A Computationally Efficient algorithm to estimate the Parameters of a Two-Dimensional Chirp Model

Abhinek Shukla[1,3], Rhythm Grover[2], Debasis Kundu[1], and Amit Mitra[1]

[1]Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Kanpur - 208016, India

[2]Mehta Family School of Data Science and Artificial Intelligence, Indian Institute of Technology Guwahati, Assam-781039, India

[3]Corresponding author. Email: abhushukla@gmail.com

Abstract

Starting from the radar problem, chirp signals and its several generalizations appear in many natural and man-made systems. Various methods have been proposed for their parameter estimation but these methods are either statistically sub-optimal or computationally expensive. This paper bridges the gap between optimality and practical feasibility of estimators of the parameters of chirp signal model. The proposed estimators are significantly much faster to compute than the conventional least squares method and at the same time, they have the same rates of convergence as the latter. Moreover, the estimators of chirp rate parameters are shown to be asymptotically optimal. Extensive numerical simulations have been executed to assess the performance of the proposed estimators. In addition, a comparative simulation study demonstrates that our algorithm is fast as well as provides accurate results.

Keywords: Chirp model, least squares, stationary linear process, linear frequency modulated signal, asymptotic normality.

1 Introduction

This paper addresses the problem of parameter estimation of two-dimensional (2D) chirp signal model defined as follows:

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

Here, $y(m,n)$ is the observed real valued signal and $X(m,n)$ is the additive noise term. A^0, B^0 are amplitude parameters, α^0, γ^0 are frequency, β^0, δ^0 are frequency rates and μ^0 is the coefficient of product term. $\boldsymbol{\xi}^0 = (\alpha^0, \beta^0, \gamma^0, \delta^0, \mu^0)^\top$ represents vector of non-linear parameters. This model can be used to describe signals having constant amplitude with frequency to be a linear function of spatial co-ordinates. Both of the frequency parameters contain a common parameter μ^0 acting as interaction for both dimensions.

Due to wide-spread applicability of the underlying model, this parameter estimation problem is encountered in many real-life applications such as 2D-homomorphic signal processing, magnetic resonance imaging (MRI), optical imaging, interferometric synthetic aperture radar (INSAR), modeling non-homogeneous patterns in the texture image captured by a camera due to perspective or orientation, see [6],[8] and [9]. 2D chirps 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 [24]. Newton's rings are one of the important interference fringes in the domain of optical metrology. It is predominantly used in testing spherical and flatting optical surface and curvature radius measurement. 2D chirp signals have also been used to model Newton's rings [13].

Many algorithms based on different approaches have been put forward in the literature to solve such problems. Polynomial phase differencing (PD) operator was introduced in [6] as an extension of the polynomial phase transform proposed in [20]. Several works [7], [8] and [9] utilized PD operator to develop computationally efficient algorithms for estimating similar polynomial phase signals. Application of another algorithm using Cubic phase function (CPF) proposed in [19], was extended in [23] for similar 2D chirp signals. Further CPF was utilized for estimating 2D cubic phase signal using genetic algorithm in [3]. Consistency and asymptotic normality of least squares estimators (LSEs) for a general order 2D polynomial phase signal (PPS) model has been shown in [16]. Another algorithm of finite step computationally efficient procedure for similar 2D chirp signal model proposed in [14] was proved asymptotically equivalent to LSEs. Quasi-

Maximum Likelihood (QML) algorithm [4] proposed for 1D PPS, was generalized for 2D PPS in [5]. Further approximate least squares estimators (ALSEs) proposed in [11] have been proved to be asymptotically equivalent to LSEs. An efficient estimation procedure based on fixed dimension technique, presented in [12] was shown equivalent to that of optimal LSEs.

Estimators based on phase differencing strategies or high order ambiguity function (HAF) or some of their modifications are computationally easier to obtain because of algorithmic simplicity. Their performance degrades considerably faster below a certain signal-to-noise ratio (SNR) threshold and are sub-optimal. Methods that use PD at a certain stage of estimation usually estimate coefficients of highest degree first, and then subsequently estimate the coefficients of a lower degree from the demodulated or de-chirped signal. Therefore, the estimation error of highest degree coefficients passes on and affects estimation accuracy of lower degree coefficients quite seriously. For more details, one can refer to [1], [5] and [22]. There has been less detailed study of the theoretical properties of the estimators such as CPF and QML to support their computational efficacy. Although optimal estimators for a similar model have already been developed in [12], but these cannot be generalized directly for the underlying model (1). The model considered in this paper is more general, as it takes into account the interaction term $\mu^0 mn$ and the error is linear stationary, which is a broader class of errors than i.i.d. errors considered in [12]. Due to the presence of this interaction term coefficient μ^0, the estimation becomes difficult as the estimators of α^0 and γ^0 are now dependent, making their computation as well as theoretical analysis challenging.

The main contributions of this paper are the proposal of an algorithm to estimate the parameters of the model defined in (1) which is computationally efficient and a detailed study of the asymptotic properties of the proposed estimators. This algorithm is motivated by the fact that a 2D chirp model with five non-linear parameters can be visualized as many 1D chirp models with two non-linear parameters and hence simplifies the computational cost.

Therefore, the key attribute of this method is that it is computationally faster than the conventional optimal methods such as LSEs, maximum likelihood estimators, or ALSEs and at the same time have good statistical properties. In fact, if we consider only the proposed estimators of the chirp rate parameters, then these are observed to have the same asymptotic variance as that of the traditional LSEs, and hence are statistically optimal.

The methodology to obtain the proposed estimators is presented in Section 2. The model assumptions as well as the derived theoretical results are stated in Section 3. In Section 4, the finite sample performance of the proposed estimators is demonstrated. In this section, a comparison of the performance of these estimators with the state-of-the-art methods such as least squares method, approximate least squares method, and 2D multilag HAF method is also presented alongwith the ability of the proposed estimators to extract a gray-scale texture from the one which is contaminated with noise. Finally, Section 5 concludes the paper, followed by the detailed proofs in appendices.

2 Estimation Methodology

If we arrange the observed data into a matrix of order $M \times N$, then the proposed algorithm rests on the observation that for each fixed column (or row), the 2D chirp model breaks down to a cascade of 1D chirp models. This is shown below mathematically. So our methodology is developed by estimating parameters of these 1D chirps based on a particular column (or row) vector of data matrix, rather than estimating the whole 2D chirp parameters based on the full data matrix.
If we fix one dimension say $n = n_0$ in (1), then the 2D chirp can be seen as 1-D chirp for $m = 1, 2, \ldots, M$, as follows:

$$\begin{aligned} y(m, n_0) = A^0(n_0) \cos\left((\alpha^0 + n_0\mu^0)m + \beta^0 m^2\right) + B^0(n_0) \sin\left((\alpha^0 + n_0\mu^0)m + \beta^0 m^2\right) \\ + X(m, n_0), \end{aligned} \tag{2}$$

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

Similarly, for a fixed $m = m_0$, we have 1-D chirp for $n = 1, 2, \ldots, N$

$$\begin{aligned} y(m_0, n) = A^0(m_0) \cos\left((\gamma^0 + m_0\mu^0)n + \delta^0 n^2\right) + B^0(m_0) \sin\left((\gamma^0 + m_0\mu^0)n + \delta^0 n^2\right) \\ + X(m_0, n). \end{aligned} \tag{3}$$

Equation (2) represents 1-D chirp signal model with $\alpha^0 + n_0\mu^0$ and β^0 as the frequency and frequency rate parameters. Similarly, equation (3) represents 1-D chirp signal model with $\gamma^0 + m_0\mu^0$ and δ^0 as the frequency and frequency rate parameters.

Let the data matrix be denoted as

$$\boldsymbol{Y} = \begin{bmatrix} y(1,1) & y(1,2) & \ldots & y(1,N) \\ y(2,1) & y(2,2) & \ldots & y(2,N) \\ \vdots & \vdots & \ddots & \vdots \\ y(M,1) & y(M,2) & \ldots & y(M,N) \end{bmatrix}_{M \times N}.$$

Further suppose column vector $\boldsymbol{Y}_{Mn_0}$ denotes the n_0^{th} column of data matrix $\boldsymbol{Y}$ and column vector $\boldsymbol{Y}_{m_0N}$ denotes the transpose of m_0^{th} row of data matrix $\boldsymbol{Y}$. Define $Z_k(\alpha_1, \alpha_2)$ matrix as

$$Z_k(\alpha_1, \alpha_2) = \begin{bmatrix} \cos(\alpha_1 + \alpha_2) & \sin(\alpha_1 + \alpha_2) \\ \cos(\alpha_1 2 + \alpha_2 2^2) & \sin(\alpha_1 2 + \alpha_2 2^2) \\ \vdots & \vdots \\ \cos(\alpha_1 k + \alpha_2 k^2) & \sin(\alpha_1 k + \alpha_2 k^2) \end{bmatrix}_{k\times 2}. \tag{4}$$

We need to estimate the model parameters of (2) which is a 1D chirp model, so we use LSEs to serve the purpose. We obtain LSEs of non-linear parameters in model (2) for fixed n_0 by defining following reduced sum of squares, see [12]:

$$R_{Mn_0}(\alpha_1, \alpha_2) = \boldsymbol{Y}_{Mn_0}^T \Big(I_{M\times M} - P_{Z_M}(\alpha_1, \alpha_2)\Big)\boldsymbol{Y}_{Mn_0}, \tag{5}$$

where, $P_{Z_M}(\alpha_1, \alpha_2) = Z_M(\alpha_1, \alpha_2)\Big(Z_M(\alpha_1, \alpha_2)^\top Z_M(\alpha_1, \alpha_2)\Big)^{-1} Z_M(\alpha_1, \alpha_2)^\top$ and $I_{M\times M}$ is the $M \times M$ identity matrix.

Then,

$$(\widehat{\alpha}_{n_0}, \widehat{\beta}_{n_0})^\top = \arg\min_{\alpha_1,\alpha_2} R_{Mn_0}(\alpha_1, \alpha_2) \tag{6}$$

is the proposed estimator of $(\alpha^0 + n_0\mu^0, \beta^0)^\top$ based on minimizing the sum of squares corresponding to n_0 column of the data matrix $\boldsymbol{Y}$.

Similarly, we can obtain LSEs of parameters in model (3) by defining

$$R_{m_0N}(\alpha_1, \alpha_2) = \boldsymbol{Y}_{m_0N}^T \Big(I_{N\times N} - P_{Z_N}(\alpha_1, \alpha_2)\Big)\boldsymbol{Y}_{m_0N}, \tag{7}$$

where, $P_{Z_N}(\alpha_1, \alpha_2) = Z_N(\alpha_1, \alpha_2)\Big(Z_N(\alpha_1, \alpha_2)^\top Z_N(\alpha_1, \alpha_2)\Big)^{-1} Z_N(\alpha_1, \alpha_2)^\top$ and $I_{N\times N}$ is the $N \times N$ identity matrix.

Then,

$$(\widehat{\gamma}_{m_0}, \widehat{\delta}_{m_0})^\top = \arg\min_{\alpha_1,\alpha_2} R_{m_0N}(\alpha_1, \alpha_2) \tag{8}$$

will be the proposed estimator of $(\gamma^0 + m_0\mu^0, \delta^0)^\top$ based on minimizing the sum of squares corresponding to m_0 row of data matrix $\boldsymbol{Y}$.

We observe that for each fixed column, we get estimate of same chirp rate parameter β^0 in (6), and also estimate of frequency parameter which is a linear combination of α^0 and μ^0. Similarly estimates of δ^0 and a linear combination of γ^0 and μ^0 for a fixed row in (8) has been obtained.

We stress the fact that linearity of parameters of 1D-chirp models is playing crucial role in getting proposed estimators of α^0, γ^0 and μ^0 by fitting a linear regression model as follows.

Once the parameters corresponding to each $(M+N)$ 1-D chirp models have been estimated. We apply the following three steps to obtain final estimates of parameters of the model (1):

Step-1. Let $\mathbf{\Gamma}^\top = \begin{bmatrix} 1 & 1 & \cdots & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 1 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & N & 1 & 2 & \cdots & M \end{bmatrix}$, $\mathbf{\Lambda}^\top = \begin{bmatrix} \widehat{\alpha}_1 & \widehat{\alpha}_2 & \cdots & \widehat{\alpha}_N & \widehat{\gamma}_1 & \widehat{\gamma}_2 & \cdots & \widehat{\gamma}_M \end{bmatrix}$, combine the obtained estimates as follows:

$$\mathbf{\Lambda} = \mathbf{\Gamma} \begin{bmatrix} \alpha \\ \gamma \\ \mu \end{bmatrix}. \tag{9}$$

Then estimate of $(\alpha^0, \gamma^0, \mu^0)^\top$ is $\left(\mathbf{\Gamma}^\top \mathbf{\Gamma}\right)^{-1} \mathbf{\Gamma}^\top \mathbf{\Lambda}$.

Step-2. The estimators of β^0 and δ^0 are simply the averages $\widehat{\beta} = \frac{1}{N}\sum_{n=1}^{N} \widehat{\beta}_n$ and $\widehat{\delta} = \frac{1}{M}\sum_{m=1}^{M} \widehat{\delta}_m$, respectively. The estimates of non-linear parameters obtained as $\widehat{\boldsymbol{\xi}} = (\widehat{\alpha}, \widehat{\beta}, \widehat{\gamma}, \widehat{\delta}, \widehat{\mu})^\top$.

Step-3. After getting estimates of non-linear parameters $\widehat{\boldsymbol{\xi}} = (\widehat{\alpha}, \widehat{\beta}, \widehat{\gamma}, \widehat{\delta}, \widehat{\mu})^\top$, the amplitude parameter estimators are as follows:

$$\begin{bmatrix} \widehat{A} \\ \widehat{B} \end{bmatrix} = \begin{bmatrix} \frac{2}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N} y(m,n)\cos\widehat{\phi} \\ \frac{2}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N} y(m,n)\sin\widehat{\phi} \end{bmatrix}. \tag{10}$$

3 Theoretical Results

In this section, we first state the model assumptions explicitly. These are as follows:

Assumption 1. *$X(m,n)$ can be expressed as a linear combination of a double array sequence of independently and identically distributed (i.i.d.) random variables $\{\epsilon(m,n)\}$ with mean 0, variance σ^2 and finite fourth moment.*

$$X(m,n) = \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} a(i,j)\epsilon(m-i, n-j), \tag{11}$$

such that

$$\sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty} |a(i,j)| < \infty. \tag{12}$$

Assumption 2. *True parameter* $\boldsymbol{\theta}^0 = (A^0, B^0, \alpha^0, \beta^0, \gamma^0, \delta^0, \mu^0)^\top$ *is an interior point of parameter space* Θ*, where* $\Theta = (-\infty, \infty) \times (-\infty, \infty) \times [0, 2\pi] \times [0, \pi/2] \times [0, 2\pi] \times [0, \pi/2] \times [0, 2\pi]$*. Amplitude of the model is positive, i.e.* $A^{0^2} + B^{0^2} > 0$*.*

Assumption 1 puts the model under a very general condition on noise contamination as it incorporates the dependent relationship too. Assumption 2 is taken to assure the absence of any identifiability problem and non-zero deterministic part of the signal. Under these general assumptions, we have derived strong consistency and asymptotic normality of the estimators. The obtained results are stated in the following theorems.

Theorem 1. *Under assumptions 1 and 2, the proposed estimator of parameter* $\boldsymbol{\theta}^0$ *is strongly consistent, i.e.*

$$\widehat{\boldsymbol{\theta}} \xrightarrow{a.s.} \boldsymbol{\theta}^0 \quad \text{as } \min\{M, N\} \to \infty.$$

Proof. Please see Appendix A for the proof. □

Theorem 2. *Under assumptions 1 and 2, the proposed estimators of* $\boldsymbol{\theta}^0$ *is asymptotically normally distributed.*

$$\boldsymbol{D}^{-1}(\widetilde{\boldsymbol{\theta}} - \boldsymbol{\theta}^0) \xrightarrow{d} \mathcal{N}_7(0, \boldsymbol{\Sigma}) \quad \text{as } M = N \to \infty,$$

where $c = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} a(i,j)^2$, $\quad \boldsymbol{D}^{-1} = diag(\ M^{\frac{3}{2}}N^{\frac{1}{2}},\ M^{\frac{5}{2}}N^{\frac{1}{2}},\ M^{\frac{1}{2}}N^{\frac{3}{2}},\ M^{\frac{1}{2}}N^{\frac{5}{2}},\ M^{\frac{3}{2}}N^{\frac{3}{2}}\ )$*, and*

$$\boldsymbol{\Sigma} = \frac{c\sigma^2}{(A^{0^2} + B^{0^2})} \begin{bmatrix} 2A^{0^2} + 187B^{0^2} & -185A^0B^0 & -378B^0 & 60B^0 & -378B^0 & 60B^0 & 612B^0 \\ -185A^0B^0 & 2B^{0^2} + 187A^{0^2} & 378A^0 & -60A^0 & 378A^0 & -60A^0 & -612A^0 \\ -378B^0 & 378A^0 & 996 & -360 & 612 & 0 & -1224 \\ 60B^0 & -60A^0 & -360 & 360 & 0 & 0 & 0 \\ -378B^0 & 378A^0 & 612 & 0 & 996 & -360 & -1224 \\ 60B^0 & -60A^0 & 0 & 0 & -360 & 360 & 0 \\ 612B^0 & -612A^0 & -1224 & 0 & -1224 & 0 & 2448 \end{bmatrix}.$$

Here $diag(a_1, a_2, ..., a_k)$ *represents* $k \times k$ *diagonal matrix with elements* $a_1, a_2, ..., a_k$ *in the principal diagonal and* $\mathcal{N}_k(\boldsymbol{\mathcal{M}}, \boldsymbol{\mathcal{S}})$ *represents the* k*-variate normal distribution with mean vector* $\boldsymbol{\mathcal{M}}$ *and variance-covariance matrix* $\boldsymbol{\mathcal{S}}$*.*

Proof. Please see Appendix B for the proof. □

Although Theorem 2 has been proved for the increasing sample size assuming $M = N \to \infty$, however asymptotic normality will still hold even if $M/N \to p$ as $M, N \to \infty$, for some $p > 0$. It is interesting to note that the asymptotic properties of the proposed estimators of chirp rates β^0, δ^0 will remain the same even if we take $\min\{M, N\} \to \infty$. The asymptotic variance-covariance matrix of $(\alpha^0, \gamma^0, \mu^0)$ will however change depending on the value of p among non-linear parameters.

If we further assume that the errors in (1) are i.i.d. Gaussian distributed random variables, then it can be observed that the proposed estimators of chirp rates parameters β^0 and δ^0 attain Cramer-Rao lower bound (CRLB). CRLB for estimators of other non-linear parameters α^0, γ^0, and μ^0 are $\dfrac{456c\sigma^2}{(A^{0^2}+B^{0^2})}$, $\dfrac{456c\sigma^2}{(A^{0^2}+B^{0^2})}$ and $\dfrac{288c\sigma^2}{(A^{0^2}+B^{0^2})}$, see [16].

4 Simulation Results

Simulation studies done in this paper are divided into three parts. The first part demonstrates the evaluation of finite sample size performance of proposed estimators. The paper considers error general than i.i.d. Gaussian too, so instead of comparing with CRLB, we compare the performance of the proposed estimators with the asymptotically optimal estimators such as LSEs, and ALSEs, and fast but sub-optimal 2D-multilag-HAF estimators. The second part shows the lower computational cost of the proposed estimators in comparison with the LSEs. Finally, the third part exemplifies the ability of the proposed estimators to extract original gray-scale texture from one polluted with noise and reproduce the original texture. We have performed simulations on the complex counterpart of the model (1), see [2] for comparison purposes.

4.1 Finite Sample Performance

To provide a detailed assessment of the performance of proposed estimators, we have chosen sample sizes
$M = N = 20, 40, 60, 80$ and 100. The fixed values of all parameters to obtain complex-valued chirp data are:

$$A^0 = 1, \quad \alpha^0 = 0.4, \quad \beta^0 = 0.1429, \quad \gamma^0 = 0.25, \quad \delta^0 = 0.1250, \quad \mu^0 = 0.1667. \tag{13}$$

Obtained data from the model is then contaminated with noise $X(m,n)$. We consider two distinct noise structures for these simulations. These are:

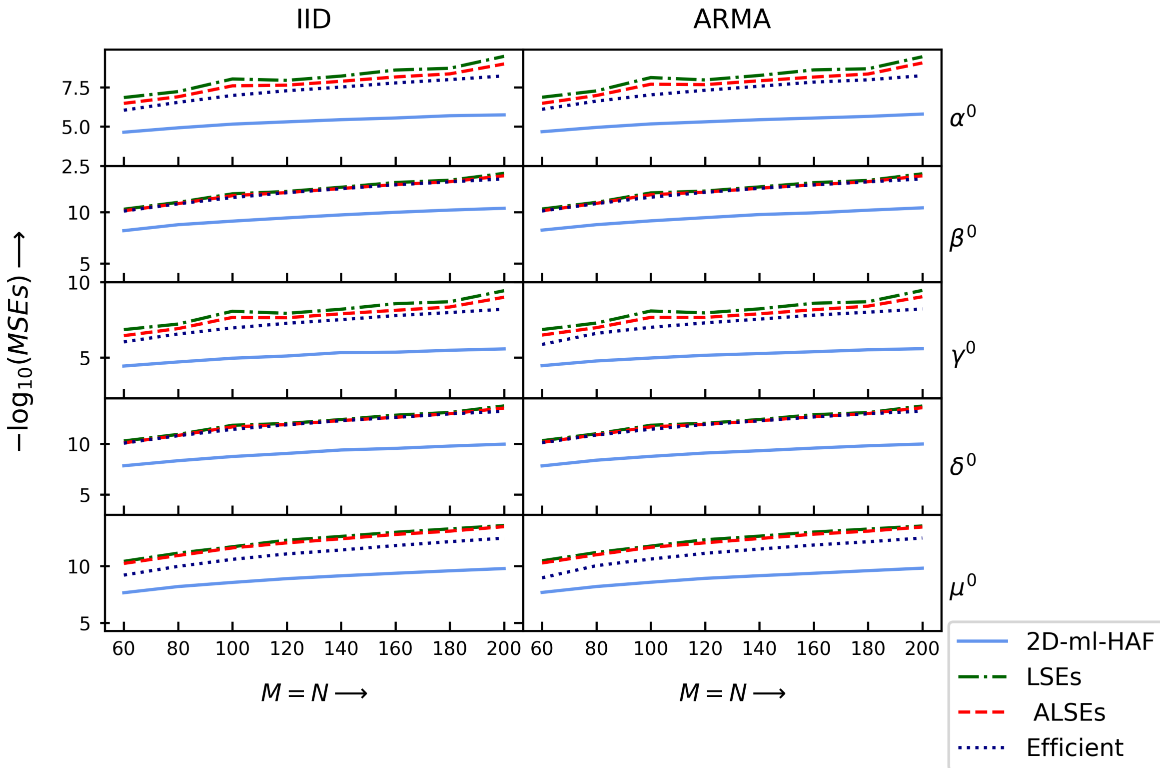
- Independently and identically distributed (i.i.d.) normal errors with mean 0 and variance σ^2;
- Autoregressive–moving-average (ARMA) errors with following representation:

$$\begin{aligned} X(m,n) =& 0.06X(m-1,n-1) - 0.054X(m,n-1) + 0.087X(m-1,n) + \\ & \epsilon(m,n) + .01\epsilon(m-1,n-1) + 0.035\epsilon(m,n-1) + 0.042\epsilon(m-1,n). \end{aligned} \tag{14}$$

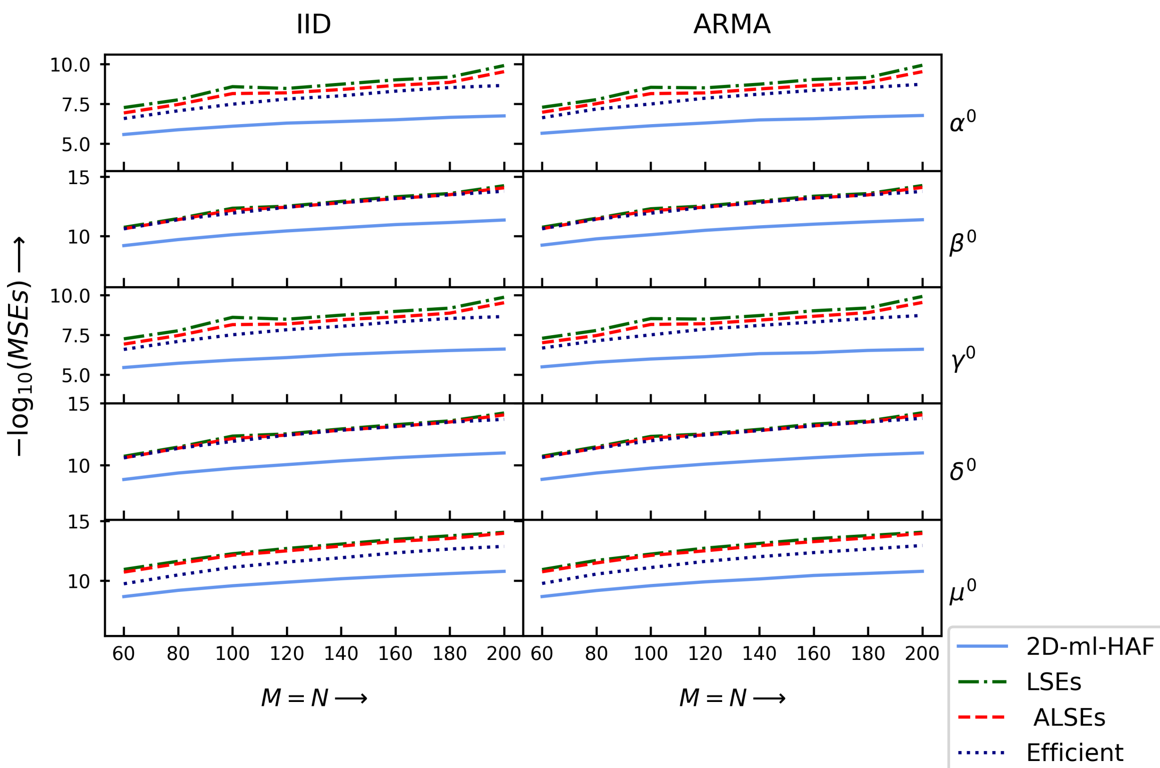
where $\epsilon(m,n)$ is a sequence of i.i.d. Gaussian random variables with mean 0 and variance σ^2.

We have obtained the estimates for 1000 iterations for a fixed sample size $M = N$, type of error i.e. IID or ARMA, and error variance σ^2. The estimators do not possess any explicit closed form expression, so we use Nelder–Mead algorithm (using "optim" function in R software) for optimization of the objective function and to obtain the estimators. Obtained mean square errors (MSEs) are displayed in the figures 1a, 1b, 2a and 2b for four different values of $\sigma = 0.1,0.5,0.9,1$. The MSEs are plotted on negative logarithmic scale. The essence of these simulations results can be summarized as follows:

- MSEs of the proposed estimators tend to decrease as error variance decreases.
- MSEs of the proposed estimators decrease rapidly and become closer to LSEs and ALSEs as sample size increases, which further supports the consistency property of the estimators.
- The obtained MSEs of proposed estimators of β^0, δ^0 are at par with those of LSEs and ALSEs.
- If we compare Figures 1a and 2b, then the gap between MSEs of optimal estimators and the 2D-multilag-HAF estimators increase on increasing noise level, i.e. σ. On the other hand, the proposed estimators do not suffer in such a situation. Their performance stays close to that of LSEs and ALSEs.

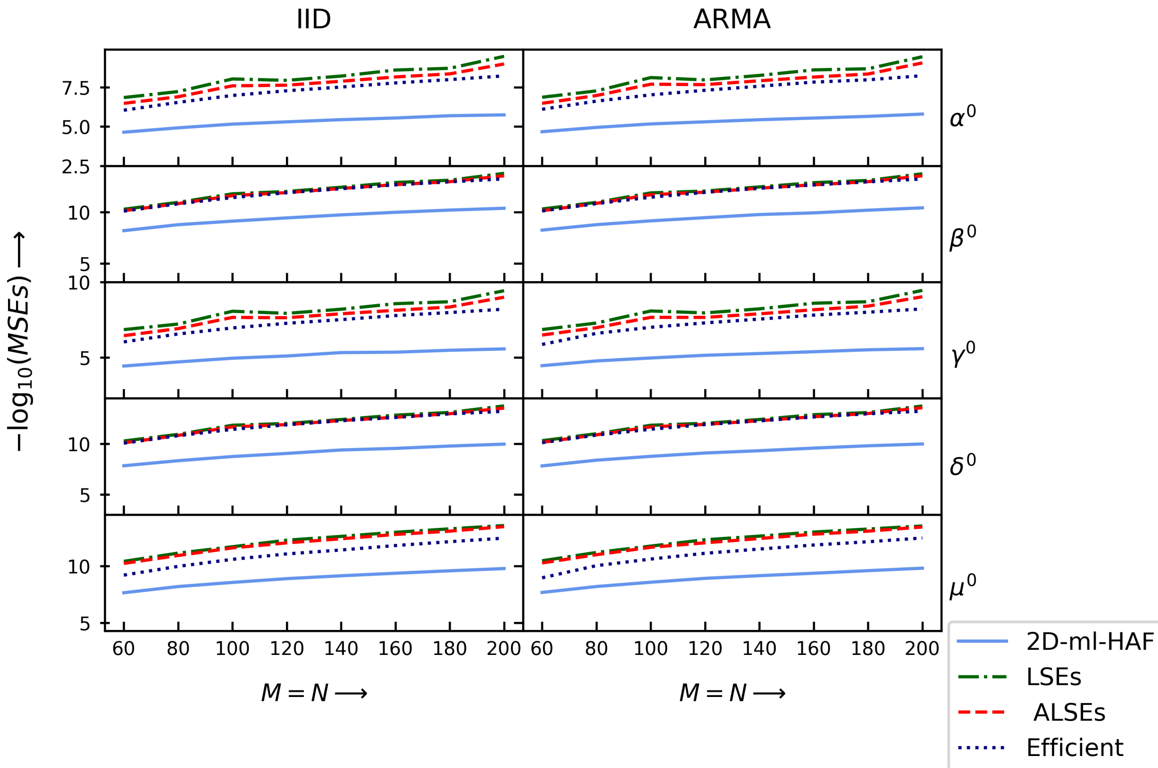


(a) Standard deviation of $\epsilon(m,n)$ is $\sigma = 0.1$

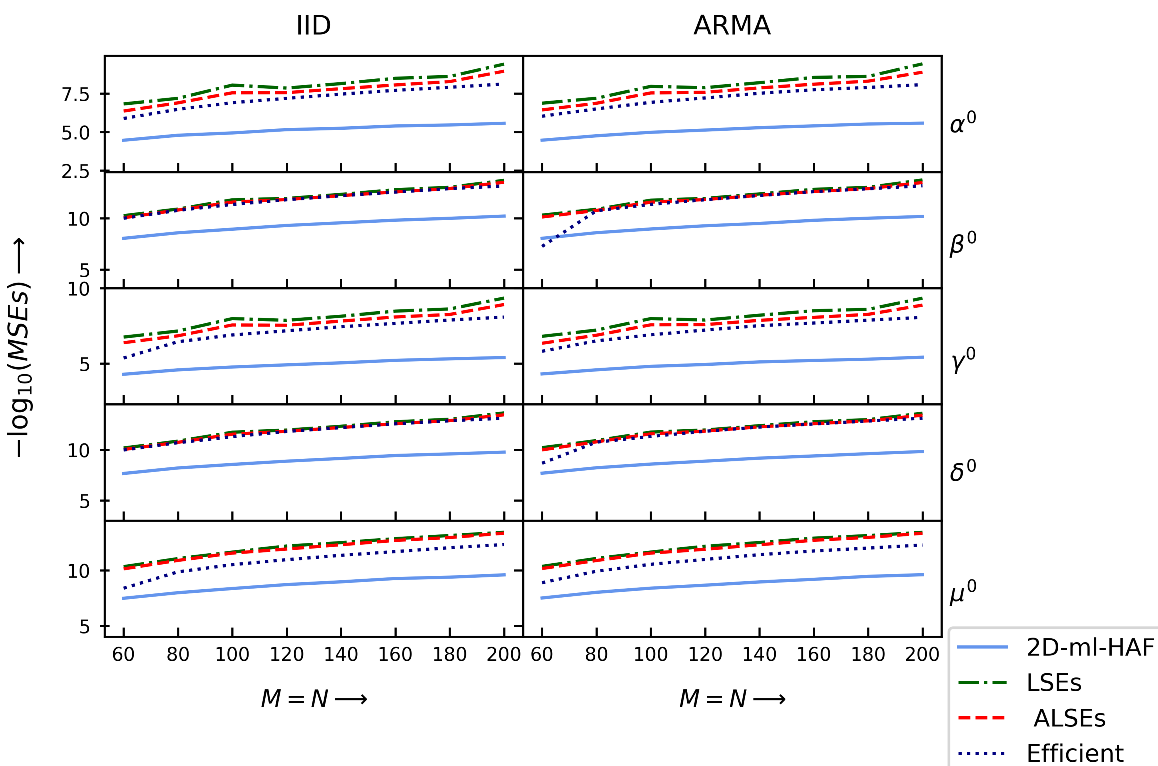


(b) Standard deviation of $\epsilon(m,n)$ is $\sigma = 0.5$

Figure 1: Plots of $-\log(\text{MSEs})$ versus the sample size for estimators of non-linear parameters.



(a) Standard deviation of $\epsilon(m, n)$ is $\sigma = 0.9$



(b) Standard deviation of $\epsilon(m, n)$ is $\sigma = 1$

Figure 2: Plots of $-\log(\text{MSEs})$ versus the sample size for estimators of non-linear parameters.

4.2 Time Comparison

The computational advantage of the proposed estimators over the conventional LSEs is magnificent. In order to compare the two methods, we measure their computational complexities in terms of the number of points in the grid that are needed to find the initial guesses of these estimators. Once we have the precise initial guesses, applying an iterative algorithm like Nelder-Mead is a matter of seconds. "gridSearch" function from the package "NMOF" is used to calculate the initial guesses. We report observed time to get the estimates for fixed sample size and the total number of grid points over which cost function evaluations are required, in table 1. The choice of parameters is same as in (13) along-with i.i.d. normal errors with mean 0 and standard deviation $\sigma = 0.9$. For a fixed sample size $M = N$, the order of computation for LSEs is $M^4N^4 = M^8$, but for the proposed method, the order of computation is $M^3N + N^3M = 2M^4$.

The numerical experiments for comparing time efficiency were performed on a system with processor: Intel(R) Core(TM) i3-5005U CPU @ 2.00GHz 2.00GHz; installed memory(RAM): 4.00 GB; and system type: 64-bit Operating System. Codes were written and run in R version 4.0.4 (2021-02-15) – "Lost Library Book", a free software environment for statistical computing and graphics.

Table 1: Time and number of grid points taken to compute the estimates

Sample Size	Efficient		LSEs	
$M = N$	Time(seconds)	Total No. of grid points	Time(seconds)	Number of grid points
2	0.04400301	12	0.1280069	27
3	0.041	96	0.308017	2048
4	0.06900986	360	4.1012349	30375
5	0.119	960	34.608979	221184
6	0.1730089	2100	198.038328	1071875
7	0.25001502	4032	856.37598	3981312

For the considered machine, it was not feasible to perform grid-search to get LSEs for more than $M = N = 7$. When we plot $M = N$ and the time to get initial guess of LSEs, then it is observed to be linear. So, to get an idea of the time deviation of LSEs with that of proposed estimates at larger sample sizes, we predict time to obtain LSEs based on grid-search by fitting a simple linear regression model between log of sample size and log of time to get LSEs. From the results in Table 2, we can clearly observe the massive time difference of getting proposed estimates and LSEs. For example, if we go for sample size, say $M = N = 40$, then it will take more than 20 years to obtain LSEs using grid-search over the same machine (even if we assume a large amount of RAM in a machine), while it took less than 10 minutes to obtain the proposed estimates.

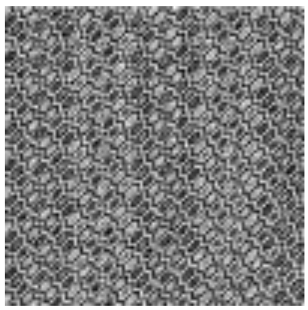
Table 2: Comparing Time efficiency with Predicted time for LSEs

Sample Size	Efficient		LSEs	
$M = N$	Computed Time	Total No. of grid points	Predicted Time	Number of grid points
8	0.53003 sec	7.06E+03	32.16544 min	1.23E+07
9	0.7340422 sec	1.15E+04	1.375421 hr	3.28E+07
10	1.545088 sec	1.78E+04	3.195031 hr	7.86E+07
15	5.8313339 sec	9.41E+04	3.411263 days	2.20E+09
20	26.479515 sec	3.03E+05	34.06977 days	2.29E+10
25	1.015608 min	7.49E+05	203.0514 days	1.40E+11
30	2.443423 min	1.56E+06	2.391816 yr	6.11E+11
35	5.369024 min	2.91E+06	8.208658 yr	2.12E+12
40	8.525054 min	4.99E+06	23.88811 yr	6.22E+12
45	14.53046 min	8.02E+06	61.28852 yr	1.61E+13
50	24.243203 min	1.22E+07	142.37 yr	3.75E+13

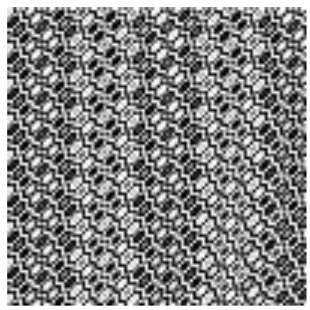
4.3 Texture Pattern Estimation

2D-chirp signals create interesting gray-scale texture patterns. In order to analyze the effectiveness of the proposed estimators for estimating texture patterns accurately, we generate data from the complex counterpart of the model with same set of parameters as in (13). Then real and imaginary part of the obtained data is contaminated independently with i.i.d. normal errors having mean 0 and variance $\sigma^2 = 0.09$. The data matrix obtained is of size 100×100. We analyze this data using the proposed estimator, the optimal LSEs, ALSEs, and also with the 2D-multilag-HAF method. Note that we have used true values as initial guess for obtaining optimal estimators LSEs and ALSEs because of the computational complexity and grid-search method for obtaining 2D-multilag-HAF estimators and proposed estimators.

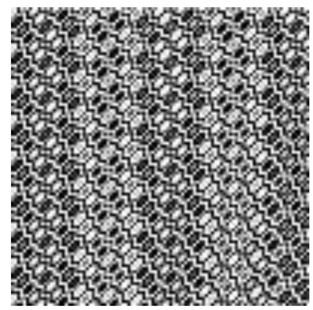
Plugging these estimators in the deterministic part of the model, we get estimated texture patterns as the real part of the reproduced data. We also present real part of the original dataset to compare the original and estimated textures. It is clear from the figures that texture pattern obtained using the proposed method is visually the same as that obtained using optimal LSEs and ALSEs, while the 2D-multilag-HAF estimator gives a slightly different pattern than the original one.



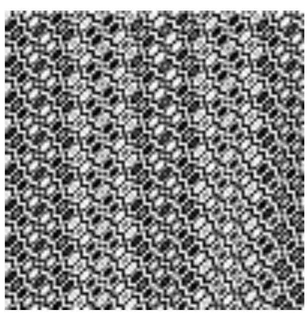
(a) Contaminated Texture



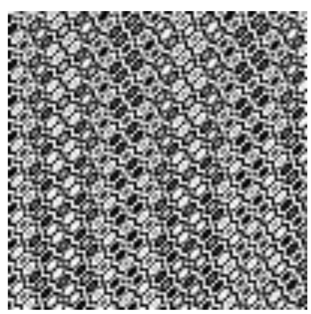
(b) Original Texture



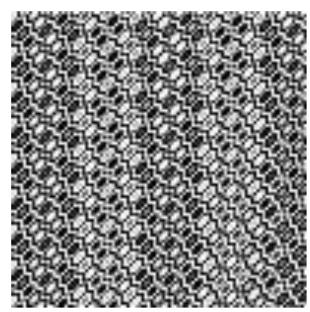
(c) Texture estimated using LSEs



(d) Texture estimated using ALSEs



(e) Texture estimated using 2D-multilag-HAF method



(f) Texture estimated using proposed method

Figure 3: A comparison of estimated textures using LSEs, ALSEs, 2D-multilag-HAF and the proposed estimators.

5 Conclusion

We have presented a computationally efficient method which produces estimators having the same rate of convergence as the LSEs or ALSEs. The key idea is to develop a stratagem by disintegrating the 2D-model into several 1D-chirp models and then perform the optimal estimation. We believe that this idea can lead to several other computationally efficient algorithms which can further be employed for estimating higher order polynomial phase signals' parameters also. Thorough theoretical derivations show that the proposed estimators have the same rates of convergence as that of the optimal estimators, the LSEs and the ALSEs. The proposed estimators are not only asymptotically unbiased but also have an asymptotic normal distribution. Moreover, the estimators converge strongly to the true value of the parameters. Extensive numerical simulations firmly support the theoretical results and also unveil the gigantic gap in time duration for obtaining proposed estimates and the LSEs. In the end, exemplification of a simulated data analysis shows how the proposed estimators can extract original textures effectively.

Appendix A

Proof of Theorem 1: Given the data matrix $\boldsymbol{Y}$, we compute the LSEs of $\alpha^0 + n_0\mu^0$ and β^0 corresponding to n_0^{th} column vector of $\boldsymbol{Y}$. We denote the obtained estimators by $\widehat{\alpha}_{n_0}$ and $\widehat{\beta}_{n_0}$ to emphasize that these depend on n_0. Similarly, for fixed m_0^{th} row of $\boldsymbol{Y}$, we have denoted the LSEs of $\gamma^0 + m_0\mu^0$ and δ^0 by $\widehat{\gamma}_{m_0}$ and $\widehat{\delta}_{m_0}$. Under assumptions that $X(m_0, n_0)$ are stationary (11) and (12), see [18], we have

$$\widehat{\beta}_{n_0} = \beta^0 + o\left(\frac{1}{M^2}\right), \qquad \widehat{\alpha}_{n_0} = \alpha^0 + n_0\mu^0 + o\left(\frac{1}{M}\right), \tag{15}$$

$$\widehat{\gamma}_{m_0} = \gamma^0 + m_0\mu^0 + o\left(\frac{1}{N}\right), \qquad \widehat{\delta}_{m_0} = \delta^0 + o\left(\frac{1}{N^2}\right). \tag{16}$$

The final estimator of β^0 given by

$$\widehat{\beta} = \frac{\sum_{n_0=1}^{N} \widehat{\beta}_{n_0}}{N}$$

is strongly consistent estimate of β^0 which is observed by (15) and also that $\widehat{\beta}_{n_0}$ is strongly consistent for β^0 as $M \to \infty$. For proof, one may refer to [15].

Similarly $\widehat{\delta} = \dfrac{1}{M}\sum_{m_0=1}^{M} \widehat{\delta}_{m_0}$ is strongly consistent estimator of δ^0.

Denote $\boldsymbol{\tau}^\top = \Bigg[\underbrace{o\Big(\frac{1}{M}\Big) \quad \dots \quad o\Big(\frac{1}{M}\Big)}_{N \text{ times}} \quad \underbrace{o\Big(\frac{1}{N}\Big) \quad \dots \quad o\Big(\frac{1}{N}\Big)}_{M \text{ times}}\Bigg]_{1\times(M+N)}$.

We now prove the consistency of the frequency parameter estimators $\widehat{\alpha}, \widehat{\gamma}$ and that of $\widehat{\mu}$, estimator of the interaction term parameter. From (15) and (16), we have the following:

$$\begin{bmatrix} \widehat{\alpha} \\ \widehat{\gamma} \\ \widehat{\mu} \end{bmatrix} = \Big(\boldsymbol{\Gamma}^\top\boldsymbol{\Gamma}\Big)^{-1}\boldsymbol{\Gamma}^\top\boldsymbol{\Lambda} = (\boldsymbol{\Gamma}^\top\boldsymbol{\Gamma})^{-1}\boldsymbol{\Gamma}^\top\left(\boldsymbol{\Gamma}\begin{bmatrix} \alpha^0 \\ \gamma^0 \\ \mu^0 \end{bmatrix} + \boldsymbol{\tau}\right), \tag{17}$$

where $\boldsymbol{\Gamma}^\top = \begin{bmatrix} 1 & 1 & \cdots & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 1 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & N & 1 & 2 & \cdots & M \end{bmatrix}_{3\times(M+N)}$, and $\boldsymbol{\Lambda}^\top = \Big[\widehat{\alpha}_1 \quad \widehat{\alpha}_2 \quad \cdots \quad \widehat{\alpha}_N \quad \widehat{\gamma}_1 \quad \widehat{\gamma}_2 \quad \cdots \quad \widehat{\gamma}_M\Big]$.

This implies that

$$\begin{bmatrix} \widehat{\alpha} \\ \widehat{\gamma} \\ \widehat{\mu} \end{bmatrix} = \begin{bmatrix} \alpha^0 \\ \gamma^0 \\ \mu^0 \end{bmatrix} + (\boldsymbol{\Gamma}^\top\boldsymbol{\Gamma})^{-1}\begin{bmatrix} N \times o\Big(\frac{1}{M}\Big) \\ M \times o\Big(\frac{1}{N}\Big) \\ \frac{N(N+1)}{2} \times o\Big(\frac{1}{M}\Big) + \frac{M(M+1)}{2} \times o\Big(\frac{1}{N}\Big) \end{bmatrix}. \tag{18}$$

Now we look at the matrix $(\boldsymbol{\Gamma}^\top\boldsymbol{\Gamma})^{-1}\begin{bmatrix} N \times o\Big(\frac{1}{M}\Big) \\ M \times o\Big(\frac{1}{N}\Big) \\ \frac{N(N+1)}{2} \times o\Big(\frac{1}{M}\Big) + \frac{M(M+1)}{2} \times o\Big(\frac{1}{N}\Big) \end{bmatrix}$.

The first element of above matrix can be written as $a_1 + a_2 - a_3$, where

$$a_1 = \left(MK - \frac{M^2(M+1)^2}{4}\right)N \times o\Big(\frac{1}{M}\Big)/\Xi,$$
$$a_2 = \frac{MN(M+1)(N+1)}{4}M \times o\Big(\frac{1}{N}\Big)/\Xi,$$
$$a_3 = \frac{MN(N+1)}{2}\left(\frac{N(N+1)}{2} \times o\Big(\frac{1}{M}\Big) + \frac{M(M+1)}{2} \times o\Big(\frac{1}{N}\Big)\right)/\Xi,$$

and

$$\Xi = \frac{MN}{12}\Big(N(N^2-1)+M(M^2-1)\Big), K = \frac{N(N+1)(2N+1)}{6} + \frac{M(M+1)(2M+1)}{6}.$$

Now we look at a_1, a_2 and a_3 individually and compute their limits.

$$a_1 = \Big(MK - \frac{M^2(M+1)^2}{4}\Big)N \times o\Big(\frac{1}{M}\Big)/\Xi$$

$$= N \times \frac{o(1)}{\Xi} \times \Big(\frac{N(N+1)(2N+1)}{6} + \frac{M(M+1)(M-1)}{12}\Big)$$

$$= \frac{o(1)}{M} \times \frac{\Big(2N(N+1)(2N+1) + M(M^2-1)\Big)}{\Big(N(N^2-1)+M(M^2-1)\Big)}.$$

This implies that $a_1 \xrightarrow{a.s.} 0$ as $\min\{M, N\} \to \infty$.
Here

$$a_2 = \frac{MN(M+1)(N+1)}{4}M \times o\Big(\frac{1}{N}\Big)/\Xi$$

$$= o(1) \times \frac{M(M+1)}{\Big(N(N^2-1)+M(M^2-1)\Big)}.$$

This implies that $a_2 \xrightarrow{a.s.} 0$ as $\min\{M, N\} \to \infty$.

$$a_3 = \frac{MN(N+1)}{2}\Big(\frac{N(N+1)}{2} \times o\Big(\frac{1}{M}\Big) + \frac{M(M+1)}{2} \times o\Big(\frac{1}{N}\Big)\Big)/\Xi$$

$$= \frac{(N+1)}{2\Xi}\Big(N^2(N+1) \times o(1) + M^2(M+1) \times o(1)\Big).$$

This implies that $a_3 \xrightarrow{a.s.} 0$ as $\min\{M, N\} \to \infty$. Hence, $\widehat{\alpha}$ is strongly consistent estimate of α^0. Similarly strong consistency of $\widehat{\gamma}$ for γ^0 can be derived.

Now, the third element of $(\mathbf{\Gamma}^\top\mathbf{\Gamma})^{-1}\begin{bmatrix} N\times o\left(\frac{1}{M}\right) \\ M\times o\left(\frac{1}{N}\right) \\ \frac{N(N+1)}{2}\times o\left(\frac{1}{M}\right)+\frac{M(M+1)}{2}\times o\left(\frac{1}{N}\right)\end{bmatrix}$ can be written as $-b_1-b_2+b_3$, where

$$\begin{aligned} b_1 &= \frac{MN(N+1)}{2}\times N\times o\left(\frac{1}{M}\right)/\Xi \\ &= o(1)\frac{N(N+1)}{M\big(N(N^2-1)+M(M^2-1)\big)}. \end{aligned}$$

This implies that $b_1 \xrightarrow{a.s.} 0$ as $\min\{M,N\}\to\infty$.

$$\begin{aligned} b_2 &= \frac{NM(M+1)}{2}\times M\times o\left(\frac{1}{N}\right)/\Xi \\ &= o(1)\times\frac{M(M+1)}{N\big(N(N^2-1)+M(M^2-1)\big)}. \end{aligned}$$

This implies that $b_2 \xrightarrow{a.s.} 0$ as $\min\{M,N\}\to\infty$.

$$\begin{aligned} b_3 &= MN\times\left(\frac{N(N+1)}{2}\times o\left(\frac{1}{M}\right)+\frac{M(M+1)}{2}\times o\left(\frac{1}{N}\right)\right)/\Xi \\ &= \left(\frac{N(N+1)}{\big(N(N^2-1)+M(M^2-1)\big)}\times o\left(\frac{1}{M}\right)+\frac{M(M+1)}{\big(N(N^2-1)+M(M^2-1)\big)}\times o\left(\frac{1}{N}\right)\right). \end{aligned}$$

Hence $b_3 \xrightarrow{a.s.} 0$ as $\min\{M,N\}\to\infty$. Hence, $\widehat{\mu}$ is a strongly consistent estimate of μ^0. The strong consistency of estimators of amplitude parameters follows from the continuity mapping theorem.

Appendix B

Proof of Theorem 2: Suppose $\boldsymbol{\kappa}^\top=(A,B,\alpha,\beta)$ and $\boldsymbol{\kappa}^{0^\top}=(A^0(n_0),B^0(n_0),\alpha^0+n_0\mu^0,\beta^0)$. We define sum of squares as follows:

$$Q_{n_0}(\boldsymbol{\kappa})=\sum_{m_0=1}^{M}\Big(y(m_0,n_0)-A\cos(\alpha m_0+\beta m_0^2)-B\sin(\alpha m_0+\beta m_0^2)\Big)^2.$$

Let $\widehat{\boldsymbol{\kappa}}$ be the minimizer of $Q_{n_0}(\boldsymbol{\kappa})$, then using Taylor Series expansion on the first derivative vector $Q'_{n_0}(\boldsymbol{\kappa})$ around the point $\boldsymbol{\kappa}^0$, we get:

$$\boldsymbol{Q}'_{n_0}(\widehat{\boldsymbol{\kappa}}) - \boldsymbol{Q}'_{n_0}(\boldsymbol{\kappa}^0) = \boldsymbol{Q}''_{n_0}(\breve{\boldsymbol{\kappa}})(\widehat{\boldsymbol{\kappa}} - \boldsymbol{\kappa}^0), \tag{19}$$

where $\breve{\boldsymbol{\kappa}}$ is a point between $\widehat{\boldsymbol{\kappa}}$ and $\boldsymbol{\kappa}^0$. Also $\boldsymbol{Q}'_{n_0}(\widehat{\boldsymbol{\kappa}}) = 0$ as $\widehat{\boldsymbol{\kappa}}$ is LSE of $\boldsymbol{\kappa}^0$.
Let us denote $\boldsymbol{D}_1^{-1} = diag(M^{1/2}, M^{1/2}, M^{3/2}, M^{5/2})$. Then on multiplying $\boldsymbol{D}_1^{-1}$ both sides of equation (19), it gives:

$$-\left[\boldsymbol{D_1}\boldsymbol{Q}''_{n_0}(\breve{\boldsymbol{\kappa}})\boldsymbol{D_1}\right]^{-1}\boldsymbol{\Sigma}_{n_0}\boldsymbol{\Sigma}_{n_0}^{-1}\boldsymbol{D_1}\boldsymbol{Q}'_{n_0}(\boldsymbol{\kappa}^0), = \boldsymbol{D_1}^{-1}(\widehat{\boldsymbol{\kappa}} - \boldsymbol{\kappa}^0) \tag{20}$$

$$\text{where} \lim_{M\to\infty} \left[\boldsymbol{D_1}\boldsymbol{Q}''_{n_0}(\breve{\boldsymbol{\kappa}})\boldsymbol{D_1}\right]^{-1}\boldsymbol{\Sigma}_{n_0} = I_{4\times 4},$$

and

$$\boldsymbol{\Sigma}_{n_0}^{-1} = \frac{2}{A^{0^2} + B^{0^2}}\begin{bmatrix} \dfrac{A^{0^2}(n_0) + 9B^{0^2}(n_0)}{2} & -4A^0(n_0)B^0(n_0) & -18B^0(n_0) & 15B^0(n_0) \\ -4A^0(n_0)B^0(n_0) & \dfrac{9A^{0^2}(n_0) + B^{0^2}(n_0)}{2} & 18A^0(n_0) & -15A^0(n_0) \\ -18B^0(n_0) & 18A^0(n_0) & 96 & -90 \\ 15B^0(n_0) & -15A^0(n_0) & -90 & 90 \end{bmatrix}.$$

The expression of $\boldsymbol{\Sigma}_{n_0}^{-1}$ can be obtained by proof of Theorem 2 shown in [15]. By using equation (20) and above expression of $\boldsymbol{\Sigma}_{n_0}^{-1}$, we get following asymptoically equivalent (a.e.) expression of the third element of vector $\boldsymbol{D_1}^{-1}(\widehat{\boldsymbol{\kappa}} - \boldsymbol{\kappa}^0)$ as:

$$\left[M^{3/2}\big(\widehat{\alpha}_{n_0} - (\alpha^0 + n_0\mu^0)\big)\right] \overset{\text{a.e.}}{=} \frac{4}{A^{0^2} + B^{0^2}}\left[18\sum_{m}\frac{\eta(m_0, n_0)}{\sqrt{M}} - 96\sum_{m_0}\frac{m_0\eta(m_0, n_0)}{M\sqrt{M}} + 90\sum_{m_0}\frac{m_0^2\eta(m_0, n_0)}{M^2\sqrt{M}}\right], \tag{21}$$

where, $\eta(m_0, n_0) = X(m_0, n_0)\big(A^0 \sin\phi(m_0, n_0, \boldsymbol{\xi}^0) - B^0 \cos\phi(m_0, n_0, \boldsymbol{\xi}^0)\big)$,
$\phi(m_0, n_0, \boldsymbol{\xi}^0) = \alpha^0 m_0 + \beta^0 m_0^2 + \gamma^0 n_0 + \delta^0 n_0^2 + \mu^0 m_0 n_0$. We have used these notations for brevity. Similarly expressions of fourth, fifth and sixth element of $\boldsymbol{D_1}^{-1}(\widehat{\boldsymbol{\kappa}} - \boldsymbol{\kappa}^0)$ can be written as in equations (22), (23) and (24) respectively:

$$\left[M^{5/2}\big(\widehat{\beta}_{n_0} - \beta^0\big)\right] \overset{\text{a.e.}}{=} \frac{4}{A^{0^2} + B^{0^2}}\left[-15\sum_{m_0}\frac{\eta(m_0, n_0)}{\sqrt{M}} + 90\sum_{m_0}\frac{m_0\eta(m_0, n_0)}{M\sqrt{M}} - 90\sum_{m_0}\frac{m_0^2\eta(m_0, n_0)}{M^2\sqrt{M}}\right], \tag{22}$$

$$\left[N^{3/2}\big(\widehat{\gamma}_{m_0}-(\gamma^0+m_0\mu^0)\big)\right]\overset{\text{a.e.}}{=}\frac{4}{A^{0^2}+B^{0^2}}\left[18\sum_{n_0}\frac{\eta(m_0,n_0)}{\sqrt{N}}-96\sum_{n_0}\frac{n_0\eta(m_0,n_0)}{N\sqrt{N}}+90\sum_{n_0}\frac{n_0^2\eta(m_0,n_0)}{N^2\sqrt{N}}\right], \tag{23}$$

$$\left[N^{5/2}\big(\widehat{\delta}_{m_0}-\delta^0\big)\right]\overset{\text{a.e.}}{=}\frac{4}{A^{0^2}+B^{0^2}}\left[-15\sum_{n_0}\frac{\eta(m_0,n_0)}{\sqrt{N}}+90\sum_{n_0}\frac{n_0\eta(m_0,n_0)}{N\sqrt{N}}-90\sum_{n_0}\frac{n_0^2\eta(m_0,n_0)}{N^2\sqrt{N}}\right]. \tag{24}$$

From equations (22) and (24) above, we can see that estimators $\widehat{\beta}$ and $\widehat{\delta}$ of β^0 and δ^0 are asymptotically equivalent to the LSEs as they have same asymptotic variances by applying central limit theorem for stationary linear processes, see [10]. So now, remaining is to show asymptotic properties of estimators $(\widehat{\alpha},\widehat{\gamma},\widehat{\mu})^\top$.

The expression of proposed estimators of $(\alpha^0,\gamma^0,\mu^0)^\top$ obtained is:

$$\begin{bmatrix}\widehat{\alpha}\\ \widehat{\gamma}\\ \widehat{\mu}\end{bmatrix}=\begin{bmatrix}\left(MK-\frac{M^2\big(M+1\big)^2}{4}\right)c_\alpha+\frac{MN(M+1)(N+1)}{4}c_\gamma-\frac{MN(N+1)}{2}c_{\alpha\gamma}\\[2ex] \frac{MN(M+1)(N+1)}{4}c_\alpha+\left(KN-\frac{N^2(N+1)^2}{4}\right)c_\gamma-\frac{NM(M+1)}{2}c_{\alpha\gamma}\\[2ex] -\frac{MN(N+1)}{2}c_\alpha-\frac{NM(M+1)}{2}c_\gamma+MNc_{\alpha\gamma}\end{bmatrix}/|\mathbf{\Gamma}^\top\mathbf{\Gamma}|, \tag{25}$$

$$c_\alpha=\sum_{n_0=1}^{N}\widehat{\alpha}_{n_0},\, c_\gamma=\sum_{m_0=1}^{M}\widehat{\gamma}_{m_0},\, c_{\alpha\gamma}=\sum_{n_0=1}^{N}n_0\widehat{\alpha}_{n_0}+\sum_{m_0=1}^{M}m_0\widehat{\gamma}_{m_0},$$

$$K=\frac{N(N+1)(2N+1)}{6}+\frac{M(M+1)(2M+1)}{6},\, \big|\mathbf{\Gamma}^\top\mathbf{\Gamma}\big|=\frac{MN}{12}\Big(N(N^2-1)+M(M^2-1)\Big).$$

From (25), we get:

$$M^{3/2}N^{1/2}\big(\widehat{\alpha}-\alpha^0\big) = \Big(\frac{2N(N+1)(2N+1)+M(M^2-1)}{N(N^2-1)+M(M^2-1)}\Big)\frac{1}{\sqrt{N}}\sum_{n_0=1}^{N} M^{3/2}\big(\widehat{\alpha}_{n_0}-(\alpha^0+n_0\mu^0)\big)$$

$$+\Big(\frac{3M^2(M+1)}{N(N^2-1)+M(M^2-1)}\Big)\frac{1}{\sqrt{M}}\sum_{m_0=1}^{M} N^{3/2}\big(\widehat{\gamma}_{m_0}-(\gamma^0+m_0\mu^0)\big)$$

$$-\Big(\frac{6M^{3/2}N^{1/2}(N+1)}{N(N^2-1)+M(M^2-1)}\Big)\Big[\sum_{n_0=1}^{N} n_0\big(\widehat{\alpha}_{n_0}-(\alpha^0+n_0\mu^0)\big)$$

$$+\sum_{m_0=1}^{M} m_0\big(\widehat{\gamma}_{m_0}-(\gamma^0+m_0\mu^0)\big)\Big].$$

For sufficiently large M and N, we have:

$$M^{3/2}N^{1/2}\big(\widehat{\alpha}-\alpha^0\big) = \Big(\frac{4N^3+M^3}{N^3+M^3}\Big)\frac{1}{\sqrt{N}}\sum_{n_0=1}^{N} M^{3/2}\big(\widehat{\alpha}_{n_0}-(\alpha^0+n_0\mu^0)\big)$$

$$+\Big(\frac{3M^3}{N^3+M^3}\Big)\frac{1}{\sqrt{M}}\sum_{m_0=1}^{M} N^{3/2}\big(\widehat{\gamma}_{m_0}-(\gamma^0+m_0\mu^0)\big)$$

$$-\Big(\frac{6N^3}{N^3+M^3}\Big)\frac{1}{N^{3/2}}\sum_{n_0=1}^{N} n_0 M^{3/2}\big(\widehat{\alpha}_{n_0}-(\alpha^0+n_0\mu^0)\big)$$

$$-\Big(\frac{6M^3}{N^3+M^3}\Big)\frac{1}{M^{3/2}}\sum_{m_0=1}^{M} m_0 N^{3/2}\big(\widehat{\gamma}_{m_0}-(\gamma^0+m_0\mu^0)\big). \tag{26}$$

Asymptotic normality of the estimators for $M=N\to\infty$ with given rates of convergence follows by applying central limit theorem for stationary processes [10].
Next, we will present some results that will help in getting exact expression of the asymptotic variance-covariance matrix.

Using equations (21) and (23), we have the following observations:

1. $AsyVar\left(\frac{1}{2\sqrt{N}}\sum_{n_0=1}^{N} M^{3/2}\big(\widehat{\alpha}_{n_0}-(\alpha^0+n_0\mu^0)\big)\right)=\frac{c\sigma^2 96}{A^{0^2}+B^{0^2}}$,

2. $AsyVar\left(\frac{1}{2\sqrt{M}}\sum_{m_0=1}^{M} N^{3/2}\big(\widehat{\gamma}_{m_0}-(\gamma^0+m_0\mu^0)\big)\right)=\frac{c\sigma^2 96}{A^{0^2}+B^{0^2}}$,

3. $AsyVar\left(\frac{1}{2N^{3/2}}\sum_{n_0=1}^{N} n_0 M^{3/2}\big(\widehat{\alpha}_{n_0}-(\alpha^0+n_0\mu^0)\big)\right)=\frac{c\sigma^2 32}{A^{0^2}+B^{0^2}}$,

4. $AsyVar\left(\frac{1}{2M^{3/2}}\sum_{m_0=1}^{M} m_0 N^{3/2}\big(\widehat{\gamma}_{m_0}-(\gamma^0+m_0\mu^0)\big)\right)=\frac{c\sigma^2 32}{A^{0^2}+B^{0^2}}$,

5. $AsyCovar\left(\frac{1}{2\sqrt{N}}\sum_{n_0=1}^{N} M^{3/2}\big(\widehat{\alpha}_{n_0}-(\alpha^0+n_0\mu^0)\big),\frac{1}{2\sqrt{M}}\sum_{m_0=1}^{M} N^{3/2}\big(\widehat{\gamma}_{m_0}-(\gamma^0+m_0\mu^0)\big)\right)=0$,

6. $AsyCovar\left(\frac{1}{2\sqrt{N}}\sum_{n_0=1}^{N} M^{3/2}\big(\widehat{\alpha}_{n_0}-(\alpha^0+n_0\mu^0)\big),\frac{1}{2N^{3/2}}\sum_{n_0=1}^{N} n_0 M^{3/2}\big(\widehat{\alpha}_{n_0}-(\alpha^0+n_0\mu^0)\big)\right)=$ $\frac{c\sigma^2 48}{A^{0^2}+B^{0^2}}$,

7. $AsyCovar\left(\frac{1}{2\sqrt{N}}\sum_{n_0=1}^{N} M^{3/2}\big(\widehat{\alpha}_{n_0}-(\alpha^0+n_0\mu^0)\big),\frac{1}{2M^{3/2}}\sum_{m_0=1}^{M} m_0 N^{3/2}\big(\widehat{\gamma}_{m_0}-(\gamma^0+m_0\mu^0)\big)\right)=0$,

8. $AsyCovar\left(\frac{1}{2\sqrt{M}}\sum_{m_0=1}^{M} N^{3/2}\big(\widehat{\gamma}_{m_0}-(\gamma^0+m_0\mu^0)\big),\frac{1}{2N^{3/2}}\sum_{n_0=1}^{N} n_0 M^{3/2}\big(\widehat{\alpha}_{n_0}-(\alpha^0+n_0\mu^0)\big)\right)=0$,

9. $AsyCovar\left(\frac{1}{2\sqrt{M}}\sum_{m_0=1}^{M} N^{3/2}\big(\widehat{\gamma}_{m_0}-(\gamma^0+m_0\mu^0)\big),\frac{1}{2M^{3/2}}\sum_{m_0=1}^{M} m_0 N^{3/2}\big(\widehat{\gamma}_{m_0}-(\gamma^0+m_0\mu^0)\big)\right)=$ $\frac{c\sigma^2 48}{A^{0^2}+B^{0^2}}$,

10. $AsyCovar\left(\frac{1}{2N^{3/2}}\sum_{n_0=1}^{N} n_0 M^{3/2}\big(\widehat{\alpha}_{n_0}-(\alpha^0+n_0\mu^0)\big),\frac{1}{2M^{3/2}}\sum_{m_0=1}^{M} m_0 N^{3/2}\big(\widehat{\gamma}_{m_0}-(\gamma^0+m_0\mu^0)\big)\right)=$ $\frac{c\sigma^2}{2(A^{0^2}+B^{0^2})}$.

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

Using the above results in equation (26), we get asymptotic variance-covariance matrix of $\begin{bmatrix} M^{3/2}N^{1/2}(\widehat{\alpha}-\alpha^0) \\ N^{3/2}M^{1/2}(\widehat{\gamma}-\gamma^0) \\ M^{3/2}N^{3/2}(\widehat{\mu}-\mu^0) \end{bmatrix}$ as:

$$\frac{c\sigma^2}{(A^{0^2}+B^{0^2})}\begin{bmatrix} 996 & 612 & -1224 \\ 612 & 996 & -1224 \\ -1224 & -1224 & 2448 \end{bmatrix}. \tag{27}$$

From (21),(22) and (23) it is further observed that:

1. $AsyCovar\Bigg(M^{5/2}N^{1/2}\big(\widehat{\beta}-\beta^0\big), M^{1/2}N^{5/2}\big(\widehat{\delta}-\delta^0\big)\Bigg) = 0,$

2. $AsyCovar\Bigg(M^{5/2}N^{1/2}\big(\widehat{\beta}-\beta^0\big), \dfrac{1}{\sqrt{N}}\sum_{n_0=1}^{N} M^{3/2}\big(\widehat{\alpha}_{n_0}-(\alpha^0+n_0\mu^0)\big)\Bigg) = \dfrac{-\ 360c\sigma^2}{A^{0^2}+B^{0^2}},$

3. $AsyCovar\Bigg(M^{5/2}N^{1/2}\big(\widehat{\beta}-\beta^0\big), \dfrac{1}{\sqrt{M}}\sum_{m_0=1}^{M} N^{3/2}\big(\widehat{\gamma}_{m_0}-(\gamma^0+m_0\mu^0)\big)\Bigg) = 0,$

4. $AsyCovar\Bigg(M^{5/2}N^{1/2}\big(\widehat{\beta}-\beta^0\big), \dfrac{1}{N\sqrt{N}}\sum_{n_0=1}^{N} M^{3/2}n_0\big(\widehat{\alpha}_{n_0}-(\alpha^0+n_0\mu^0)\big)\Bigg) = \dfrac{-\ 180c\sigma^2}{A^{0^2}+B^{0^2}},$

5. $AsyCovar\Bigg(M^{5/2}N^{1/2}\big(\widehat{\beta}-\beta^0\big), \dfrac{1}{M\sqrt{M}}\sum_{m_0=1}^{M} m_0N^{3/2}\big(\widehat{\gamma}_{m_0}-(\gamma^0+m_0\mu^0)\big)\Bigg) = 0.$

Similar results can be derived for $M^{1/2}N^{5/2}\big(\widehat{\delta}-\delta^0\big)$.

By combining all results, the asymptotic variance-covariance matrix of the non-linear parameters' estimators is as follows:

$$\frac{c\sigma^2}{(A^{0^2}+B^{0^2})}\begin{bmatrix} 996 & -360 & 612 & 0 & -1224 \\ -360 & 360 & 0 & 0 & 0 \\ 612 & 0 & 996 & -360 & -1224 \\ 0 & 0 & -360 & 360 & 0 \\ -1224 & 0 & -1224 & 0 & 2448 \end{bmatrix}. \tag{28}$$

To derive the asymptotics of amplitude estimators, please recall that by using taylor series expansion of $\cos\phi(m_0, n_0, \widehat{\boldsymbol{\xi}})$ around the point $\boldsymbol{\xi}^0$, we can write:

$$\cos\widehat{\phi} - \cos\phi^0 = -\sin\breve{\phi}\big(\widehat{\boldsymbol{\xi}} - \boldsymbol{\xi}^0\big)^\top \begin{bmatrix} m_0 \\ m_0^2 \\ n_0 \\ n_0^2 \\ m_0 n_0 \end{bmatrix}.$$

For brevity, we have denoted $\cos\phi(m_0, n_0, \widehat{\boldsymbol{\xi}})$ by $\cos\widehat{\phi}$, $\cos\phi(m_0, n_0, \widehat{\boldsymbol{\xi}})$ by $\cos\phi^0$, and $\sin\breve{\phi}$ by $\sin\phi(m_0, n_0, \breve{\boldsymbol{\xi}})$, where $\breve{\boldsymbol{\xi}}$ is a point lying between $\widehat{\boldsymbol{\xi}}$ and $\boldsymbol{\xi}^0$.

Now consider first element of the following vector,

$$\sqrt{MN}\left(\begin{bmatrix} \frac{2}{MN}\sum_{m_0=1}^{M}\sum_{n_0=1}^{N} y(m_0, n_0)\cos\widehat{\phi} \\ \frac{2}{MN}\sum_{m_0=1}^{M}\sum_{n_0=1}^{N} y(m_0, n_0)\sin\widehat{\phi} \end{bmatrix} - \begin{bmatrix} A^0 \\ B^0 \end{bmatrix}\right),$$

we get:

$$\sqrt{MN}\left(\frac{2}{MN}\sum_{m_0=1}^{M}\sum_{n_0=1}^{N} y(m_0, n_0)\cos\phi^0 - \frac{2}{MN}\sum_{m_0=1}^{M}\sum_{n_0=1}^{N} y(m_0, n_0)\sin\breve{\phi}R(m_0, n_0) - A^0\right), \tag{29}$$

where $R(m_0, n_0) = \big(\widehat{\boldsymbol{\xi}} - \boldsymbol{\xi}^0\big)^\top \begin{bmatrix} m_0 \\ m_0^2 \\ n_0 \\ n_0^2 \\ m_0 n_0 \end{bmatrix}$, and second element of the above amplitude vector can be written as:

$$\sqrt{MN}\left(\frac{2}{MN}\sum_{m_0=1}^{M}\sum_{n_0=1}^{N} y(m_0, n_0)\sin\phi^0 + \frac{2}{MN}\sum_{m_0=1}^{M}\sum_{n_0=1}^{N} y(m_0, n_0)\sin\breve{\phi}R(m_0, n_0) - B^0\right). \tag{30}$$

Now let us look at the first and the last term of equation (29) and putting value of $y(m_0, n_0)$ from the model (1),

$$\sqrt{MN}\Bigg(\frac{2}{MN}\sum_{m_0=1}^{M}\sum_{n_0=1}^{N} y(m_0,n_0)\cos\phi^0 - A^0\Bigg)$$

$$= \sqrt{MN}\Bigg(\frac{2}{MN}\sum_{m_0=1}^{M}\sum_{n_0=1}^{N} A^0\cos^2\phi^0 - A^0 + \frac{2}{MN}\sum_{m_0=1}^{M}\sum_{n_0=1}^{N} B^0\sin\phi^0\cos\phi^0 + \frac{2}{MN}\sum_{m_0=1}^{M}\sum_{n_0=1}^{N} X(m_0,n_0)\cos\phi^0\Bigg)$$

$$\stackrel{\text{a.e.}}{=} \frac{2}{\sqrt{MN}}\sum_{m_0=1}^{M}\sum_{n_0=1}^{N} X(m_0,n_0)\cos\phi^0. \qquad (31)$$

The above result has been obtained from a famous number theory conjecture by [17]. The interested reader is referred to [21] for more details.

In second term of (29), $R(m_0, n_0)$ is a sum of five terms, so now consider first term of $\frac{2}{MN}\sum_{m_0=1}^{M}\sum_{n_0=1}^{N} y(m_0,n_0)\sin\breve{\phi}R(m_0,n_0)$,

$$\frac{2}{\sqrt{MN}}\sum_{m_0=1}^{M}\sum_{n_0=1}^{N} y(m_0,n_0)\sin\breve{\phi}m_0(\widehat{\alpha}-\alpha^0)$$

$$= \frac{2}{\sqrt{MN}}\sum_{m_0=1}^{M}\sum_{n_0=1}^{N} A^0\cos\phi^0\sin\breve{\phi}m_0(\widehat{\alpha}-\alpha^0) + \frac{2}{\sqrt{MN}}\sum_{m_0=1}^{M}\sum_{n_0=1}^{N} B^0\sin\phi^0\sin\breve{\phi}m_0(\widehat{\alpha}-\alpha^0)$$

$$+ \frac{2}{\sqrt{MN}}\sum_{m_0=1}^{M}\sum_{n_0=1}^{N} X(m_0,n_0)\sin\breve{\phi}m_0(\widehat{\alpha}-\alpha^0)$$

$$= \Bigg(\frac{2}{M^2N}\sum_{m_0=1}^{M}\sum_{n_0=1}^{N} A^0\cos\phi^0\sin\breve{\phi}m_0\Bigg)M\sqrt{MN}(\widehat{\alpha}-\alpha^0) +$$

$$\Bigg(\frac{2}{M^2N}\sum_{m_0=1}^{M}\sum_{n_0=1}^{N} B^0\sin\phi^0\sin\breve{\phi}m_0\Bigg)M\sqrt{MN}(\widehat{\alpha}-\alpha^0) +$$

$$\Bigg(\frac{2}{M^2N}\sum_{m_0=1}^{M}\sum_{n_0=1}^{N} X(m_0,n_0)\sin\breve{\phi}m_0\Bigg)M\sqrt{MN}(\widehat{\alpha}-\alpha^0)$$

$$\stackrel{\text{a.e.}}{=} \Bigg(\frac{2}{M^2N}\sum_{m_0=1}^{M}\sum_{n_0=1}^{N} B^0\sin\phi^0\sin\breve{\phi}m_0\Bigg)M\sqrt{MN}(\widehat{\alpha}-\alpha^0), \qquad \text{by using Proposition 1 of [16].}$$

So, finding the asymptotic distribution of (29) boils down to finding asymptotic distribution of

$$\frac{2}{\sqrt{MN}}\sum_{m_0=1}^{M}\sum_{n_0=1}^{N} X(m_0,n_0)\cos\phi^0 - \Bigg(\frac{2}{\sqrt{MN}}\sum_{m_0=1}^{M}\sum_{n_0=1}^{N} B^0 \sin\phi^0 \sin\breve{\phi}\Bigg)R(m_0,n_0),$$

which is further asymptotically equivalent to:

$$\frac{2}{\sqrt{MN}}\sum_{m_0=1}^{M}\sum_{n_0=1}^{N} X(m_0,n_0)\cos\phi^0 - B^0\Bigg(\frac{1}{2}M^{3/2}N^{1/2}(\widehat{\alpha}-\alpha^0)+\frac{1}{3}M^{5/2}N^{1/2}(\widehat{\beta}-\beta^0)$$
$$+\frac{1}{2}N^{3/2}M^{1/2}(\widehat{\gamma}-\gamma^0)+\frac{1}{3}N^{5/2}M^{1/2}(\widehat{\delta}-\delta^0)+\frac{1}{4}M^{3/2}N^{3/2}(\widehat{\mu}-\mu^0)\Bigg). \tag{32}$$

Asymptotic normality of amplitude estimators is thus proved by equation (32). Now we need to derive the expression of their asymptotic variances.

After lengthy calculations, we get the asymptotic variance of $\widehat{A}$ as follows:

$$\frac{c\sigma^2}{(A^{0^2}+B^{0^2})}(2A^{0^2}+187B^{0^2}).$$

Similarly by calculating other terms too, we get complete variance co-variance matrix $\boldsymbol{\Sigma}$ as mentioned in Theorem 2.

Hence the result.

References

[1] Barbarossa, S., Scaglione, A., & Giannakis, G. B. (1998). Product high-order ambiguity function for multicomponent polynomial-phase signal modeling. In IEEE Transactions on Signal Processing, vol. 46, no. 3, pp. 691-708.

[2] Barbarossa, S., Di Lorenzo, P., & Vecchiarelli, P. (2014). Parameter estimation of 2D multi-component polynomial phase signals: an application to SAR imaging of moving targets. IEEE Transactions on Signal Processing, 62(17), 4375-4389.

[3] 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.

[4] Djurović, I., & Stanković, L. (2014). Quasi-maximum-likelihood estimator of polynomial phase signals. IET Signal Processing, 8(4), 347-359.

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

[6] Francos, J. M., & 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.

[7] Friedlander, B., & Francos, J. M. (1996). Model based phase unwrapping of 2-D signals. IEEE Transactions on Signal Processing, 44(12), 2999-3007.

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

[9] 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.

[10] Fuller, W. A. (1996). Introduction to Statistical Time Series, second Edition, John Wiley and Sons, New York.

[11] 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.

[12] Grover, R., Kundu, D., & Mitra, A. (2021). An efficient methodology to estimate the parameters of a two-dimensional Chirp signal model. Multidimensional Systems and Signal Processing, 32(1), 49-75.

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

[14] Lahiri, A., Kundu, D., & Mitra, A. (2013). Efficient algorithm for estimating the parameters of two dimensional Chirp signal. Sankhya B, 75(1), 65-89.

[15] Lahiri, A., Kundu, D., & Mitra, A. (2015). Estimating the parameters of multiple chirp signals. Journal of Multivariate Analysis, 139, 189-206.

[16] Lahiri, A., & Kundu, D. (2017). On parameter estimation of two-dimensional polynomial phase signal model. Statistica Sinica, 1779-1792.

[17] Montgomery, H. L. (1994). Ten lectures on the interface between analytic number theory and harmonic analysis (No. 84). American Mathematical Soc..

[18] Nandi, S., & Kundu, D. (2004). Asymptotic properties of the least squares estimators of the parameters of the chirp signals. Annals of the Institute of Statistical Mathematics, 56(3), 529-544.

[19] O'shea, P. (2002). A new technique for instantaneous frequency rate estimation. IEEE Signal Processing Letters, 9(8), 251-252.

[20] Peleg, S., & Porat, B. (1991). Estimation and classification of polynomial-phase signals. IEEE Transactions on Information Theory, 37(2), 422-430.

[21] Shukla, A., Grover, R., Kundu, D., & Mitra, A. (2021). Approximate Least Squares Estimators of a Two-Dimensional Chirp Model. Submitted in JMVA.

[22] Wu, Y., So, H. C., & Liu, H. (2008). Subspace-based algorithm for parameter estimation of polynomial phase signals. IEEE Transactions on Signal Processing, 56(10), 4977-4983.

[23] Zhang, K., Wang, S., & 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.

[24] Zhang, Y., Mobasseri, B. G., Dogahe, B. M., & Amin, M. G. (2010). Image-adaptive watermarking using 2D Chirps. Signal, Image and Video Processing, 4(1), 105-121.