



A computationally efficient algorithm to estimate the parameters of a two-dimensional chirp model with the product term

Abhinek Shukla¹ · Rhythm Grover² · Debasis Kundu¹ · Amit Mitra¹

Received: 8 April 2022 / Revised: 13 April 2023 / Accepted: 23 April 2023 /
Published online: 11 May 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract

Chirp signal models and their generalizations have been used to model many natural and man-made phenomena in signal processing and time series literature. In recent times, several methods have been proposed for parameter estimation of these models. However, these methods are either statistically sub-optimal or computationally burdensome, especially for two dimensional chirp models. In this paper, we consider the problem of parameter estimation of two dimensional chirp models and propose a computationally efficient estimator and establish asymptotic theoretical properties of the proposed estimators. Moreover, the proposed estimators are observed to have the same rates of convergence as the least squares estimators. Further, the proposed estimators of chirp rate parameters are shown to be asymptotically optimal. Extensive and detailed numerical simulations are conducted, which support the theoretical results of the proposed estimators.

Keywords 2D chirp model · Least squares estimators · Stationary linear process · Consistency · Asymptotic normality.

1 Introduction

The parameter estimation problem for two dimensional (2D) chirp models 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 e.g., Francos & Friedlander, 1995, 1998, 1999 and the references cited therein). 2D chirp signals have also been used to model Newton's rings (Guo & Li, 2018).

✉ Abhinek Shukla
abhushukla@gmail.com

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, Guwahati, Assam 781039, India

These rings are predominantly used in testing spherical and flattening optical surface and curvature radius measurement. 2D chirp signals have been used as a spreading function or base for digital watermarking (Stankovic et al., 2001), which is also helpful in data security, medical safety, fingerprinting, and observing content manipulations, see (Zhang et al., 2010). 2D chirps with linear frequency modulation have been extensively used because of their linear feature in time-frequency domain. These models have been employed in many applications due to other important properties also, e.g., spectral shaping of the watermark by choosing appropriately the frequency and frequency rate parameters, which allows minimum overlap of the spreading function with the image data. 2D chirps also allow adaptive watermarking for enhanced robustness to stationary filtering attacks, i.e., filtering whose parameters do not change over an image, see (Stankovic et al., 2001).

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 Friedlander and Francos (1995) as an extension of the polynomial phase transform proposed in Peleg and Porat (1991). Several works Friedlander and Francos (1996), (1998) and (1999) utilized PD operator to develop computationally efficient algorithms for estimating similar polynomial phase signals. Cubic phase function (CPF) proposed in O’Shea (2002), was extended in Zhang et al. (2008) for similar 2D chirp signal modeling. Further, CPF was utilized to estimate 2D cubic phase signal using genetic algorithm in Djurović et al. (2010). Consistency and asymptotic normality of LSEs for a general 2D polynomial phase signal (PPS) model have been derived in Lahiri and Kundu (2017). A finite step computationally efficient procedure for a similar 2D chirp signal model proposed in Lahiri et al. (2013) was proved asymptotically equivalent to LSEs. Quasi-Maximum Likelihood (QML) algorithm Djurović and Stanković 2014 proposed for 1D PPS, was generalized for 2D PPS in Djurović (2017). Further approximate least squares estimators (ALSEs) proposed in Grover and Kundu (2018) have been proved to be asymptotically equivalent to LSEs. An efficient estimation procedure based on fixed dimension technique, presented in Grover et al. (2021) was shown to be asymptotically equivalent to the optimal LSEs, for a 2D chirp model without the product term mn , e.g., compare model (1) and the one considered in Grover et al. (2021).

Estimators based on phase differencing strategies or high order ambiguity function (HAF) or some of their modifications are computationally easier to obtain. However, the performance of estimation deteriorates below a relatively high signal-to-noise ratio (SNR) threshold and are sub-optimal. Methods that use PD in the steps of estimation, usually estimate coefficients of the 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 accumulates and affects estimation accuracy of lower degree coefficients quite seriously. For more details, one can refer to Barbarossa et al. (1998), Djurović (2017) and Wu et al. (2008). Till date, there is no detailed study of the theoretical properties of the estimators CPF and QML, such as strong consistency and asymptotic normality. Recently, optimal estimators for a simpler 2D chirp model without the interaction term have been developed in Grover et al. (2021). However, the results in Grover et al. (2021) cannot be generalized directly for the underlying model (1). It may be noted that the model considered in this paper is more general, as it takes into account the interaction term $\mu^0 mn$. Due to the presence of this interaction term coefficient μ^0 , the estimation becomes more difficult as the estimators of α^0 and γ^0 are no longer independent (as in the case for Grover et al., 2021), and hence making their computation as well as the study of theoretical analysis becomes more challenging. The problem becomes more complicated under the assumption of general stationary linear process error assumption.

The main contributions of this paper are; providing a computationally efficient algorithm to estimate the parameters of the model defined in (1), and further establishing theoretical asymptotic properties of the proposed estimators. The proposed algorithm is motivated by the fact that a 2D chirp model with five non-linear parameters can be viewed through a number of 1D chirp models with two non-linear parameters, and hence computational complexity of estimators for the 2D models can be reduced.

The key attributes of the proposed method are that it is computationally faster than the conventional optimal methods such as LSEs, maximum likelihood estimators, or ALSEs and at the same time, having desirable statistical properties such as, attaining the same rates of convergence as the optimal LSEs. In fact the proposed estimators of the chirp rate parameters have the same asymptotic variance as that of the traditional LSEs, and hence are asymptotically optimal.

The rest of the paper is organised as follows: the mathematical model and the methodology to obtain the proposed estimators is presented in Sect. 2. The model assumptions and the asymptotic theoretical results are given in Sect. 3. In Sect. 4, the finite sample performance of the proposed estimators is demonstrated through simulation studies. In this section, a comparison of the performance of the proposed estimators with the state-of-the-art methods such as the least squares method, approximate least squares method, and 2D multilag HAF method is also presented. Finally, Sect. 5 concludes the paper, followed by detailed proofs in appendices.

2 Estimation methodology

This paper addresses the problem of parameter estimation of a 2D chirp signal model defined as follows:

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

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 parameters, β^0, δ^0 are frequency rates or chirp rates, and μ^0 is the coefficient of product term. $\xi^0 = (\alpha^0, \beta^0, \gamma^0, \delta^0, \mu^0)^T$ 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. The product term mn in such chirp models (1), is an important characteristic of numerous measurement interferometric signals, radar signal returns and detecting digital watermarking.

Now we discuss the proposed method of estimation. Let the data matrix for model (1) be denoted as

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

The proposed algorithm uses the fact that for each fixed column (or row) of \mathbf{Y} , the 2D chirp model breaks down to a cascade of 1D chirp models.

Realise that if we fix one dimension say $n = n_0$ in (1), then the 2D chirp can be seen as 1-D chirp for $m = 1, 2, \dots, M$, as follows:

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

where, $A^0(n_0) = A^0 \cos(\gamma^0 n_0 + \delta^0 n_0^2) + B^0 \sin(\gamma^0 n_0 + \delta^0 n_0^2)$,
 $B^0(n_0) = -A^0 \sin(\gamma^0 n_0 + \delta^0 n_0^2) + B^0 \cos(\gamma^0 n_0 + \delta^0 n_0^2)$. (2)

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

$$\begin{aligned}
 y(m_0, n) &= \tilde{A}^0(m_0) \cos((\gamma^0 + m_0\mu^0)n + \delta^0 n^2) + \tilde{B}^0(m_0) \sin((\gamma^0 + m_0\mu^0)n + \delta^0 n^2) \\
 &\quad + X(m_0, n).
 \end{aligned}$$
 (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 respectively. Similarly, Eq. (3) represents 1-D chirp signal model with $\gamma^0 + m_0\mu^0$ and δ^0 as the frequency and frequency rate parameters respectively.

Hence, 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. Therefore this procedure reduces computational burden to estimate model parameters drastically. Further suppose column vector \mathbf{Y}_{Mn_0} denotes the n_0^{th} column of data matrix \mathbf{Y} and column vector \mathbf{Y}_{m_0N} denotes the transpose of m_0^{th} row of data matrix \mathbf{Y} . Define $Z_k(\alpha_1, \alpha_2)$ matrix as

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

We need to estimate the model parameters of (2) which is a 1D chirp model, so we use LSEs to estimate the parameters. We obtain LSEs of non-linear parameters in model (2) for a fixed n_0 by defining following reduced sum of squares, see (Grover et al., 2021):

$$R_{Mn_0}(\alpha_1, \alpha_2) = \mathbf{Y}_{Mn_0}^T \left(I_{M \times M} - P_{Z_M}(\alpha_1, \alpha_2) \right) \mathbf{Y}_{Mn_0},$$
 (5)

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

Then,

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

is the proposed estimator of $(\alpha^0 + n_0\mu^0, \beta^0)^T$ based on minimizing the sum of squares corresponding to n_0^{th} column of the data matrix \mathbf{Y} . Similarly, we can obtain LSEs of parameters in model (3) by defining

$$R_{m_0N}(\alpha_1, \alpha_2) = \mathbf{Y}_{m_0N}^T \left(I_{N \times N} - P_{Z_N}(\alpha_1, \alpha_2) \right) \mathbf{Y}_{m_0N},$$
 (7)

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

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

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^{th} row of data matrix Y .

We observe that for each fixed column, we get an estimate of the same chirp rate parameter β^0 in (6), and also an 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. It is important to note that the linearity of parameters of 1D-chirp models plays a 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 $\mathbf{\Gamma}^\top = \begin{bmatrix} 1 & 1 & \dots & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 1 & 1 & \dots & 1 \\ 1 & 2 & \dots & N & 1 & 2 & \dots & M \end{bmatrix}$, and $\mathbf{\Lambda}^\top = [\widehat{\alpha}_1 \ \widehat{\alpha}_2 \ \dots \ \widehat{\alpha}_N \ \widehat{\gamma}_1 \ \widehat{\gamma}_2 \ \dots \ \widehat{\gamma}_M]$.

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 $(\mathbf{\Gamma}^\top \mathbf{\Gamma})^{-1} \mathbf{\Gamma}^\top \mathbf{\Lambda}$.

Step-2. The estimates 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.

Step-3. After getting estimates of non-linear parameters, $\widehat{\xi} = (\widehat{\alpha}, \widehat{\beta}, \widehat{\gamma}, \widehat{\delta}, \widehat{\mu})^\top$, the amplitude parameter estimates can be provided 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 required to derive the theoretical asymptotic properties 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 $\theta^0 = (A^0, B^0, \alpha^0, \beta^0, \gamma^0, \delta^0, \mu^0)^\top$ is an interior point of parameter space Θ , where $\Theta = (-M, M) \times (-M, M) \times [0, 2\pi] \times [0, \pi/2] \times [0, 2\pi] \times [0, \pi/2] \times [0, 2\pi]$, and $A^0 + B^0 > 0$, for some $M > 0$.

Assumption 1 puts the model under a very general set-up of noise contamination as it includes 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 θ^0 is strongly consistent, i.e.,

$$\hat{\theta} \xrightarrow{a.s.} \theta^0 \quad \text{as } \min\{M, N\} \rightarrow \infty.$$

Proof Please see Appendix A for the proof. □

Theorem 2 Under assumptions 1 and 2, the proposed estimators of θ^0 is asymptotically normally distributed.

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

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

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

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

Proof Please see Appendix B for the proof. □

Although Theorem 2 has been proved for the increasing sample size assuming $M = N \rightarrow \infty$, however asymptotic normality will still hold even if $M/N \rightarrow p$ as $M, N \rightarrow \infty$, 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 $\min\{M, N\} \rightarrow \infty$. 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 , asymptotically attain Cramer-Rao lower bound (CRLB). CRLB for estimators of other non-linear parameters α^0, γ^0 , and μ^0 are $\frac{456c\sigma^2}{(A^{0^2} + B^{0^2})}$, $\frac{456c\sigma^2}{(A^{0^2} + B^{0^2})}$ and $\frac{288c\sigma^2}{(A^{0^2} + B^{0^2})}$, see (Lahiri & Kundu, 2017).

4 Simulation results

Simulation studies performed in this paper are divided into three parts. The first part demonstrates the evaluation of finite sample size performance of proposed estimators. 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 as compared to the LSEs. Finally, the third part exemplifies the ability of the proposed estimators to extract original gray-scale texture from one contaminated with noise and reproduce the original texture. We have performed simulations on the complex counterpart of the model (1), (see Barbarossa, 2014) 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 our 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:

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

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 replications for a fixed sample size $M = N$, under a particular error structure with fixed σ^2 . The estimators do not have any explicit closed

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

Sample size $M = N$	Efficient		LSEs	
	Time (seconds)	Total no. of grid points	Time (seconds)	Number of grid points
2	0.044	12	0.128	27
3	0.041	96	0.308	2048
4	0.069	360	4.101	30375
5	0.119	960	34.609	221184
6	0.173	2100	198.038	1071875
7	0.250	4032	856.375	3981312

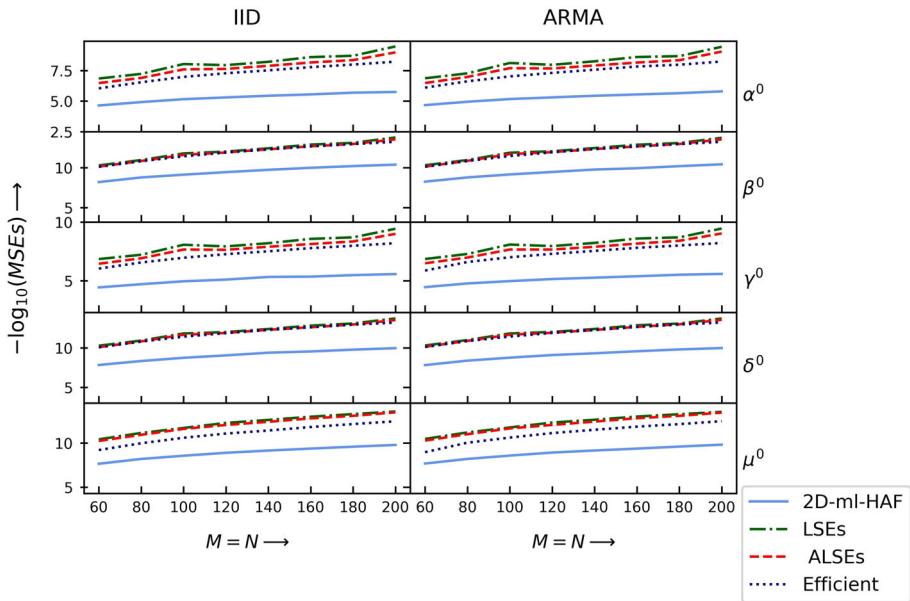
form expression, so we use Nelder-Mead algorithm (using “optim” function in R software) for optimization of the objective function and to obtain the estimators. Mean square errors (MSEs) obtained under 1000 replications are displayed in the Figs. 1a, b, 2a and b for four different values of $\sigma = 0.1, 0.5, 0.9, 1$. The MSEs are plotted on negative logarithmic scale. The findings of these simulation results can be summarized as follows:

- MSEs of the proposed estimators decrease rapidly as sample size increases, which supports the consistency property of the estimators. Further, as sample size increases, the gap between the MSEs of the proposed estimators and LSEs decreases.
- The obtained MSEs of proposed estimators of β^0 and δ^0 , are at par with those of LSEs and ALSEs.

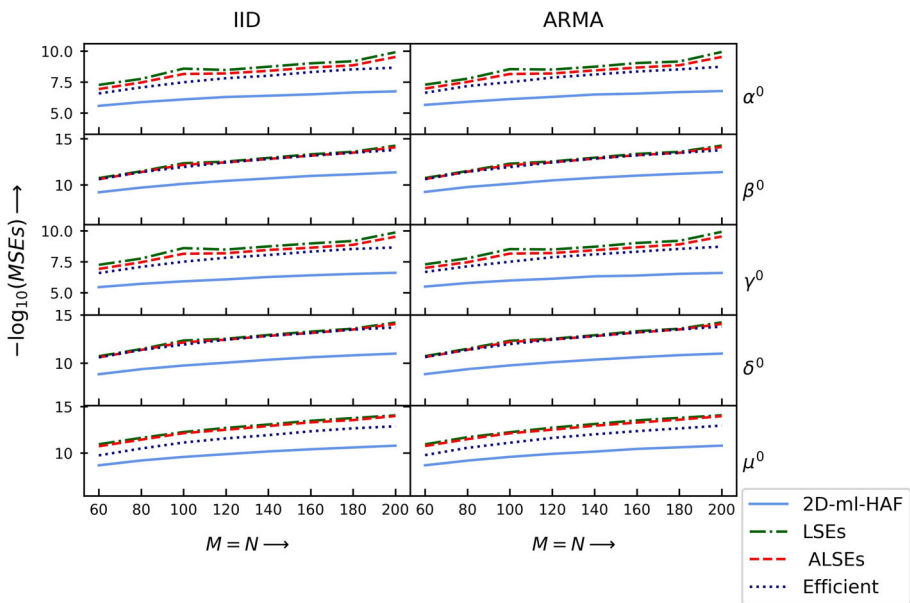
4.2 Time comparison

The computational advantage of the proposed estimators over the conventional LSEs is quite significant. 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 a 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 $M^4 N^4 = M^8$, but for the proposed method, the order of computation is $M^3 N + N^3 M = 2M^4$. 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.

For the considered machine, it was not feasible to perform grid-search to get LSEs for more than $M = N = 7$. When we plot logarithm sample size $M (= N)$ against the logarithm time to compute initial guess for 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,

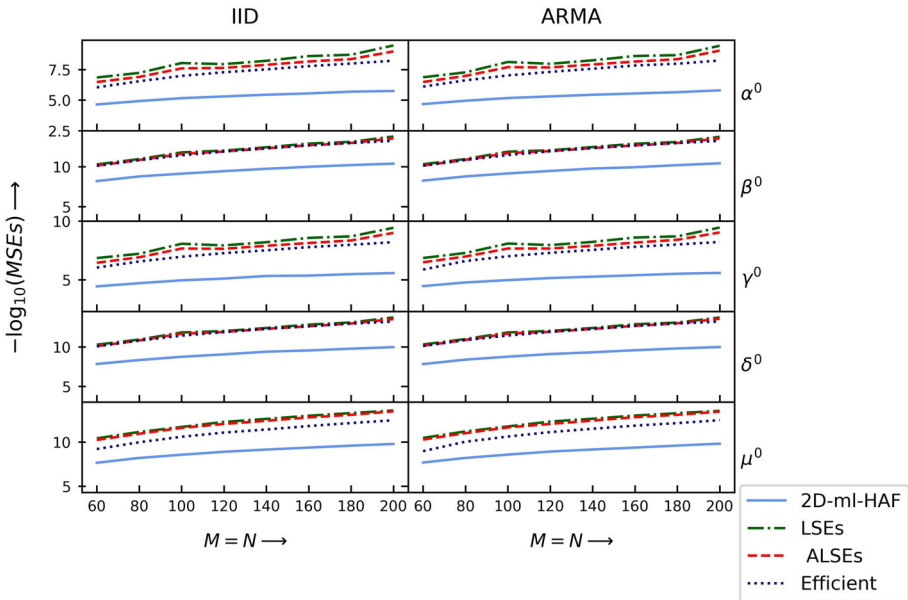


(a) Standard deviation of $\epsilon(m, n)$ is $\sigma = 0.1$

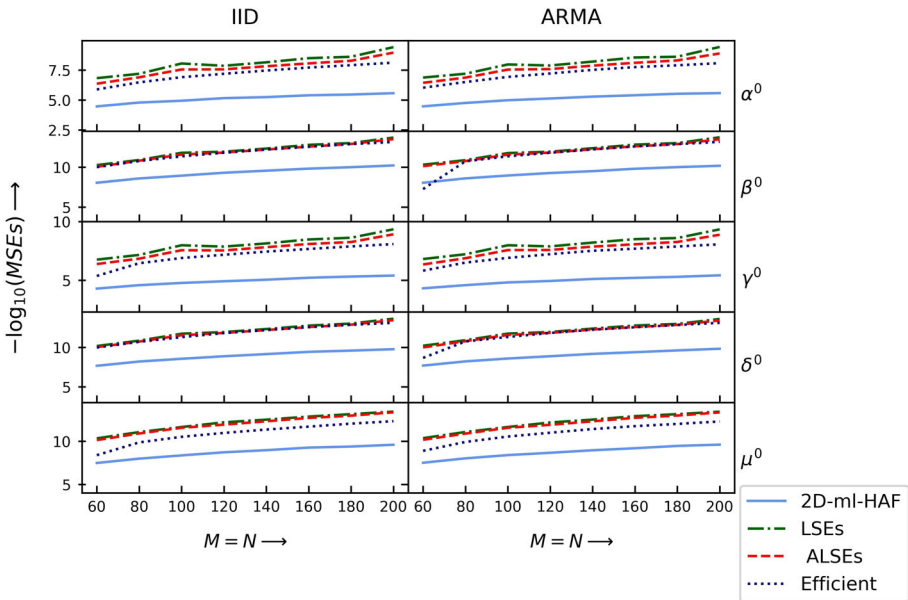


(b) Standard deviation of $\epsilon(m, n)$ is $\sigma = 0.5$

Fig. 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 $\sigma = 1$

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

Table 2 Comparing Time efficiency of proposed estimators with Predicted time for LSEs

Sample size $M = N$	Efficient Computed time	Total no. of grid points	LSEs	
			Predicted time	Number of grid points
8	0.530 sec	7.06E+03	32.165 min	1.23E+07
9	0.734 sec	1.15E+04	1.375 hr	3.28E+07
10	1.545 sec	1.78E+04	3.195 hr	7.86E+07
15	5.831 sec	9.41E+04	3.411 days	2.20E+09
20	26.479 sec	3.03E+05	34.070 days	2.29E+10
25	1.015 min	7.49E+05	203.051 days	1.40E+11
30	2.443 min	1.56E+06	2.392 yr	6.11E+11
35	5.369 min	2.91E+06	8.209 yr	2.12E+12
40	8.525 min	4.99E+06	23.888 yr	6.22E+12
45	14.530 min	8.02E+06	61.288 yr	1.61E+13
50	24.243 min	1.22E+07	142.37 yr	3.75E+13

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 min to obtain the proposed estimates.

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 (Fig. 3).

5 Conclusion

The paper proposes a computationally efficient estimators with the same convergence rate as the LSEs or ALSEs. The key idea is to develop a strategy by disintegrating the 2D model into several 1D chirp models and then design an optimal estimation method to obtain estimates of the 2D model parameters. The proposed estimators are not only asymptotically unbiased but also have an asymptotic normal distribution and same rate of convergence as that of the LSEs.

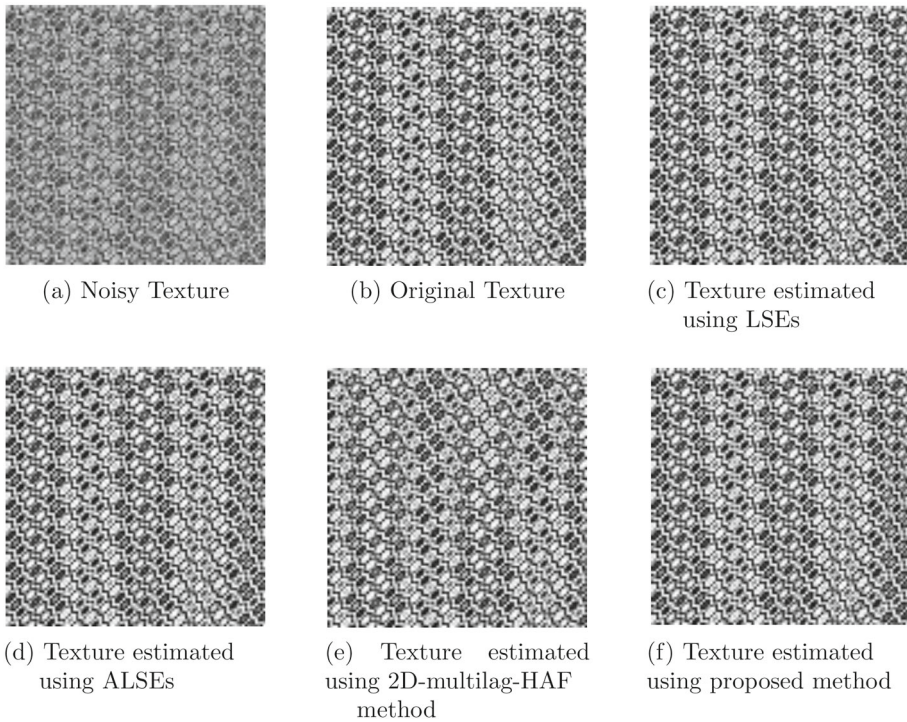


Fig. 3 A comparison of estimated textures using LSEs, ALSEs, 2D-multilag-HAF and the proposed estimators

Furthermore, 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 between time required for obtaining proposed estimates and the LSEs. Synthetic data analysis illustrates the effectiveness of the proposed estimators in recovering 2D gray-scale textures contaminated with noise. For several applications, like enhancing the robustness of the watermarking from frequent attacks e.g., cropping, rotating, compressing, etc., multi-component chirp signals are required for the base watermarking, see Stankovic et al. (2001). The proposed method can effectively be used to address such problems. Further, our proposed methodology can be adapted for parameter estimation of more general signal models by utilizing a sequential procedure similar to Grover et al. (2021). We believe that the estimation technique and the results obtained in this paper can lead to designing computationally efficient algorithms for estimating higher order polynomial phase signals and thus making a significant contribution towards research in this area.

Acknowledgements We would like to thank the Editor for his encouragement and continuous support, and anonymous reviewer for providing critical and constructive comments. Part of the work of the third author has been supported by a research grant from the Science and Engineering Research Board, Government of India.

Author Contributions All authors contributed to study the conceptualization, methodology, review and editing. Formal analysis of proofs and simulations analysis were performed by AS. The first draft of the manuscript was written by AS and all authors commented on revising the previous versions of the manuscript. All authors read and approved the final manuscript. AM and DK had supervised the whole research project.

Funding Part of the work of the third author has been supported by a research grant from the Science and Engineering Research Board, Government of India.

Data Availability Not applicable.

Code Availability The authors can share the code used for simulations, on individual requests.

Declarations

Conflict of interest The authors declares that they have no conflict of interest.

Consent for Publication Not applicable.

Ethical Approval Not applicable.

Consent to Participate Not applicable.

Appendix A

Proof of Theorem 1: Given the data matrix \mathbf{Y} , we compute the LSEs of $\alpha^0 + n_0\mu^0$ and β^0 corresponding to n_0^{th} column vector of \mathbf{Y} . We denote the obtained estimators by $\hat{\alpha}_{n_0}$ and $\hat{\beta}_{n_0}$ to emphasize that these depend on n_0 . Similarly, for fixed m_0^{th} row of \mathbf{Y} , we have denoted the LSEs of $\gamma^0 + m_0\mu^0$ and δ^0 by $\hat{\gamma}_{m_0}$ and $\hat{\delta}_{m_0}$. Under assumptions that $X(m_0, n_0)$ are stationary (11) and (12), see Nandi and Kundu (2004), we have

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

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

The final estimator of β^0 given by

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

is strongly consistent estimate of β^0 which is observed by (15) and also that $\hat{\beta}_{n_0}$ is strongly consistent for β^0 as $M \rightarrow \infty$. For proof, one may refer to Lahiri et al. (2015).

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

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

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

$$\begin{bmatrix} \hat{\alpha} \\ l\hat{\gamma} \\ \hat{\mu} \end{bmatrix} = (\boldsymbol{\Gamma}^\top \boldsymbol{\Gamma})^{-1} \boldsymbol{\Gamma}^\top \mathbf{A} = (\boldsymbol{\Gamma}^\top \boldsymbol{\Gamma})^{-1} \boldsymbol{\Gamma}^\top \left(\boldsymbol{\Gamma} \begin{bmatrix} \alpha^0 \\ \gamma^0 \\ \mu^0 \end{bmatrix} + \boldsymbol{\tau} \right), \tag{17}$$

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

This implies that

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

Now we look at the first element $a_1 + a_2 - a_3$ of the following matrix

$$(\mathbf{\Gamma}^\top \mathbf{\Gamma})^{-1} \begin{bmatrix} N \times o\left(\frac{1}{M}\right) \\ M \times o\left(\frac{1}{N}\right) \\ \frac{N(N+1)}{2} \times o\left(\frac{1}{M}\right) + \frac{M(M+1)}{2} \times o\left(\frac{1}{N}\right) \end{bmatrix}.$$

where,

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

and

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

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

$$\begin{aligned} a_1 &= \left(MK - \frac{M^2(M+1)^2}{4} \right) N \times o\left(\frac{1}{M}\right) / \Xi \\ &= N \times \frac{o(1)}{\Xi} \times \left(\frac{N(N+1)(2N+1)}{6} + \frac{M(M+1)(M-1)}{12} \right) \\ &= \frac{o(1)}{M} \times \frac{\left(2N(N+1)(2N+1) + M(M^2-1) \right)}{\left(N(N^2-1) + M(M^2-1) \right)}. \end{aligned}$$

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

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

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

$$\begin{aligned} a_3 &= \frac{MN(N+1)}{2} \left(\frac{N(N+1)}{2} \times o\left(\frac{1}{M}\right) + \frac{M(M+1)}{2} \times o\left(\frac{1}{N}\right) \right) / \Xi \\ &= \frac{(N+1)}{2\Xi} \left(N^2(N+1) \times o(1) + M^2(M+1) \times o(1) \right). \end{aligned}$$

This implies that $a_3 \xrightarrow{a.s.} 0$ as $\min\{M, N\} \rightarrow \infty$. Hence, $\hat{\alpha}$ is strongly consistent estimate of α^0 . Similarly strong consistency of $\hat{\gamma}$ 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

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

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

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

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

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

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

Appendix B

Proof of Theorem 2: Suppose $\kappa^\top = (A, B, \alpha, \beta)$ and

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

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

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

$$Q'_{n_0}(\hat{\kappa}) - Q'_{n_0}(\kappa^0) = Q''_{n_0}(\check{\kappa})(\hat{\kappa} - \kappa^0), \tag{19}$$

where $\check{\kappa}$ is a point between $\hat{\kappa}$ and κ^0 . Also $Q'_{n_0}(\hat{\kappa}) = 0$ as $\hat{\kappa}$ is LSE of κ^0 .

Let us denote $D_1^{-1} = \text{diag}(M^{1/2}, M^{1/2}, M^{3/2}, M^{5/2})$. Then on multiplying D_1^{-1} both sides of Eq. (19), it gives:

$$\begin{aligned} & - [D_1 Q''_{n_0}(\check{\kappa}) D_1]^{-1} \Sigma_{n_0} \Sigma_{n_0}^{-1} D_1 Q'_{n_0}(\kappa^0), = D_1^{-1}(\hat{\kappa} - \kappa^0) \\ & \text{where } \lim_{M \rightarrow \infty} [D_1 Q''_{n_0}(\check{\kappa}) D_1]^{-1} \Sigma_{n_0} = I_{4 \times 4}, \end{aligned} \tag{20}$$

and

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

The expression of $\Sigma_{n_0}^{-1}$ can be obtained by proof of Theorem 2 shown in Lahiri et al. (2015). By using Eq. (20) and above expression of $\Sigma_{n_0}^{-1}$, we get following asymptotically equivalent (a.e.) expression of the third element of vector $D_1^{-1}(\hat{\kappa} - \kappa^0)$ as:

$$\begin{aligned} & [M^{3/2}(\hat{\alpha}_{n_0} - (\alpha^0 + n_0\mu^0))] \\ & \stackrel{\text{a.e.}}{=} \frac{4}{A^{0^2} + B^{0^2}} \left[18 \sum_m \frac{\eta(m_0, n_0)}{\sqrt{M}} \right. \\ & \left. - 96 \sum_{m_0} \frac{m_0 \eta(m_0, n_0)}{M \sqrt{M}} + 90 \sum_{m_0} \frac{m_0^2 \eta(m_0, n_0)}{M^2 \sqrt{M}} \right], \end{aligned} \tag{21}$$

where, $\eta(m_0, n_0) = X(m_0, n_0)(A^0 \sin \phi(m_0, n_0, \xi^0) - B^0 \cos \phi(m_0, n_0, \xi^0))$,

$\phi(m_0, n_0, \xi^0) = \alpha^0 m_0 + \beta^0 m_0^2 + \gamma^0 n_0 + \delta^0 n_0^2 + \mu^0 m_0 n_0$. We have used these notations for brevity. Similarly expressions of fourth, fifth and sixth element of $D_1^{-1}(\hat{\kappa} - \kappa^0)$ can be written as in Eqs. (22), (23) and (24) respectively:

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

$$[N^{3/2}(\hat{\gamma}_{m_0} - (\gamma^0 + m_0\mu^0))]$$

$$\begin{aligned} &\stackrel{\text{a.e.}}{=} \frac{4}{A^2 + B^2} \left[18 \sum_{n_0} \frac{\eta(m_0, n_0)}{\sqrt{N}} - 96 \sum_{n_0} \frac{n_0 \eta(m_0, n_0)}{N \sqrt{N}} \right. \\ &\quad \left. + 90 \sum_{n_0} \frac{n_0^2 \eta(m_0, n_0)}{N^2 \sqrt{N}} \right], \end{aligned} \tag{23}$$

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

From Eqs. (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 Fuller (1996). 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:

$$\begin{aligned} \begin{bmatrix} \widehat{\alpha} \\ \widehat{\gamma} \\ \widehat{\mu} \end{bmatrix} &= \begin{bmatrix} \left(MK - \frac{M^2(M+1)^2}{4} \right) c_\alpha + \frac{MN(M+1)(N+1)}{4} c_\gamma - \frac{MN(N+1)}{2} c_{\alpha\gamma} \\ \frac{MN(M+1)(N+1)}{4} c_\alpha + \left(KN - \frac{N^2(N+1)^2}{4} \right) c_\gamma - \frac{NM(M+1)}{2} c_{\alpha\gamma} \\ - \frac{MN(N+1)}{2} c_\alpha - \frac{NM(M+1)}{2} c_\gamma + MN c_{\alpha\gamma} \end{bmatrix} \\ &/|\mathbf{\Gamma}^\top \mathbf{\Gamma}|, \\ c_\alpha &= \sum_{n_0=1}^N \widehat{\alpha}_{n_0}, \quad c_\gamma = \sum_{m_0=1}^M \widehat{\gamma}_{m_0}, \\ c_{\alpha\gamma} &= \sum_{n_0=1}^N n_0 \widehat{\alpha}_{n_0} + \sum_{m_0=1}^M m_0 \widehat{\gamma}_{m_0}, \\ K &= \frac{N(N+1)(2N+1)}{6} + \frac{M(M+1)(2M+1)}{6}, \quad |\mathbf{\Gamma}^\top \mathbf{\Gamma}| \\ &= \frac{MN}{12} \left(N(N^2 - 1) + M(M^2 - 1) \right). \end{aligned} \tag{25}$$

From (25), we get:

$$\begin{aligned} M^{3/2} N^{1/2} (\widehat{\alpha} - \alpha^0) &= \left(\frac{2N(N+1)(2N+1) + M(M^2 - 1)}{N(N^2 - 1) + M(M^2 - 1)} \right) \\ &\quad \frac{1}{\sqrt{N}} \sum_{n_0=1}^N M^{3/2} (\widehat{\alpha}_{n_0} - (\alpha^0 + n_0 \mu^0)) \end{aligned}$$

$$\begin{aligned}
 & + \left(\frac{3M^2(M+1)}{N(N^2-1) + M(M^2-1)} \right) \\
 & \frac{1}{\sqrt{M}} \sum_{m_0=1}^M N^{3/2} (\widehat{\gamma}_{m_0} - (\gamma^0 + m_0\mu^0)) \\
 & - \left(\frac{6M^{3/2}N^{1/2}(N+1)}{N(N^2-1) + M(M^2-1)} \right) \left[\sum_{n_0=1}^N n_0 (\widehat{\alpha}_{n_0} - (\alpha^0 + n_0\mu^0)) \right. \\
 & \left. + \sum_{m_0=1}^M m_0 (\widehat{\gamma}_{m_0} - (\gamma^0 + m_0\mu^0)) \right].
 \end{aligned}$$

For sufficiently large M and N , we have:

$$\begin{aligned}
 M^{3/2}N^{1/2}(\widehat{\alpha} - \alpha^0) & = \left(\frac{4N^3 + M^3}{N^3 + M^3} \right) \frac{1}{\sqrt{N}} \sum_{n_0=1}^N M^{3/2}(\widehat{\alpha}_{n_0} - (\alpha^0 + n_0\mu^0)) \\
 & + \left(\frac{3M^3}{N^3 + M^3} \right) \frac{1}{\sqrt{M}} \sum_{m_0=1}^M N^{3/2}(\widehat{\gamma}_{m_0} - (\gamma^0 + m_0\mu^0)) \\
 & - \left(\frac{6N^3}{N^3 + M^3} \right) \frac{1}{N^{3/2}} \sum_{n_0=1}^N n_0 M^{3/2}(\widehat{\alpha}_{n_0} - (\alpha^0 + n_0\mu^0)) \\
 & - \left(\frac{6M^3}{N^3 + M^3} \right) \frac{1}{M^{3/2}} \sum_{m_0=1}^M m_0 N^{3/2}(\widehat{\gamma}_{m_0} - (\gamma^0 + m_0\mu^0)). \tag{26}
 \end{aligned}$$

Asymptotic normality of the estimators for $M = N \rightarrow \infty$ with given rates of convergence follows by applying central limit theorem for stationary processes (Fuller, 1996).

We now present some important results which will be used to find the asymptotic variance-covariance matrix of the proposed estimators of non-linear parameters. Using Eqs. (21) and (23) in the paper, we have the following observations:

1. $AsyVar\left(\frac{1}{2\sqrt{N}} \sum_{n_0=1}^N M^{3/2}(\widehat{\alpha}_{n_0} - (\alpha^0 + n_0\mu^0))\right) = \frac{c\sigma^2 96}{A^0 + B^0},$
2. $AsyVar\left(\frac{1}{2\sqrt{M}} \sum_{m_0=1}^M N^{3/2}(\widehat{\gamma}_{m_0} - (\gamma^0 + m_0\mu^0))\right) = \frac{c\sigma^2 96}{A^0 + B^0},$
3. $AsyVar\left(\frac{1}{2N^{3/2}} \sum_{n_0=1}^N n_0 M^{3/2}(\widehat{\alpha}_{n_0} - (\alpha^0 + n_0\mu^0))\right) = \frac{c\sigma^2 32}{A^0 + B^0},$
4. $AsyVar\left(\frac{1}{2M^{3/2}} \sum_{m_0=1}^M m_0 N^{3/2}(\widehat{\gamma}_{m_0} - (\gamma^0 + m_0\mu^0))\right) = \frac{c\sigma^2 32}{A^0 + B^0},$
5. $AsyCovar\left(\frac{1}{2\sqrt{N}} \sum_{n_0=1}^N M^{3/2}(\widehat{\alpha}_{n_0} - (\alpha^0 + n_0\mu^0)), \frac{1}{2\sqrt{M}} \sum_{m_0=1}^M N^{3/2}(\widehat{\gamma}_{m_0} - (\gamma^0 + m_0\mu^0))\right) = 0,$

6. $AsyCovar\left(\frac{1}{2\sqrt{N}}\sum_{n_0=1}^N M^{3/2}(\hat{\alpha}_{n_0} - (\alpha^0 + n_0\mu^0)), \frac{1}{2N^{3/2}}\sum_{n_0=1}^N n_0 M^{3/2}(\hat{\alpha}_{n_0} - (\alpha^0 + n_0\mu^0))\right) = \frac{c\sigma^2 48}{A^2 + B^2},$
7. $AsyCovar\left(\frac{1}{2\sqrt{N}}\sum_{n_0=1}^N M^{3/2}(\hat{\alpha}_{n_0} - (\alpha^0 + n_0\mu^0)), \frac{1}{2M^{3/2}}\sum_{m_0=1}^M m_0 N^{3/2}(\hat{\gamma}_{m_0} - (\gamma^0 + m_0\mu^0))\right) = 0,$
8. $AsyCovar\left(\frac{1}{2\sqrt{M}}\sum_{m_0=1}^M N^{3/2}(\hat{\gamma}_{m_0} - (\gamma^0 + m_0\mu^0)), \frac{1}{2N^{3/2}}\sum_{n_0=1}^N n_0 M^{3/2}(\hat{\alpha}_{n_0} - (\alpha^0 + n_0\mu^0))\right) = 0,$
9. $AsyCovar\left(\frac{1}{2\sqrt{M}}\sum_{m_0=1}^M N^{3/2}(\hat{\gamma}_{m_0} - (\gamma^0 + m_0\mu^0)), \frac{1}{2M^{3/2}}\sum_{m_0=1}^M m_0 N^{3/2}(\hat{\gamma}_{m_0} - (\gamma^0 + m_0\mu^0))\right) = \frac{c\sigma^2 48}{A^2 + B^2},$
10. $AsyCovar\left(\frac{1}{2N^{3/2}}\sum_{n_0=1}^N n_0 M^{3/2}(\hat{\alpha}_{n_0} - (\alpha^0 + n_0\mu^0)), \frac{1}{2M^{3/2}}\sum_{m_0=1}^M m_0 N^{3/2}(\hat{\gamma}_{m_0} - (\gamma^0 + m_0\mu^0))\right) = \frac{c\sigma^2}{2(A^2 + B^2)}.$

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

Using the above results in Eq. (26) from the paper, we get asymptotic variance-covariance matrix of

$$\begin{bmatrix} M^{3/2}N^{1/2}(\hat{\alpha} - \alpha^0) \\ N^{3/2}M^{1/2}(\hat{\gamma} - \gamma^0) \\ M^{3/2}N^{3/2}(\hat{\mu} - \mu^0) \end{bmatrix} \text{ as: } \frac{c\sigma^2}{(A^2 + B^2)} \begin{bmatrix} 996 & 612 & -1224 \\ 612 & 996 & -1224 \\ -1224 & -1224 & 2448 \end{bmatrix}.$$

From (21), (22) and (23) equations of the paper, it is further observed that:

1. $AsyCovar\left(M^{5/2}N^{1/2}(\hat{\beta} - \beta^0), M^{1/2}N^{5/2}(\hat{\delta} - \delta^0)\right) = 0,$
2. $AsyCovar\left(M^{5/2}N^{1/2}(\hat{\beta} - \beta^0), \frac{1}{\sqrt{N}}\sum_{n_0=1}^N M^{3/2}(\hat{\alpha}_{n_0} - (\alpha^0 + n_0\mu^0))\right) = \frac{-360c\sigma^2}{A^2 + B^2},$
3. $AsyCovar\left(M^{5/2}N^{1/2}(\hat{\beta} - \beta^0), \frac{1}{\sqrt{M}}\sum_{m_0=1}^M N^{3/2}(\hat{\gamma}_{m_0} - (\gamma^0 + m_0\mu^0))\right) = 0,$
4. $AsyCovar\left(M^{5/2}N^{1/2}(\hat{\beta} - \beta^0), \frac{1}{N\sqrt{N}}\sum_{n_0=1}^N M^{3/2}n_0(\hat{\alpha}_{n_0} - (\alpha^0 + n_0\mu^0))\right) = \frac{-180c\sigma^2}{A^2 + B^2},$
5. $AsyCovar\left(M^{5/2}N^{1/2}(\hat{\beta} - \beta^0), \frac{1}{M\sqrt{M}}\sum_{m_0=1}^M m_0 N^{3/2}(\hat{\gamma}_{m_0} - (\gamma^0 + m_0\mu^0))\right) = 0.$

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

Asymptotic variance-covariance matrix of the proposed estimators of non-linear parameters is given by:

$$\frac{c\sigma^2}{(A^0 + B^0)} \begin{bmatrix} 996 & -360 & 612 & 0 & -1224 \\ -360 & 360 & 0 & 0 & 0 \\ 612 & 0 & 996 & -360 & -1224 \\ 0 & 0 & -360 & 360 & 0 \\ -1224 & 0 & -1224 & 0 & 2448 \end{bmatrix}. \tag{27}$$

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

$$\cos \hat{\phi} - \cos \phi^0 = -\sin \check{\phi} (\hat{\xi} - \xi^0)^\top \begin{bmatrix} m_0 \\ m_0^2 \\ n_0 \\ n_0^2 \\ m_0 n_0 \end{bmatrix}.$$

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

Now consider first element of the following vector,

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

we get:

$$\sqrt{MN} \left(\frac{2}{MN} \sum_{m_0=1}^M \sum_{n_0=1}^N y(m_0, n_0) \cos \phi^0 - \frac{2}{MN} \sum_{m_0=1}^M \sum_{n_0=1}^N y(m_0, n_0) \sin \check{\phi} R(m_0, n_0) - A^0 \right), \tag{28}$$

where $R(m_0, n_0) = (\hat{\xi} - \xi^0)^\top \begin{bmatrix} m_0 \\ m_0^2 \\ n_0 \\ n_0^2 \\ m_0 n_0 \end{bmatrix}$, and second element of the above amplitude vector

can be written as:

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

Now let us look at the first and the last term of Eq. (28) and putting value of $y(m_0, n_0)$ from the model (1),

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

$$\begin{aligned}
 &= \sqrt{MN} \left(\frac{2}{MN} \sum_{m_0=1}^M \sum_{n_0=1}^N A^0 \cos^2 \phi^0 - A^0 + \frac{2}{MN} \sum_{m_0=1}^M \sum_{n_0=1}^N B^0 \sin \phi^0 \cos \phi^0 \right. \\
 &\quad \left. + \frac{2}{MN} \sum_{m_0=1}^M \sum_{n_0=1}^N X(m_0, n_0) \cos \phi^0 \right) \\
 &\stackrel{\text{a.e.}}{=} \frac{2}{\sqrt{MN}} \sum_{m_0=1}^M \sum_{n_0=1}^N X(m_0, n_0) \cos \phi^0. \tag{30}
 \end{aligned}$$

The above result has been obtained from a famous number theory conjecture by Montgomery (1994).

In second term of (28), $R(m_0, n_0)$ is a sum of five terms, so now consider first term of

$$\begin{aligned}
 &\frac{2}{MN} \sum_{m_0=1}^M \sum_{n_0=1}^N y(m_0, n_0) \sin \check{\phi} R(m_0, n_0), \\
 &\frac{2}{\sqrt{MN}} \sum_{m_0=1}^M \sum_{n_0=1}^N y(m_0, n_0) \sin \check{\phi} m_0 (\hat{\alpha} - \alpha^0) \\
 &= \frac{2}{\sqrt{MN}} \sum_{m_0=1}^M \sum_{n_0=1}^N A^0 \cos \phi^0 \sin \check{\phi} m_0 (\hat{\alpha} - \alpha^0) \\
 &\quad + \frac{2}{\sqrt{MN}} \sum_{m_0=1}^M \sum_{n_0=1}^N B^0 \sin \phi^0 \sin \check{\phi} m_0 (\hat{\alpha} - \alpha^0) \\
 &\quad + \frac{2}{\sqrt{MN}} \sum_{m_0=1}^M \sum_{n_0=1}^N X(m_0, n_0) \sin \check{\phi} m_0 (\hat{\alpha} - \alpha^0) \\
 &= \left(\frac{2}{M^2 N} \sum_{m_0=1}^M \sum_{n_0=1}^N A^0 \cos \phi^0 \sin \check{\phi} m_0 \right) M \sqrt{MN} (\hat{\alpha} - \alpha^0) \\
 &\quad + \left(\frac{2}{M^2 N} \sum_{m_0=1}^M \sum_{n_0=1}^N B^0 \sin \phi^0 \sin \check{\phi} m_0 \right) M \sqrt{MN} (\hat{\alpha} - \alpha^0) \\
 &\quad + \left(\frac{2}{M^2 N} \sum_{m_0=1}^M \sum_{n_0=1}^N X(m_0, n_0) \sin \check{\phi} m_0 \right) M \sqrt{MN} (\hat{\alpha} - \alpha^0) \\
 &\stackrel{\text{a.e.}}{=} \left(\frac{2}{M^2 N} \sum_{m_0=1}^M \sum_{n_0=1}^N B^0 \sin \phi^0 \sin \check{\phi} m_0 \right) M \sqrt{MN} (\hat{\alpha} - \alpha^0), \text{ by using Proposition 1 of [16].}
 \end{aligned}$$

So, finding the asymptotic distribution of (28) boils down to finding asymptotic distribution of

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

which is further asymptotically equivalent to:

$$\frac{2}{\sqrt{MN}} \sum_{m_0=1}^M \sum_{n_0=1}^N X(m_0, n_0) \cos \phi^0 - B^0 \left(\frac{1}{2} M^{3/2} N^{1/2} (\hat{\alpha} - \alpha^0) + \frac{1}{3} M^{5/2} N^{1/2} (\hat{\beta} - \beta^0) \right)$$

$$+ \frac{1}{2}N^{3/2}M^{1/2}(\widehat{\gamma} - \gamma^0) + \frac{1}{3}N^{5/2}M^{1/2}(\widehat{\delta} - \delta^0) + \frac{1}{4}M^{3/2}N^{3/2}(\widehat{\mu} - \mu^0) \Big). \quad (31)$$

Asymptotic normality of amplitude estimators is thus proved by Eq. (31). 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 Σ as mentioned in Theorem 2.

Hence the result. \square

References

- Barbarossa, S., Scaglione, A., & Giannakis, G. B. (1998). Product high-order ambiguity function for multicomponent polynomial-phase signal modeling. *IEEE Transactions on Signal Processing*, 46(3), 691–708.
- 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.
- 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.
- Djurović, I., & Stanković, L. (2014). Quasi-maximum-likelihood estimator of polynomial phase signals. *IET Signal Processing*, 8(4), 347–359.
- Djurović, I. (2017). Quasi ML algorithm for 2-D PPS estimation. *Multidimensional Systems and Signal Processing*, 28(2), 371–387.
- Francois, J. M., & Friedlander, B. (1995). The polynomial phase difference operator for modeling of nonhomogeneous images. In *Proceedings., International Conference on Image Processing* (Vol. 2, pp. 276–279). IEEE.
- Friedlander, B., & Francois, J. M. (1996). Model based phase unwrapping of 2-D signals. *IEEE Transactions on Signal Processing*, 44(12), 2999–3007.
- Francois, J. M., & Friedlander, B. (1998). Two-dimensional polynomial phase signals: Parameter estimation and bounds. *Multidimensional Systems and Signal Processing*, 9(2), 173–205.
- Francois, J. M., & Friedlander, B. (1999). Parameter estimation of 2-D random amplitude polynomial-phase signals. *IEEE Transactions on Signal Processing*, 47(7), 1795–1810.
- Fuller, W. A. (1996). *Introduction to Statistical Time Series* (2nd ed.). John Wiley and Sons.
- 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.
- 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.
- 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.
- Lahiri, A., Kundu, D., & Mitra, A. (2013). Efficient algorithm for estimating the parameters of two dimensional Chirp signal. *Sankhya B*, 75(1), 65–89.
- Lahiri, A., Kundu, D., & Mitra, A. (2015). Estimating the parameters of multiple chirp signals. *Journal of Multivariate Analysis*, 139, 189–206.
- Lahiri, A., & Kundu, D. (2017). On parameter estimation of two-dimensional polynomial phase signal model. *Statistica Sinica*, 27, 1779–1792.
- Montgomery, H. L. (1994). Ten lectures on the interface between analytic number theory and harmonic analysis (No. 84). American Mathematical Soc.
- 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.
- O'shea, P. (2002). A new technique for instantaneous frequency rate estimation. *IEEE Signal Processing Letters*, 9(8), 251–252.
- Peleg, S., & Porat, B. (1991). Estimation and classification of polynomial-phase signals. *IEEE Transactions on Information Theory*, 37(2), 422–430.

- Stankovic, S., Djurovic, I., & Pitas, I. (2001). Watermarking in the space/spatial-frequency domain using two-dimensional Radon-Wigner distribution. *IEEE Transactions on Image Processing*, 10(4), 650–658.
- 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.
- Zhang, K., Wang, S., & Cao, F. (2008). Product cubic phase function algorithm for estimating the instantaneous frequency rate of multicomponent two-dimensional Chirp signals. In *2008 Congress on Image and Signal Processing* (Vol. 5, pp. 498–502). IEEE.
- Zhang, Y., Mobasser, B. G., Dogahe, B. M., & Amin, M. G. (2010). Image-adaptive watermarking using 2D Chirps. *Signal, Image and Video Processing*, 4(1), 105–121.

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.