol{\theta}^0)\right\}$$

$$\frac{\partial^2 S(\boldsymbol{\theta}^0)}{\partial \beta \partial \gamma} = 2\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right)\left\{m^2n(As_{mn}(\boldsymbol{\phi}^0) - Bc_{mn}(\boldsymbol{\phi}^0))^2 + m^2nX(m,n)\mu(\boldsymbol{\theta}^0)\right\}$$

$$\frac{\partial^2 S(\boldsymbol{\theta}^0)}{\partial \beta \partial \delta} = 2\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right)\left\{m^2n^2(As_{mn}(\boldsymbol{\phi}^0) - Bc_{mn}(\boldsymbol{\phi}^0))^2 + m^2n^2X(m,n)\mu(\boldsymbol{\theta}^0)\right\}$$

$$\frac{\partial^2 S(\boldsymbol{\theta}^0)}{\partial \gamma \partial \delta} = 2\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right)\left\{(Ans_{mn}(\boldsymbol{\phi}^0) - Bnc_{mn}(\boldsymbol{\phi}^0))^2 + n^3 X(m,n)\mu(\boldsymbol{\theta}^0)\right\}.$$

Now using Central limit theorem, it follows that

$$S'(\boldsymbol{\theta}^0)\boldsymbol{D} \xrightarrow{d} N_6(\mathbf{0}, 2\sigma^2\boldsymbol{\Sigma}).$$

Also observe that $\overline{\boldsymbol{\theta}} \to \boldsymbol{\theta}^0$, and

$$\lim_{M,N\to\infty} \boldsymbol{D}^{-1}S''(\overline{\boldsymbol{\theta}})\boldsymbol{D}^{-1} = \lim_{M,N\to\infty} \boldsymbol{D}^{-1}S''(\boldsymbol{\theta}^0)\boldsymbol{D}^{-1} = \boldsymbol{\Gamma},$$

hence, the result follows. ∎

References

[1] Barbieri, M.M. and Barone, P. (1992), "A two dimensional Prony's method for spectral estimation", *IEEE Transactions on Signal Processing*, vol. 40, 2747 – 2756.

[2] Cabrera S. D. and Bose, N.K. (1993), "Prony's method for two dimensional complex exponentials modeling", *Applied Control*, S. G. Tzafestas (ed.), Chapter 15, Dekker, New York, 401–411.

[3] Chun J. and N. K. Bose (1995), "Parameter estimation via signal selectivity of signal subspaces (PESS) and its applications to 2-D wave number estimation", *Digital Signal Processing*, vol. 5, 58–76.

[4] Djurovíc, I., Wang, P. and Ioana, C. (2010). Parameter estimation of 2-D cubic phase signal using cubic phase function with genetic algorithm. Signal Processing, 90(9), 2698-2707.

[5] Fan, C. and Liu, C. (2015), "A Novel Algorithm for Circle Curve Fitting Based on the Least Square Method by the points of the Newton's rings", *IEEE International Conference on Computers, Communications, and Systems*, 256 – 260.

[6] Fu, L., Zhang, M., Liu, Z. and Li, H. (2019), "Robust Frequency Estimation of Multi-sinusoidal Signals Using Orthogonal Matching Pursuit with Weak Derivatives Criterion", *Circuits, Systems, and Signal Processing*, vol. 38, 1194–1205.

[7] Francos, J. M. and Friedlander, B. (1995, October). The polynomial phase difference operator for modeling of nonhomogeneous images. In Proceedings., International Conference on Image Processing (Vol. 2, pp. 276-279). IEEE.

[8] 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.

[9] 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.

[10] Francos, J.M., Meiri, A.Z. and Porat, B. (1993), "A united texture model based on 2-D Wald like decomposition", *IEEE Transactions on Signal Processing*, vol. 41, 2665 – 2678.

[11] Friedlander, B. and Francos, J. M. (1996), "Model based phase unwrapping of 2-D signals", *IEEE Transactions on Signal Processing*, vol. 44, 2999–3007.

[12] Gorthi, S.S. and Rastogi, P. (2009), "Estimation of phase derivatives using discrete chirp-Fourier-transform-based method", *Optics Letters*, vol. 34, 2396–2398.

[13] Gorthi, S.S. and Rastogi, P. (2009), "Simultaneous measurement of displacement, strain and curvature in digital holographic interferometry using high-order instantaneous moments", *Optical Express*, vol. 17, 17784–17791.

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

[15] Grover, R., Kundu, D. and Mitra, A. (2021), "An efficient methodology to estimate the parameters of a two-dimensional chirp signal model", *Multidimensional Systems and Signal Processing*, vol. 32, no. 1, 49 –75.

[16] Grover, R., Sharma, A., Delcourt, T. and Kundu, D. (2022), "Computationally efficient algorithm for frequency estimation of a two-dimensional sinusoidal model", *Circuits, Systems and Signal Processing*, vol. 41, no. 1, 346 – 371.

[17] Guo, Y. and Li, B.Z. (2018), "Novel method for parameter estimation of Newton's rings based on CFRFT and ER-WCA", *Signal Processing*, vol. 144, 118–126.

[18] Guo, Z., Lu, M-F., Wu, J-M., Jin, X-W., Zhang, F. and Tao, R. (2019), "Fast FRFT-based method for estimating physical parameters from Newton's rings", *Applied Optics*, vol. 58, No. 14, 3926 – 3931.

[19] Gupta, R.D. and Kundu, D. (1998), "Asymptotic Properties of the Least Squares Estimators of a Two Dimensional Texture Model", *Metrika* , Vol. 48, No. 2, 83-97, 1998.

[20] Hua Y. (1992), "Estimating two dimensional frequencies by matrix enhancement and matrix pencil", *IEEE Transaction on Signal Processing*, vol. 40, 2267–2280.

[21] Kliger, M. and Francos, J.M. (2008), "Strong Consistency of a Family of Model Order Selection Rules for Estimating 2-D Sinusoids in Noise", *Statistics and Probability Letters*, vol. 78, 3075–3081.

[22] Kliger, M. and Francos, J.M. (2013), "A Strongly Consistent Model Order Selection Rule for Estimating the Parameters of 2-D Sinusoids in Colored Noise", *IEEE Transactions on Information Theory*, vol. 59, no. 7, 4408 –4422.

[23] Kundu, D. (2024), "On robust method of estimation in polynomial sinusoidal regression", *Signal Processing*, vol. 217, Paper No. 109328.

[24] Kundu, D. (2024), "Robust parameter estimation in two-channel sinusoidal frequency model", *Franklin Open*, vol. 6, Paper No. 100059.

[25] Kundu, D. (2025), "Robust parameter estimation of Newton's Rings", *Circuits, Systems and Signal Processing*, (to appear).

[26] Kundu, D. and Grover, R. (2023), "Robust estimators of two dimensional sinusoidal model parameters", *Signal Processing*, vol. 212, Paper No. 109167.

[27] Kundu, D. and Nandi, S. (2003), "Determination of Discrete Spectrum in a Stationary Random Field", *Statistica Neerlandica*, vol. 57, no. 2, 258–284.

[28] Kundu, D., Nandi, S. and Grover, R. (2024), "On Weighted Least Squares Estimators of Parameters of a Chirp Model", *Circuits, Systems and Signal Processing*, Vol. 42, No. 1, 493 - 521.

[29] Kundu, D., Nandi, S. and Grover, R. (2024), “On weighted least squares estimators for chirp like model”, *Sankhya A*, vol. 86, no. 1, 27 – 66.

[30] Lahiri, A. and Kundu, D. (2017), “On parameter estimation of two-dimensional polynomial phase signal model”, *Statistica Sinica*, vol. 27, no. 4, 1779 – 1792.

[31] Lahiri, A., Kundu, D. and Mitra, A. (2013), “Efficient algorithm for estimating the parameters of two dimensional chirp signal”, *Sankhya, Ser. B*, vol. 75, no. 1, B, 65 – 89.

[32] Lahiri, A., Kundu, D. and Mitra, A. (2014), “On least absolute deviation estimator of one dimensional chirp model”, *Statistics*, vol. 48, no. 2, 405 – 420.

[33] Lang S. W. and J. H. McClellan (1982), “The extension of Pisarenkos method to multiple dimensions”, *Proceedings ICASSP-82*, 125–128.

[34] Mittal, A., Kundu, D. and Mitra, A. (2025), “On weighted least squares estimation of elementary chirp model”, *IEEE Transactions on Aerospace and Electronic Systems*, (to appear).

[35] Mondal, S.B. and Achilefu, S. (2021), “Virtual and Augmented Reality Technologies in Molecular and Anatomical Imaging”, *Molecular Imaging: Principles and Practice*, 2nd Edition, Chapter 79, 1591 – 1612.

[36] Nandi, S., Grover, R. and Kundu, D. (2022), “Estimating of parameters of two-dimensional random amplitude chirp signal in additive noise”, *Multidimensional Systems and Signal Processing*, vol. 33, 1045 –1068.

[37] Nandi, S., Kundu, D. and Srivastava, R.K. (2013), “Noise Space Decomposition method for two dimensional sinusoidal model”, *Computational Statistics and Data Analysis*, vol. 58, no. 2, 147 – 161.

[38] Oshea, P. (2002), "A new technique for instantaneous frequency rate estimation", *IEEE Signal Processing Letters*, vol. 9, 251–252.

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

[40] Prasad, A., Nandi, S. and Kundu, D. (2010), "An efficient and fast algorithm for estimating the parameters of two-dimensional sinusoidal signals", *Journal of Statistical Planning and Inference*, vol. 140, 153 – 168.

[41] Rao, C.R., Zhao, L.C. and Zhou, B. (1993), "A novel algorithm for 2-dimensional frequency estimation", *Conference Record of The Twenty-Seventh Asilomar Conference on Signals, Systems and Computers*, vol. 1 (2), November 1–3, IEEE Computer Soc. Press, Silver Spring, MD.

[42] Rao, C.R., Zhao, L.C. and Zhou, B. (1994), "Maximum likelihood estimation of 2-D superimposed exponential signals", *IEEE Transactions on Signal Processing*, vol. 42, 1795-1802

[43] Richards, F.S.G. (1961), "A method of maximum likelihood estimation", *Journal of the Royal Statistical Society*, Ser. B, vol. 23, 469–476

[44] Shukla, A., Grover, R., Kundu, D. and Mitra, A. (2022), "Approximate least squares estimators of a two-dimensional chirp model", *Journal of Multivariate Analysis*, Vol. 192, Artilce No. 105045.

[45] Shukla, A., Grover, R., Kundu, D. and Mitra, A. (2023), "A computationally efficient algorithm to estimate the parameters of a two-dimensional chirp model", *Multidimensional Sysytems and Signal Processing*, vol. 34, 633 - 655.

[46] Simeunovíc, M. and Djurovíc, I. (2015), “Parameter estimation of multicomponent 2D polynomial- phase signals using the 2D PHAF-based approach”, *IEEE Transactions on Signal Processing*, vol. 64, 771–782.

[47] Super, B.J. and Bovik, A.C. (1995), “Shape from texture using local spectralmoments”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, pp.333–343.

[48] Tekavec, P.F., Myers, J.A., Lewis, K.L.M., Fuller, F.D. and Ogilvie, J.P. (2010), “Effects of chirp on two-dimensional Fourier transform electronic spectra”, *Opt Express*, vol. 18, no. 11, 11015–11024.

[49] Wang, Y. and Jiang, Y. (2009), “ISAR imaging of a ship target using product high-order matched-phase transform”, *IEEE Geoscience and Remote Sensing Letters*, vol. 6, 658–661.

[50] Zhang, H. and Mandrekar, V. (2001), “Estimation of hidden frequencies for 2D stationary processes”, *Journal of Time Series Analysis*, vol. 22, 613 - 629.

[51] Zhang, Y., Mobasseri, B. G., Dogahe, B. M. and Amin, M. G. (2010), “Image-adaptive watermarking using 2D Chirps”, *Signal, Image and Video Processing*, vol. 4, 105-121.

[52] Zhang, K., Wang, S. and Cao, F. (2008, May). 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). IEEE.

[53] Zhang, C., Zhang, T., Yin, F. and Zoubir, A.M. (2023), “Data-adaptive M-estimators for robust regression via bi-level optimization”, *Signal Processing*, vol. 210, Paper # 109063.