Robust Estimators of Two Dimensional Sinusoidal Model Parameters

Debasis Kundu [*]& Rhythm Grover [†]

Abstract

In this paper we consider the two dimensional (2-D) sinusoidal model. This particular model has several applications in the statistical signal processing and in texture analysis. Extensive work has been done in developing several efficient procedures and establishing their properties. The least squares estimators (LSEs) are known to be the most efficient estimators in presence of additive noise. But it is observed that the LSEs are quite sensitive in presence of outliers. In this paper we propose to use the weighted least squares estimators (WLSEs) of the unknown parameters in presence of additive white noise. It is observed that in presence of outliers, the WLSEs are more robust than the least squares estimators (LSEs) and they behave very similarly to some of the other well known robust estimators, for example the least absolute deviation estimators (LADEs). It is observed that developing the properties of the LADEs is not immediate. We derive the consistency and asymptotic normality properties of the WLSEs. Extensive simulations have been performed to show the effectiveness of the proposed method. One synthetic data set has been analyzed to illustrate how the proposed method can be used in practice.

Key Words and Phrases: Two-dimensional sinusoidal model; least squares estimators; weighted least squares estimators; robust estimators; consistency; asymptotic normality.

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

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

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

1 Introduction

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

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

Here A^0 and B^0 are non-zero real numbers, $A^{0^2} + B^{0^2}$ is known as the amplitude of the signal, μ^0 and λ^0 are frequencies, $X(m,n)$ is the additive random errors. It is assumed that $\{X(m,n)\}$ is a two-dimensional sequence of independent and identically distributed (i.i.d.) random variables with mean zero and finite variance σ^2. Based on a random sample $\{y(m,n); m = 1,\ldots,M, n = 1,\ldots,N\}$, the problem is to estimate the unknown parameters namely A^0, B^0, μ^0 and λ^0.

The 2-D sinusoidal model (1) has received a significant amount of attention in the statistical signal processing area due to its applications in image processing and in texture analysis, see for example Grover et al. [3], Porwal et al. [13], Ivanov and Lymar[6, 7], Ivanov and Savych [8], Hou et al. [5], Marchant and Jackway [11], Nandi and Kundu [12], Chen et al. [1], Kundu and Nandi [10], Zhang and Mandrekar [16], Rao et al. [15] and see the references cited therein. In all these cases attempts have been made in developing efficient algorithms, deriving their properties and compare their performances with the corresponding least squares estimators (LSEs). It may be mentioned that the LSEs are the most efficient estimators, whose variances attend the Cramer-Rao lower bound under the assumptions that the errors are i.i.d. normal random variables. Although the performances of most of the efficient estimators are quite satisfactory and they are quite comparable with the performances of the LSEs, it is observed all these estimators are quite susceptible in presence of outliers. Even if there are only few outliers present in the data set, that can affect the performance of the estimators quite significantly. One natural choice will be to use some robust estimators, like least absolute deviation LADEs in this situation, see also Fu et al. [2] in this respect.

There are mainly two major issues in using LADEs in practice for a non-linear model like (1). Even in case of one-dimensional (1-D) sinusoidal model, it has been observed by Kim et al. [9] that in establishing the consistency and asymptotic normality of the LADEs one needs to make some stronger assumption on the error random variables. Extending the 1-D results to 2-D results are not immediate, and it has not yet been established. Moreover, implementing the LADEs for model (1), one needs to solve a four dimensional optimization problem. The problem becomes more complicated for the multicomponent model. The same is true even for Huber's M-estimators also. No work has been done in this case even for one dimensional sinusoidal model also. It is not very clear how it can be extended to two dimensional model. Moreover, it is also observed that when there is no outliers present in the data set, then any robust estimator does not perform well.

The aim of this paper is to develop a robust estimation procedures and we propose to use WLSEs of the unknown parameters of the 2-D sinusoidal model (1) under the same set of assumptions as it has been mentioned above. First we consider the case when the weight function is a k-th degree polynomial in 2-D. It is observed that the WLSEs are consistent and asymptotically normally distributed. Moreover, the WLSEs can be obtained by solving a two dimensional optimization problem. Extensive simulations have been performed to compare the performances of the WLSEs with some of the other estimators. It is observed that the WLSEs are quite robust in presence of few outliers. In presence of outliers its performances are better than the LSEs, and are very similar with the LADEs. When there is no outliers present, the performances of the WLSEs are very similar to the LSEs, and they are better than the LADEs. We extend the results for general weight functions also and for multicomponent model also. One synthetic data set has been analyzed to show how the proposed estimators can be implemented in practice.

The main contributions of this paper are the following: (a) We have proposed a simple WLSEs which are robust estimators in presence of outliers. (b) We have developed theoretical properties of the proposed estimators under the same set of assumptions which are needed to establish the consistency and asymptotic normality properties of the LSEs. (c) The proposed method is very easy to implement in practice and it can be used quite conveniently even when the number of components is large.

The rest of the paper is organized as follows. In Section 2 we provide how to obtain the WLSEs. The consistency and the asymptotic normality properties have been provided in Section 3. In Section 4 we discuss about the general weight function. The numerical results are performed in Section 5, We extend the results for multicomponent model in Section 6. For illustrative purposes we have provided the analysis of a synthetic texture data set in Section 7, and finally we conclude the paper in Section 8.

2 Weighted Least Squares Estimators

In this section we provide the WLSEs of the unknown parameters namely A^0, B^0, μ^0 and λ^0, of the model (1). Suppose, the weight function $w(s,t)$ is a polynomial of degree k, defined for $0 \leq s, t \leq 1$, and it can be written as follows:

$$w(s,t) = c_{0,0} + c_{1,0}s + c_{0,1}t + c_{2,0}s^2 + c_{1,1}st + c_{0,2}t^2 + \ldots + c_{k,0}s^k + \ldots + c_{0,k}t^k. \qquad (2)$$

There are certain restrictions on the coefficients of the polynomial (2), we will make it explicit later. Let us define the weighted residual sums of squares as follows:

$$Q(A,B,\mu,\lambda) = \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right)\left(y(m,n) - A\cos(m\mu + n\lambda) - B\sin(m\mu + n\lambda)\right)^2. \qquad (3)$$

The WLSEs can be obtained by minimizing (3) with respect to A, B, μ and λ. We use the separable regression of Richards [14] technique to minimize (3). It reduces solving a four dimensional optimization problem to a two dimensional optimization problem. If we denote the WLSEs of A and B for a given μ and λ as $\widehat{A}(\mu,\lambda)$ and $\widehat{B}(\mu,\lambda)$, respectively, then

$$\begin{aligned}\widehat{A}(\mu,\lambda) &= \frac{1}{a_{11}a_{22}-(a_{12})^2}\left[a_{22}d_1 - a_{12}d_2\right]\\ \widehat{B}(\mu,\lambda) &= \frac{1}{a_{11}a_{22}-(a_{12})^2}\left[a_{11}d_2 - a_{12}d_1\right],\end{aligned} \tag{4}$$

where

$$\begin{aligned}a_{11} &= \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right)\cos^2(m\mu+n\lambda)\\ a_{22} &= \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right)\sin^2(m\mu+n\lambda)\\ a_{12}=a_{21} &= \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right)\cos(m\mu+n\lambda)\sin(m\mu+n\lambda)\\ d_1 &= \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right)y(m,n)\cos(m\mu+n\lambda)\\ d_2 &= \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right)y(m,n)\sin(m\mu+n\lambda).\end{aligned}$$

Therefore, the WLSEs of μ and λ can be obtained by minimizing $Q(\widehat{A}(\mu,\lambda),\widehat{B}(\mu,\lambda),\mu,\lambda)$ with respect to μ and λ first and then obtain $\widehat{A}(\widehat{\mu},\widehat{\lambda})$ and $\widehat{B}(\widehat{\mu},\widehat{\lambda})$. It can be verified that

$$\lim_{M,N\to\infty} a_{11} = \frac{1}{2}\int_0^1\int_0^1 w(s,t)dsdt, \tag{5}$$

$$\lim_{M,N\to\infty} a_{22} = \frac{1}{2}\int_0^1\int_0^1 w(s,t)dsdt, \tag{6}$$

$$\lim_{M,N\to\infty} a_{12} = \lim_{M,N\to\infty} a_{21} = 0. \tag{7}$$

The above results, namely (5), (6) and (7), mainly follow based on the fact that $w(s,t)$ is a polynomial of finite degree. For large M and N, we consider

$$\widehat{A}(\mu,\lambda) = \frac{2d_1}{\int_0^1\int_0^1 w(s,t)dsdt} \quad \text{and} \quad \widehat{B}(\mu,\lambda) = \frac{2d_2}{\int_0^1\int_0^1 w(s,t)dsdt}. \tag{8}$$

Hence, the WLSEs of μ and λ are obtained by minimizing $Q(\widehat{A}(\mu,\lambda),\widehat{B}(\mu,\lambda),\mu,\lambda)$, where $\widehat{A}(\mu,\lambda)$ and $\widehat{B}(\mu,\lambda)$ are as defined in(8).

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

3 Large Sample Properties of the WLSEs

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

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

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

Assumption 3: (i) $\inf\limits_{0 \le s,t \le 1} w(s, t) > \gamma > 0$, and (ii) $\sup\limits_{0 \le s,t \le 1} w(s, t) \le 1$.

Theorem 1: Under Assumptions 1,2 and 3, the WLSEs of the unknown parameters namely A^0, B^0, μ^0 and λ^0 are strongly consistent.

Proof: See in the Appendix. ∎

Now we are going to provide the asymptotic normality properties of the WLSEs, and we need to use the following notations for $p, q = 0, 1, 2 \ldots$ for that purpose.

$$u_{pq} = \int_0^1 \int_0^1 s^p t^q w(s,t) ds dt, \quad v_{pq} = \int_0^1 \int_0^1 s^p t^q w^2(s,t) ds dt. \tag{9}$$

The 4×4 diagonal matrix $\boldsymbol{D}$ is as follows:

$$\boldsymbol{D} = \operatorname{diag}\left\{M^{1/2}N^{1/2}, M^{1/2}N^{1/2}, M^{3/2}N^{1/2}, M^{1/2}N^{3/2}\right\}. \tag{10}$$

The matrix $\boldsymbol{\Sigma}$ is

$$\boldsymbol{\Sigma} = \begin{bmatrix} v_{00} & 0 & B^0 v_{10} & B^0 v_{01} \\ 0 & v_{00} & -A^0 v_{10} & -A^0 v_{01} \\ B^0 v_{10} & -A^0 v_{10} & (A^{0^2} + B^{0^2}) v_{20} & (A^{0^2} + B^{0^2}) v_{11} \\ B^0 v_{01} & -A^0 v_{01} & (A^{0^2} + B^{0^2}) v_{11} & (A^{0^2} + B^{0^2}) v_{02} \end{bmatrix} \tag{11}$$

and the matrix $\boldsymbol{\Gamma}$ is

$$\boldsymbol{\Gamma} = \begin{bmatrix} u_{00} & 0 & B^0 u_{10} & B^0 u_{01} \\ 0 & u_{00} & -A^0 u_{10} & -A^0 u_{01} \\ B^0 u_{10} & -A^0 u_{10} & (A^{0^2} + B^{0^2}) u_{20} & (A^{0^2} + B^{0^2}) u_{11} \\ B^0 u_{01} & -A^0 u_{01} & (A^{0^2} + B^{0^2}) u_{11} & (A^{0^2} + B^{0^2}) u_{02} \end{bmatrix}. \tag{12}$$

Theorem 2: Under Assumptions 1,2 and 3,

$$\boldsymbol{D}\left(\widehat{A} - A^0, \widehat{B} - B^0, \widehat{\mu} - \mu^0, \widehat{\lambda} - \lambda^0\right)^\top \rightarrow N_4\left(\mathbf{0}, 2\sigma^2 \boldsymbol{\Gamma}^{-1} \boldsymbol{\Sigma} \boldsymbol{\Gamma}^{-1}\right).$$

Proof: See in the Appendix. ∎

Theorem 2 can be used for construction of the confidence intervals and also for testing purposes. The performance bounds of the WLSEs can be obtained from the diagonal entries of the matrix $2\sigma^2\boldsymbol{\Gamma}^{-1}\boldsymbol{\Sigma}\boldsymbol{\Gamma}^{-1}$. It indicates that the variance of the WLSEs of μ and λ will be in the order of $M^{-3/2}N^{-1/2}$ and $M^{-1/2}N^{-3/2}$, respectively. The performance bounds depend on the weight function $w(\cdot,\cdot)$, as u_{pq} and v_{pq} depend on the weight function. The performance bound of the LSEs can be easily obtained from here by using the fact: $u_{00} = v_{00} = 1$,$u_{10} = u_{01} = v_{10} = v_{01} = 1/2$,$u_{20} = u_{02} = v_{20} = v_{02} = 1/3$,$u_{11} = v_{11} = 1/4$, and they are the Cramer-Rao lower bounds also. It may be mentioned that the performance bounds of the LADEs are not available yet.

4 General Weight Function

So far we have considered the two-dimensional weight function $w(s,t)$ defined on $[0,1] \times [0,1]$ to be continuous and satisfies Assumption 3. Based on that we have established the consistency and asymptotic normality properties of the WLSEs. In this section we will show that the results can be extended for general continuous weight functions. We make the following assumptions on the weight function $w(s,t)$.

Assumption 4: The weight function $w(s,t)$ is a continuous function on $[0,1] \times [0,1]$. Further, (i) $\inf\limits_{0 \leq s,t \leq 1} w(s,t) > \gamma > 0$, and (ii) $\sup\limits_{0 \leq s,t \leq 1} w(s,t) \leq 1$.

Theorem 3: Under Assumptions 1,2 and 4, the WLSEs of A^0, B^0, μ^0 and λ^0 are strongly consistent.

Proof: Following the proof of Theorem 1, it may be seen that no where it has been used that $w(s,t)$ is a polynomial. The continuity property and Assumption 3, have been used. Hence, the result follows. ∎

If we use the same notations as in (9), then we have the following asymptotic distribution of the WLSEs.

Theorem 4: Under Assumptions 1, 2 and 4,

$$\boldsymbol{D}\left(\widehat{A} - A^0, \widehat{B} - B^0, \widehat{\mu} - \mu^0, \widehat{\lambda} - \lambda^0\right)^{\top} \rightarrow N_4\left(\mathbf{0}, 2\sigma^2 \mathbf{\Gamma}^{-1}\mathbf{\Sigma}\mathbf{\Gamma}^{-1}\right).$$

Proof: See in the Appendix.

5 Numerical Results

In this section we present some numerical results to show how the proposed estimators behave in practice in presence of outliers, for different model parameters, different sample sizes, different weight functions and for different error variances. It is assumed that the error random variables $X(m,n)$'s are i.i.d. Gaussian random variables with mean 0 and variance σ_0^2. The outliers are present in the middle of the data set and they are also i.i.d. Gaussian random variables with mean 0, and variance σ_1^2. For comparison purposes, we have considered LADEs also. Let us recall that the LADEs of the unknown parameters of the model (1) can be obtained by minimizing

$$Q_{L1}(A,B,\mu,\lambda) = \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\Big|y(m,n) - A\cos(m\mu+n\lambda) - B\sin(m\mu+n\lambda)\Big|, \tag{13}$$

with respect to A, B, μ and λ. Hence, to obtain the LADEs of the unknown parameters, one needs to solve a four dimensional optimization problem.

The following models have been considered:

$$\text{Model 1}: A = B = 1.0, \quad \lambda = \mu = 0.5, \quad \sigma_0^2 = 2.0, \quad \sigma_1^2 = 25$$

$$\text{Model 2}: A = B = 1.0, \quad \lambda = \mu = 0.5, \quad \sigma_0^2 = 1.0, \quad \sigma_1^2 = 100.$$

It may be seen that the ratio $\frac{\sigma_1^2}{\sigma_0^2}$ is significantly larger in Model 2 compared to Model 1, and we want to see the effect of this. For both the models, we have considered different sample sizes, and different proportions of outliers. We have also considered the case when there is no outliers in the data set. When the outliers are present in the data, it is assumed that they are present in the middle portion of the data set. Due to this reason, we have used the following two weight functions in this case:

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

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

Note that W1 is a polynomial weight function, and W2 is a continuous general weight function and both the weight functions put less weights in the middle portion of the data set. We have also reported the results for the least squares estimators also, for comparison purposes. In each case we have a generated a sample of size $M \times N$ from the given model, and we estimate the parameters based on different methods, we compute the average estimates and the square root of the mean square errors (SMSEs) based on 1000 replications. We have reported the results mainly for the frequencies, as they are the non-linear parameters, and more difficult to compute than the linear parameters namely A and B. The general trends of the results related to the linear parameters are quite similar with the non-linear parameters, hence they are not reported here. They can be obtained on request from the corresponding author.

Some of the points are quite clear from these simulation experiments. For all the methods the performance deteriorate in terms of the biases and SMSEs when the number of outliers increases, and similarly the performances become better as the sample size increases. The LSEs perform the best when there is no outliers in the data set. The performances of W1 and W2 are very similar in all the cases considered here. When the number of outliers (NOL) is small then W1 and W2 perform better than the LSEs and LADEs. When the number of outliers is quite large then the LADEs perform the best. When there is no outliers in the data set, the performances of the LADEs are not very satisfactory. Similarly, when the number of outliers present in the data set is high, then the LSEs do not perform well. In general it is observed that when there is no outliers or few outliers in the data set, then the performances of W1 and W2 are quite similar to the LSEs, and when the number of outliers is high, then the performances of W1 and W2 are quite close to LAD estimators.

Now we would like to point out two issues related to LADEs. The first issue is regarding the theoretical properties of the LADEs of the model (1). Although, Kim et al. [9] established the consistency and asymptotic normality properties of the LAD estimators, the same is not available in the literature till today for the two dimensional sinusoidal model. The second issue is regarding the computational issue. It is observed that the LADEs can be obtained by solving a 4-D optimization problems. One needs to provide a 4-D initial guesses of the unknown parameters for any iterative process to compute the LADEs. On the other hand to compute the LSEs and the WLSEs one needs to solve a 2-D optimization problem. Hence, computationally, the LADEs is more challenging than the LSEs and WLSEs. Considering all these points we recommend to use W1 or W2.

NOL	Par	LSE	W1	W2	LADE
0	μ	0.4999	0.4999	0.4999	0.4996
		(1.232E-2)	(1.229E-2)	(1.230E-2)	(1.483E-2)
	λ	0.5003	0.5002	0.5002	0.5008
		(1.191E-2)	(1.197E-2)	(1.201E-2)	(1.505E-2)
25	μ	0.5000	0.5000	0.5000	0.4995
		(1.306E-2)	(1.280E-2)	(1.268E-2)	(1.491E-2)
	λ	0.5004	0.5003	0.5003	0.5006
		(1.246E-2)	(1.232E-2)	(1.225E-2)	(1.497E-2)
49	μ	0.5002	0.5002	0.5001	0.4996
		(1.463E-2)	(1.396E-2)	(1.361E-2)	(1.514E-2)
	λ	0.5005	0.5004	0.5004	0.5006
		(1.406E-2)	(1.345E-2)	(1.314E-2)	(1.506E-2)
81	μ	0.5003	0.5003	0.5002	0.4993
		(1.779E-2)	(1.635E-2)	(1.577E-2)	(1.560E-2)
	λ	0.5013	0.5007	0.5006	0.5005
		(2.149E-2)	(1.582E-2)	(1.527E-2)	(1.550E-2)
361	μ	0.5051	0.5171	0.5224	0.5020
		(0.49481)	(0.37346)	(0.37543)	(0.37526)
	λ	0.5391	0.52186	0.5172	0.5133
		(0.50884)	(0.39425)	(0.38552)	(0.39235)

Table 1: The average estimates and the associated MSEs of the non-linear parameters for Model 1 of the different estimators when $M = N = 20$.

NOL	Par	LSE	W1	W2	LADE
0	μ	0.5001	0.5001	0.5001	0.5007
		(7.764E-3)	(7.757E-3)	(7.768E-3)	(9.460E-3)
	λ	0.5000	0.5000	0.5000	0.5002
		(7.726E-3)	(7.761E-3)	(7.784E-3)	(9.482E-3)
25	μ	0.4999	0.4999	0.5000	0.5006
		(8.008E-3)	(7.920E-3)	(7.883E-3)	(9.360E-3)
	λ	0.5001	0.5000	0.5000	0.5001
		(7.993E-3)	(7.960E-3)	(7.923E-3)	(9.406E-3)
49	μ	0.4999	0.5000	0.5000	0.5008
		(8.493E-3)	(8.266E-3)	(8.166E-3)	(9.489E-3)
	λ	0.5001	0.5000	0.5000	0.5002
		(8.401E-3)	(8.223E-3)	(8.134E-3)	(9.360E-3)
81	μ	0.5000	0.5000	0.5000	0.5005
		(9.563E-3)	(9.062E-3)	(8.864E-3)	(9.771E-3)
	λ	0.5002	0.5002	0.5001	0.5000
		(9.481E-3)	(9.002E-3)	(8.816E-3)	(9.524E-3)
361	μ	0.5010	0.4984	0.5004	0.5006
		(0.10806)	(3.969E-2)	(3.833E-2)	(4.346E-2)
	λ	0.5032	0.5033	0.5017	0.5000
		(0.10693)	(4.296E-2)	(4.472E-2)	(4.349E-2)

Table 2: The average estimates and the associated MSEs of the non-linear parameters for Model 1 of the different estimators when $M = N = 25$.

6 Multicomponent 2-D Sinusoidal Model

So far we have discussed about the one component 2-D sinusoidal model. In this section we consider the multicomponent model (14). Let us use the following notations for $k = 1, \ldots, p$;

$$\theta_k = (A_k, B_k, \mu_k, \lambda_k)^\top, \quad \theta_k^0 = (A_k^0, B_k^0, \mu_k^0, \lambda_k^0)^\top, \quad \Theta = (\theta_1, \ldots, \theta_p)^\top, \quad \Theta^0 = (\theta_1^0, \ldots, \theta_p^0)^\top.$$

The multicomponent 2-D sinusoidal model has the following form:

$$y(m,n) = \sum_{k=1}^{p} g_k(m,n;\theta_k^0) + X(m,n); \qquad m = 1, \ldots, M; \quad n = 1, \ldots, N, \tag{14}$$

where

$$g_k(m,n;\theta_k^0) = A_k^0 \cos(m\mu_k^0 + n\lambda_k^0) + B_k^0 \sin(m\mu_k^0 + n\lambda_k^0).$$

NOL	Par	LSE	W1	W2	LADE
0	μ	0.4999	0.4999	0.4999	0.4998
		(8.659E-3)	(8.638E-3)	(8.644E-3)	(1.048E-2)
	λ	0.5002	0.5001	0.5001	0.5006
		(8.370E-3)	(8.407E-3)	(8.431E-3)	(1.059E-2)
25	μ	0.5002	0.5001	0.5000	0.4996
		(1.057E-2)	(1.001E-2)	(9.666E-3)	(1.068E-2)
	λ	0.5004	0.5003	0.5003	0.5004
		(9.980E-3)	(9.511E-3)	(9.232E-3)	(1.067E-2)
49	μ	0.5021	0.5004	0.5003	0.4995
		(4.678E-2)	(1.275E-2)	(1.193E-2)	(1.099E-2)
	λ	0.4998	0.5005	0.5004	0.5003
		(5.440E-2)	(1.228E-2)	(1.148E-2)	(1.093E-2)
81	μ	0.5022	0.5028	0.5001	0.4994
		(0.14784)	(9.571E-2)	(6.078E-2)	(1.143E-2)
	λ	0.5078	0.5009	0.5030	0.5003
		(0.14775)	(9.407E-2)	(5.838E-2)	(1.135E-2)
361	μ	0.5399	0.5165	0.5215	0.5215
		(0.73039)	(0.45763)	(0.44234)	(0.45131)
	λ	0.5575	0.5296	0.5232	0.5312
		(0.75249)	(0.45342)	(0.46213)	(0.45671)

Table 3: The average estimates and the associated MSEs of the non-linear parameters for Model 2 of the different estimators when $M = N = 20$.

Here $\{(A_k^0, B_k^0); k = 1, \ldots\}$ are unknown linear parameters, similarly, $\{(\mu_k^0, \lambda_k^0); k = 1, \ldots, p\}$ are unknown frequencies, 'p', the number of sinusoidal components, is assumed to be known, $X(m,n)$'s are noise random variables satisfying Assumption 2. Rest of this section for identifiability purposes, it is assumed without loss of generality

$$0 < A_p^{0^2} + B_p^{0^2} < \ldots < A_1^{0^2} + B_1^{0^2}.$$

The maim aim of this section is to first consider the WLSEs of the unknown parameters and derive their properties. It has been shown that the WLSEs are consistent estimators, and they are asymptotically normally distributed after proper normalization. It is observed that the WLSEs of the unknown parameters can be obtained by solving a $2p$ dimensional

NOL	Par	LSE	W1	W2	LADE
0	μ	0.5000	0.5000	0.5000	0.5004
		(5.482E-3)	(5.477E-3)	(5.485E-3)	(6.783E-3)
	λ	0.5000	0.5000	0.5000	0.5001
		(5.449E-3)	(5.476E-3)	(5.489E-3)	(6.738E-3)
25	μ	0.4998	0.4998	0.4999	0.5003
		(6.157E-3)	(5.947E-3)	(5.821E-3)	(6.756E-3)
	λ	0.5001	0.5001	0.5000	0.5000
		(6.157E-3)	(5.987E-3)	(5.864E-3)	(6.648E-3)
49	μ	0.4999	0.4999	0.4999	0.5001
		(7.396E-3)	(6.864E-3)	(6.582E-3)	(6.866E-3)
	λ	0.5001	0.5001	0.5001	0.4997
		(7.317E-3)	(6.805E-3)	(6.532E-3)	(6.768E-3)
81	μ	0.4997	0.5001	0.5000	0.5000
		(1.431E-2)	(8.856E-3)	(8.261E-3)	(7.140E-3)
	λ	0.5008	0.5006	0.5002	0.4997
		(1.497E-2)	(1.399E-2)	(8.254E-3)	(6.857E-3)
361	μ	0.5251	0.5178	0.5177	0.4997
		(0.39567)	(0.24383)	(0.22833)	(0.23165)
	λ	0.4999	0.4989	0.4993	0.4992
		(0.37958)	(0.23567)	(0.21734)	(0.23124)

Table 4: The average estimates and the associated MSEs of the non-linear parameters for Model 2 of the different estimators when $M = N = 25$.

optimization problem, and this can be quite challenging if p is large. To avoid that we have proposed a sequential WLSEs of the unknown parameters, and this can be obtained by solving p two dimensional optimization problem. Hence, it can be implemented quite easily even if p is large. Moreover, we have shown that the sequential WLSEs have the same asymptotic properties as the WLSEs.

Now we will describe the WLSEs. We need the following assumptions for further development.

Assumption 5: It is assumed that there exists a M, such that $0 < |A_k^0|, |B_k^0| < M$, $0 < \mu_k^0, \lambda_k^0 < \pi$, for $k = 1, \ldots, p$.

Assumption 6: The frequencies satisfy the following conditions: $\mu_1^0 \neq \mu_2^0 \neq \ldots \neq \mu_p^0$ and $\lambda_1^0 \neq \lambda_2^0 \neq \ldots \neq \lambda_p^0$.

Based on the above assumptions the WLSEs of Θ can be obtained by minimizing

$$Q(\Theta) = \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} w\left(\frac{m}{M}, \frac{n}{N}\right) \left(y(m,n) - \sum_{k=1}^{p} g_k(m,n;\theta_k) \right)^2.$$

The WLSEs of the unknown parameters will be denoted by $\{\widehat{\theta}_k = (\widehat{A}_k, \widehat{B}_k, \widehat{\mu}_k, \widehat{\lambda}_k); k = 1, \ldots, p\}$. Clearly, the WLSEs cannot be obtained in explicit forms, they have to be obtained by solving a $2p$ dimensional optimization problem. Now we state the consistency and asymptotic normality results of the WLSEs for the model (14) similar to the Theorems 3 and 4.

Theorem 5: Under Assumptions 2, 4, 5 and 6, the WLSEs of $\{(A_k^0, B_k^0, \mu_k^0, \lambda_k^0); k = 1, \ldots, p\}$ are strongly consistent.

Proof: The proof can be obtained exactly along the same line as the proof of Theorem 3, hence the details ar avoided.

Theorem 6: Under Assumptions 2, 4, 5 and 6

$$\left((\widehat{A}_1 - A_1^0, \widehat{B}_1 - B_1^0, \widehat{\mu}_1 - \mu_1^0, \widehat{\lambda}_1 - \lambda_1^0))\boldsymbol{D}, \ldots, (\widehat{A}_p - A_p^0, \widehat{B}_p - B_p^0, \widehat{\mu}_p - \mu_p^0, \widehat{\lambda}_p - \lambda_p^0)\boldsymbol{D} \right)^\top \\ \rightarrow N_{4p}\left(\mathbf{0}, 2\sigma^2 \boldsymbol{G}\right)$$

Here, $\boldsymbol{D}$ is a 4×4 diagonal matrix as defined in (10), the matrix $\boldsymbol{G}$ is a $4p \times 4p$ block diagonal matrix, with p blocks, each with order 4×4, and the k-th block $\boldsymbol{G}_k$ is of the form

$$\boldsymbol{G}_k = \boldsymbol{\Gamma}_k^{-1} \boldsymbol{\Sigma}_k \boldsymbol{\Gamma}_k^{-1}.$$

Here $\boldsymbol{\Gamma}_k$ can be obtained from $\boldsymbol{\Gamma}$ as defined in (12) by replacing A_0 and B_0 with A_k^0 and B_k^0, respectively, for $k = 1, \ldots, p$. Similarly, $\boldsymbol{\Sigma}_k$ can be obtained from $\boldsymbol{\Sigma}$ as defined in (11) for $k = 1, \ldots, p$.

Proof: See in the Appendix.

Since computation of the WLSEs of the model (14) involves solving a $2p$ dimensional optimization problem, it might be quite challenging when p is large. To avoid that heavy computation, we propose a sequential WLSEs of the unknown parameters and that involves solving p two dimensional problems sequentially. First we will provide the algorithm to compute the sequential WLSEs, and then we will show that the proposed sequential WLSEs have the same asymptotic properties as the WLSEs.

Algorithm:

Step 1: First compute the sequential WLSEs of A_1^0, B_1^0, μ_1^0 and λ_1^0 by minimizing

$$Q_1(A, B, \mu, \lambda) = \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} w\left(\frac{m}{M}, \frac{n}{N}\right) \left(y(m,n) - A\cos(m\mu + n\lambda) - B\sin(m\mu + n\lambda)\right)^2 .$$

Let the estimators be denoted by $\widetilde{A}_1$,$\widetilde{B}_1$, $\widetilde{\mu}_1$ and $\widetilde{\lambda}_1$, respectively.

Step 2: Compute the sequential WLSEs of A_2^0, B_2^0, μ_2^0 and λ_2^0 by minimizing

$$Q_2(A, B, \mu, \lambda) = \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} w\left(\frac{m}{M}, \frac{n}{N}\right) \left(\widetilde{y}(m,n) - A\cos(m\mu + n\lambda) - B\sin(m\mu + n\lambda)\right)^2 ,$$

where for $m = 1, \ldots, M$ and $n = 1, \ldots, N$,

$$\widetilde{y}(m,n) = y(m,n) - \widetilde{A}_1 \cos(m\widetilde{\mu}_1 + n\widetilde{\lambda}_1) - \widetilde{B}\sin(m\widetilde{\mu}_1 + n\widetilde{\lambda}_1)$$

Let the estimators be denoted by $\widetilde{A}_2$,$\widetilde{B}_2$, $\widetilde{\mu}_2$ and $\widetilde{\lambda}_2$, respectively.

Step 3: The process continues p times, and we obtain the sequential WLSEs of all the parameters.

It is immediate that at each step one needs to solve a two dimensional optimization problem to compute the sequential WLSEs. The following results provide the consistency and asymptotic normality properties of the sequential WLSEs.

Theorem 7: Under Assumptions 2, 4, 5 and 6, the sequential WLSEs of A_1^0,B_1^0,μ_1^0 and λ_1^0 are strongly consistent.

Proof: See Appendix B.

Theorem 8: Under Assumptions 2, 4, 5 and 6, the sequential WLSEs of A_1^0,B_1^0,μ_1^0 have the following asymptotic distribution:

$$\boldsymbol{D}\left(\widetilde{A}_1 - A_1^0, \widetilde{B}_1 - B_1^0, \widetilde{\mu}_1 - \mu_1^0, \widetilde{\lambda}_1 - \lambda_1^0\right)^\top \rightarrow N_4\left(\mathbf{0}, 2\sigma^2\mathbf{\Gamma}_1^{-1}\mathbf{\Sigma}_1\mathbf{\Gamma}_1^{-1}\right).$$

Proof: See Appendix B. ∎

Similar results can be obtained upto p steps of the sequential process.

7 Synthetic Data Analysis

In this section we have provided the analysis of a synthetic data set for illustrative purpose, mainly to show how the sequential WLSEs can be applied in analyzing a texture data. We have generated a synthetic data set from the following model:

$$y(m,n) = \sum_{k=1}^{3}\{A_k^0\cos(m\mu_k^0 + n\lambda_k^0) + B_k^0\sin(m\mu_k^0 + n\lambda_k^0) + X(m,n),$$
$$m = 1,\ldots,25, \quad n = 1,\ldots,25.$$

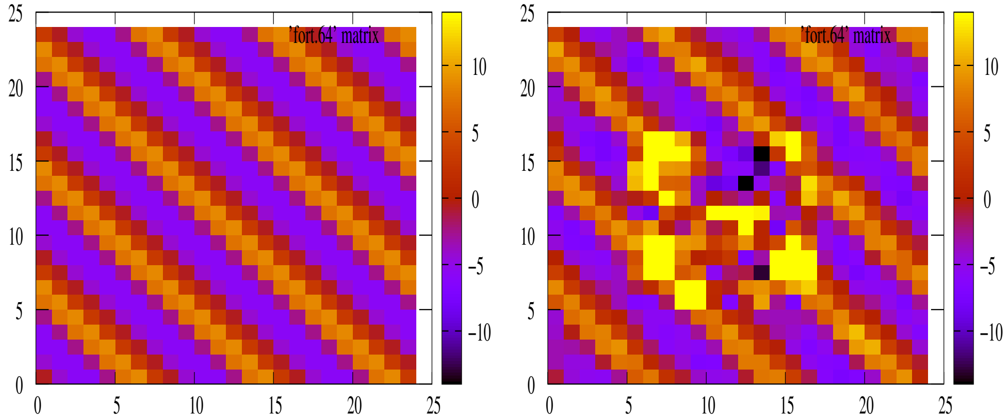
Here

$$A_1^0 = B_1^0 = 6.0, \mu_1^0 = \lambda_1^0 = 0.25\pi;$$

$$A_2^0 = B_2^0 = 5.0,\ \mu_1^0 = \lambda_1^0 = 0.75\pi;$$

$$A_3^0 = B_3^0 = 3.0,\ \mu_3^0 = \lambda_3^0 = 0.50\pi.$$

Here $X(m,n)$'s are i.i.d. Gaussian random variables with mean 0 and variance 4. There are 121 outliers in the middle portion of the data which are also i.i.d. Gaussian random



(a) (b)

Figure 1: (a) Original texture without any noise. (b) Noisy texture with outliers.

variables with mean 0 and variance 100. The original texture (without any noise) and the noisy texture are plotted in Figure 1: (a) and (b), respectively. Our problem is to extract the original texture Figure 1: (a) from the noisy texture Figure 1: (b).

Since it is quite clear from Figure 1: (b) that there are outliers in the middle portion of the data, we have used the Weight Function 1 (W1) as mentioned in Section 5. First we have first plotted $S(\mu,\lambda)$ as a function of μ and λ, where

$$S(\mu,\lambda) = \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right)\Big(y(m,n) - \widehat{A}(\mu,\lambda)\cos(m\mu+n\lambda) - \widehat{B}(\mu,\lambda)\sin(m\mu+n\lambda)\Big)^2, \tag{15}$$

where $\widehat{A}(\mu,\lambda)$ and $\widehat{B}(\mu,\lambda)$ are as defined in (4). From Figure 2 it seems $p = 3$, and hence we repeated the sequential process three times and obtained the sequential WLSEs as

$$\begin{aligned}
&\widehat{A}_1 = 6.208807, \quad \widehat{B}_1 = 5.766465 \quad \widehat{\mu}_1 = 0.249882 \quad \widehat{\lambda}_1 = 0.248472\\
&\widehat{A}_2 = 4.786206, \quad \widehat{B}_2 = 5.359531 \quad \widehat{\mu}_2 = 0.749173 \quad \widehat{\lambda}_2 = 0.750441\\
&\widehat{A}_3 = 0.742999, \quad \widehat{B}_3 = 3.640539 \quad \widehat{\mu}_3 = 0.502268 \quad \widehat{\lambda}_3 = 0.503539.
\end{aligned}$$

Based on the estimates sequential WLSEs of the unknown parameters, we have obtained the predicted texture, and it has been provided in Figure 3. It matches quite well with the

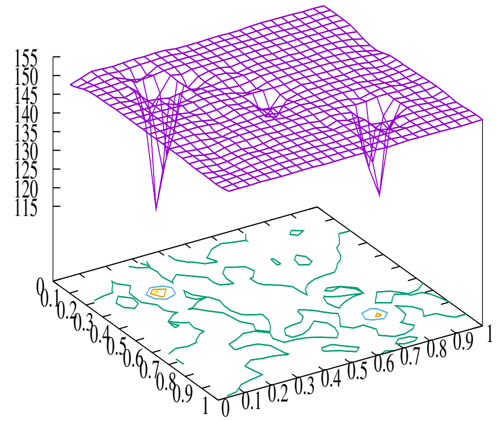


Figure 2: (a) The plot of (15) at the Fourier frequencies of the original data.

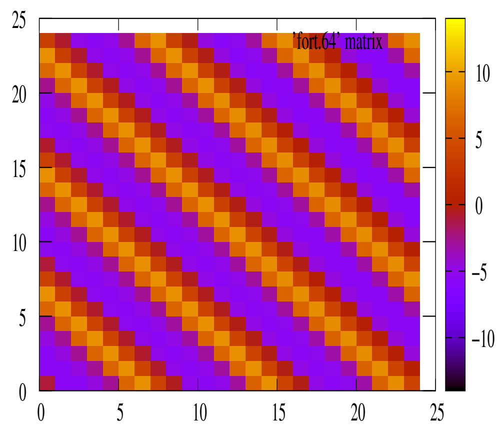


Figure 3: (a) The predicted texture bases on the sequential WLSEs.

original texture provided in Figure 1 (a).

8 Conclusions

In this manuscript we have considered 2-D sinusoidal model in presence of additive noise, and we have used the WLSEs to estimate the unknown parameters. It is observed that the WLSEs can be used as an alternative to the robust estimators as it is easy to implement in practice and the properties of the WLSEs can be established in somewhat weaker assumptions than the existing robust estimators. The results have been extended to multicomponent model also, and we have proposed a sequential WLSEs which are very easy to implement in practice. It is observed that the proposed WLSEs behave better than the ordinary LSEs in presence of outliers, and the performances are quite comparable with the robust LADEs.

It may be mentioned that the choice of the weight function is an important problem. If we have some idea about the region where the outliers might be present, then the weight function can be chosen accordingly. In that case one can choose a weight function which puts less weight in that region. In our extensive simulation experiments it has been observed that the above method works quite well. But the problem becomes complicated if one does not have any idea about the region, where the outliers might be present. It is clear that it has to be a data dependent approach in this case. In such a situation, a two-step procedure may be used. In the first step obtain the estimators of the unknown parameters by using some methods, may be by the least squares or some other non-robust methods, and from the residuals get some idea about the region where the outliers might be present. At the second step use this knowledge to choose the weight function, and obtain robust estimators of the unknown parameters based on the weighted least squares method. It is definitely an important problem, and we believe more systematic study is needed along this direction.

Acknowledgements:

The authors would like to thank Ms. Aditi Sharma for her help to prepare this manuscript. The authors would also like to thank the unknown reviewers, the Associate Editor and the handling editor Prof. Michel Kieffer for their constructive suggestions, which have helped to improve the manuscript significantly.

Conflict of Interests Statement

None of the authors has any conflict of interest in preparation of this manuscript.

Appendix A

To prove Theorem 1, we need the following result.

Lemma 1: If $\{X(m,n); m,n \in Z\}$ satisfy Assumption 2 and $w(s,t)$ satisfies Assumption 3, then

$$\sup_{\alpha,\beta}\left|\frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}X(m,n)w\left(\frac{m}{M},\frac{n}{N}\right)\cos(m\alpha)\cos(n\beta)\right| \to 0 \qquad \min\{M,N\}\to\infty.$$

Proof: Consider the following random variables

$$Z(m,n) = \begin{cases} X(m,n) & \text{if} \quad |X(m,n)| < (mn)^{3/4} \\ 0 & \text{if} \quad |X(m,n)| \geq (mn)^{3/4} \end{cases}$$

First we will show that $X(m,n)$ and $Z(m,n)$ are equivalent sequences. Consider

$$\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}P(X(m,n)\neq Z(m,n)) = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty}P\{|X(m,n)|\geq (mn)^{3/4}\}.$$

Now observe that there are at most $2^k k$ combinations of $(m,n)'s$ such that $mn < 2^k$, therefore we have

$$\begin{aligned}
&\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}P\left(|X(m,n)|\geq (mn)^{3/4}\right) \\
&\quad \leq \sum_{k=1}^{\infty}\sum_{2^{k-1}\leq r<2^k}P\left(|X(m,n)|\geq r^{3/4}\right) \quad [\text{here } r = mn] \\
&\quad \leq \sum_{k=1}^{\infty}k2^k P\left(|X(1,1)|\geq 2^{(k-1)3/4}\right) \\
&\quad \leq C\sum_{k=1}^{\infty}k2^k\frac{E|X(1,1)|^2}{2^{(k-1)3/2}} \leq C\sum_{k=1}^{\infty}\frac{k}{2^{k/2}} < \infty.
\end{aligned}$$

Here C is a constant and it may represent different constants and different places. Therefore, $X(m,n)$ and $Z(m,n)$ are equivalent sequences. So

$$P\left(X(m,n) \neq Z(m,n) \quad i.o\right) = 0.$$

Therefore, if we can show that

$$\sup_{\alpha,\beta}\left|\frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N} Z(m,n) w\left(\frac{m}{M},\frac{n}{N}\right)\cos(m\alpha)\cos(n\beta)\right| \to 0 \qquad \min\{M,N\}\to\infty,$$

we are done. Now consider $U(m,n) = Z(m,n) - E(Z(m,n))$, then

$$\sup_{\alpha,\beta}\left|\frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N} E(Z(m,n)) w\left(\frac{m}{M},\frac{n}{N}\right)\cos(m\alpha)\cos(n\beta)\right| \le \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\left|E(Z(m,n))\right|.$$

Since $E(Z(m,n)) \to 0$ as $M,N\to\infty$, therefore, as $M,N\to\infty$,

$$\frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\left|E(Z(m,n))\right| \to 0.$$

Therefore, if we can show

$$\sup_{\alpha,\beta}\left|\frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N} U(m,n) w\left(\frac{m}{M},\frac{n}{N}\right)\cos(m\alpha)\cos(n\beta)\right| \to 0 \qquad \min\{M,N\}\to\infty,$$

we are done. Now for any fixed $\epsilon > 0$, $0 < \alpha,\beta < \pi$ and $0 < h \le \dfrac{1}{2(MN)^{3/4}}$, we have

$$\begin{aligned} &P\left\{\left|\frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N} U(m,n) w\left(\frac{m}{M},\frac{n}{N}\right)\cos(m\alpha)\cos(n\beta)\right| \ge \epsilon\right\} \\ &\quad \le 2e^{-hMN\epsilon}\prod_{m=1}^{\infty}\prod_{n=1}^{\infty} E\exp\left(hU(m,n)w\left(\frac{m}{M},\frac{n}{N}\right)\cos(m\alpha)\cos(n\beta)\right). \end{aligned}$$

Since, $\left|hU(m,n)w\left(\frac{m}{M},\frac{n}{N}\right)\cos(m\alpha)\cos(n\beta)\right| \le 1/2$, using $e^x < 1 + x + x^2$, for $|x| < 1/2$, we have

$$\le 2e^{-hMN\epsilon}\prod_{m=1}^{\infty}\prod_{n=1}^{\infty} E\exp\left(hU(m,n)w\left(\frac{m}{M},\frac{n}{N}\right)\cos(m\alpha)\cos(n\beta)\right) \le 2e^{-hMN\epsilon}(1+h^2\sigma^2)^{MN}.$$

Now choose $h =\leq \dfrac{1}{2(MN)^{3/4}}$, therefore, for large M and N

$$P\left\{\left|\frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}U(m,n)w\left(\frac{m}{M},\frac{n}{N}\right)\cos(m\alpha)\cos(n\beta)\right|\geq\epsilon\right\}\leq Ce^{-(MN)^{1/4}\epsilon/2}e^{\sigma^2/4}.$$

Let $K = M^2N^2$, choose K points, $\theta_1 = (\alpha_1, \beta_1), \ldots, (\alpha_K, \beta_K)$, such that for each point $\theta = (\alpha, \beta) \in (0, \pi) \times (0, \pi)$, there is a point θ_j, satisfying

$$|\alpha_j - \alpha| + |\beta_j - \beta| \leq \frac{\pi}{M^2N^2}.$$

Observe that

$$\left|\frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}U(m,n)w\left(\frac{m}{M},\frac{n}{N}\right)\left(\cos(m\alpha)\cos(n\beta)-\cos(m\alpha_j)\cos(n\beta_j)\right)\right|\leq$$
$$\frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\left|U(m,n)w\left(\frac{m}{M},\frac{n}{N}\right)\right|\left|(\cos(m\alpha)\cos(n\beta)-\cos(m\alpha_j)\cos(n\beta_j))\right|\leq$$
$$\frac{C}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\frac{MN}{M^2N^2}(m+n)\to 0,\quad \text{as}\quad M,N\to\infty.$$

Therefore, for large M and N, we have

$$\begin{aligned}
&P\left\{\sup_{\alpha,\beta}\left|\frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}U(m,n)w\left(\frac{m}{M},\frac{n}{N}\right)\cos(m\alpha)\cos(n\beta)\right|\geq 2\epsilon\right\}\\
&\quad\leq P\left\{\max_{j\leq M^2N^2}\left|\frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}U(m,n)w\left(\frac{m}{M},\frac{n}{N}\right)\cos(m\alpha_j)\cos(n\beta_j)\right|\geq\epsilon\right\}\\
&\quad\leq CM^2N^2e^{-(MN)^{1/4}/2}.
\end{aligned}$$

Since $\displaystyle\sum_{t=1}^{\infty}t^2e^{-t^{3/4}}<\infty$,

$$\sup_{\alpha,\beta}\left|\frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}U(m,n)w\left(\frac{m}{M},\frac{n}{N}\right)\cos(m\alpha_j)\cos(n\beta_j)\right|\quad\to 0\qquad a.s.$$

Similarly, it can be shown that

Lemma 2: If $\{X(m,n); m,n\in Z\}$ satisfy Assumption 2 and $w(s,t)$ satisfies Assumption 3, then

$$\sup_{\alpha,\beta}\left|\frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}X(m,n)w\left(\frac{m}{M},\frac{n}{N}\right)\cos(m\alpha)\sin(n\beta)\right|\to 0 \qquad \min\{M,N\}\to\infty,$$
$$\sup_{\alpha,\beta}\left|\frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}X(m,n)w\left(\frac{m}{M},\frac{n}{N}\right)\sin(m\alpha)\sin(n\beta)\right|\to 0 \qquad \min\{M,N\}\to\infty.$$

Let us use the following notations: $\boldsymbol{\theta}=(A,B,\mu,\lambda)^\top$, $\boldsymbol{\theta}^0=(A^0,B^0,\mu^0,\lambda^0)^\top$

$$f(m,n;\boldsymbol{\theta})=A\cos(m\mu+n\lambda)+B\sin(m\mu+n\lambda).$$

Therefore, for any given A and B such that $|A|<M$ and $|B|<M$, using Lemma 1 and Lemma 2 it immediately follows that

$$\sup_{\alpha,\beta}\left|\frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}X(m,n)w\left(\frac{m}{M},\frac{n}{N}\right)f(m,n;\boldsymbol{\theta})\right|\to 0 \qquad \min\{M,N\}\to\infty. \tag{16}$$

Lemma 3: For any given $\delta>0$ using (16)

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

Proof: Consider

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

For any given $\delta>0$ using (16), it follows

$$\liminf_{M,N\to\infty}\inf_{|\boldsymbol{\theta}^0-\boldsymbol{\theta}|>\delta}\left\{Q(\boldsymbol{\theta})-Q(\boldsymbol{\theta}^0)\right\}=$$

$$\liminf_{M,N\to\infty} \inf_{|\boldsymbol{\theta}^0-\boldsymbol{\theta}|>\delta} \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right)\left(f(m,n;\boldsymbol{\theta}^0)-f(m,n;\boldsymbol{\theta})\right)^2 \geq$$
$$\liminf_{M,N\to\infty} \inf_{|\boldsymbol{\theta}^0-\boldsymbol{\theta}|>\delta} \frac{\gamma}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\left(f(m,n;\boldsymbol{\theta}^0)-f(m,n;\boldsymbol{\theta})\right)^2.$$

Consider the following sets:

$$\Gamma_1 = \{\boldsymbol{\theta} : \boldsymbol{\theta} \in \boldsymbol{\Theta}, |A - A^0| > \delta/2\}, \quad \Gamma_2 = \{\boldsymbol{\theta} : \boldsymbol{\theta} \in \boldsymbol{\Theta}, |B - B^0| > \delta/2\},$$

$$\Gamma_3 = \{\boldsymbol{\theta} : \boldsymbol{\theta} \in \boldsymbol{\Theta}, |\mu - \mu^0| > \delta/2\}, \quad \Gamma_4 = \{\boldsymbol{\theta} : \boldsymbol{\theta} \in \boldsymbol{\Theta}, |\lambda - \lambda^0| > \delta/2\}.$$

Note that

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

Observe that

$$\liminf_{M,N\to\infty} \inf_{\boldsymbol{\theta}\in\Gamma_1} \frac{\gamma}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\left(f(m,n;\boldsymbol{\theta}^0)-f(m,n;\boldsymbol{\theta})\right)^2 =$$
$$\liminf_{M,N\to\infty} \frac{\gamma|A-A^0|^2}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\cos^2(m\mu^0+n\lambda^0) = \frac{\gamma|A-A^0|^2}{2} > 0.$$

Similarly, it can be shown that for $j = 2,3,4$,

$$\liminf_{M,N\to\infty} \inf_{\boldsymbol{\theta}\in\Gamma_j} \frac{\gamma}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\left(f(m,n;\boldsymbol{\theta}^0)-f(m,n;\boldsymbol{\theta})\right)^2 > 0.$$

Since,

$$\liminf_{M,N\to\infty} \inf_{|\boldsymbol{\theta}^0-\boldsymbol{\theta}|>\delta} \left\{Q(\boldsymbol{\theta})-Q(\boldsymbol{\theta}^0)\right\} \geq \liminf_{M,N\to\infty} \inf_{\boldsymbol{\theta}\in\cup_{j=1}^4\Gamma_j} \frac{\gamma}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\left(f(m,n;\boldsymbol{\theta}^0)-f(m,n;\boldsymbol{\theta})\right)^2,$$

the result follows.

Proof of Theorem 1: Let us denote $\widehat{\boldsymbol{\theta}} = (\widehat{A}, \widehat{B}, \widehat{\mu}, \widehat{\lambda})$ as the least squares estimate $\boldsymbol{\theta}$. In this proof only we denote it by $\widehat{\boldsymbol{\theta}}_{MN}$ to emphasize that it is a function of M and N. Note that for any M and N

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

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

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

this contradicts Lemma 3. Hence, $\widehat{\boldsymbol{\theta}}_{MN}$ is a consistent estimate of $\boldsymbol{\theta}^0$.

Proof of Theorem 2: To prove this Theorem, we need the Taylor series expansion of

$$S(\boldsymbol{\theta}) = MNQ(\boldsymbol{\theta}) = \sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M}, \frac{n}{N}\right)\left(y(m,n) - A\cos(m\mu + n\lambda) - B\sin(m\mu + n\lambda)\right)^2,$$

as follows

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

Here

$$S'(\boldsymbol{\theta}) = \left(\frac{\partial S(\boldsymbol{\theta})}{\partial A}, \frac{\partial S(\boldsymbol{\theta})}{\partial B}, \frac{\partial S(\boldsymbol{\theta})}{\partial \mu}, \frac{\partial S(\boldsymbol{\theta})}{\partial \lambda}\right),$$

$$S''(\boldsymbol{\theta}) = \begin{bmatrix} \frac{\partial^2 S(\boldsymbol{\theta})}{\partial A^2} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial A \partial B} & \frac{\partial^2 S(\boldsymbol{\theta})}{\partial A \partial \mu} & \frac{\partial^2 S(\boldsymbol{\theta})}{\partial A \partial \lambda} \\ \frac{\partial^2 S(\boldsymbol{\theta})}{\partial B \partial A} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial B^2} & \frac{\partial^2 S(\boldsymbol{\theta})}{\partial B \partial \mu} & \frac{\partial^2 S(\boldsymbol{\theta})}{\partial B \partial \lambda} \\ \frac{\partial^2 S(\boldsymbol{\theta})}{\partial \mu \partial A} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial \mu \partial B} & \frac{\partial^2 S(\boldsymbol{\theta})}{\partial \mu^2} & \frac{\partial^2 S(\boldsymbol{\theta})}{\partial \mu \partial \lambda} \\ \frac{\partial^2 S(\boldsymbol{\theta})}{\partial \lambda \partial A} & \frac{\partial^2 Q(\boldsymbol{\theta})}{\partial \lambda \partial B} & \frac{\partial^2 S(\boldsymbol{\theta})}{\partial \lambda \partial \mu} & \frac{\partial^2 S(\boldsymbol{\theta})}{\partial \lambda^2} \end{bmatrix}$$

and $\overline{\boldsymbol{\theta}}$ is a point on the line joining $\widehat{\boldsymbol{\theta}}$ and $\boldsymbol{\theta}^0$. Since $S'(\widehat{\boldsymbol{\theta}}) = \mathbf{0}$, (18) can be written as

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

here $\boldsymbol{D}$ is the diagonal matrix as defined in (10). Now observe that

$$\begin{aligned}
\frac{\partial S(\boldsymbol{\theta}^0)}{\partial A} &= -2\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M}, \frac{n}{N}\right) X(m,n)\cos(m\mu^0 + n\lambda^0) \\
\frac{\partial S(\boldsymbol{\theta}^0)}{\partial B} &= -2\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M}, \frac{n}{N}\right) X(m,n)\sin(m\mu^0 + n\lambda^0) \\
\frac{\partial S(\boldsymbol{\theta}^0)}{\partial \mu} &= 2\sum_{m=1}^{M}\sum_{n=1}^{N} mw\left(\frac{m}{M}, \frac{n}{N}\right) X(m,n)\sin(m\mu^0 + n\lambda^0) \\
\frac{\partial S(\boldsymbol{\theta}^0)}{\partial A} &= -2\sum_{m=1}^{M}\sum_{n=1}^{N} nw\left(\frac{m}{M}, \frac{n}{N}\right) X(m,n)\cos(m\mu^0 + n\lambda^0),
\end{aligned}$$

similarly,

$$
\begin{aligned}
\frac{\partial^2 S(\boldsymbol{\theta}^0)}{\partial A^2} &= 2\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right)\cos^2(m\mu^0+n\lambda^0)\\
\frac{\partial^2 S(\boldsymbol{\theta}^0)}{\partial B^2} &= 2\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right)\sin^2(m\mu^0+n\lambda^0)\\
\frac{\partial^2 S(\boldsymbol{\theta}^0)}{\partial \mu^2} &= 2\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right)\big\{X(m,n)(m^2A^0\cos(m\mu^0+n\lambda^0)+m^2B^0\sin(m\mu^0+n\lambda^0)+\\
&\quad (mA^0\sin(m\mu^0+n\lambda^0)-mB^0\cos(m\mu^0+n\lambda^0))^2\big\}\\
\frac{\partial^2 S(\boldsymbol{\theta}^0)}{\partial \lambda^2} &= 2\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right)\big\{X(m,n)(n^2A^0\cos(m\mu^0+n\lambda^0)+n^2B^0\sin(m\mu^0+n\lambda^0)+\\
&\quad (nA^0\sin(m\mu^0+n\lambda^0)-nB^0\cos(m\mu^0+n\lambda^0))^2\big\}\\
\frac{\partial^2 S(\boldsymbol{\theta}^0)}{\partial A\partial B} &= 2\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right)\cos(m\mu^0+n\lambda^0)\sin(m\mu^0+n\lambda^0)\\
\frac{\partial^2 S(\boldsymbol{\theta}^0)}{\partial A\partial \mu} &= 2\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right)\big\{X(m,n)(m\sin(m\mu^0+n\lambda^0))-\\
&\quad \cos(m\mu^0+n\lambda^0)(mA^0\sin(m\mu^0+n\lambda^0)-mB^0\cos(m\mu^0+n\lambda^0))\big\}\\
\frac{\partial^2 S(\boldsymbol{\theta}^0)}{\partial A\partial \lambda} &= 2\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right)\big\{X(m,n)(n\sin(m\mu^0+n\lambda^0))-\\
&\quad \cos(m\mu^0+n\lambda^0)(nA^0\sin(m\mu^0+n\lambda^0)-nB^0\cos(m\mu^0+n\lambda^0))\big\}\\
\frac{\partial^2 S(\boldsymbol{\theta}^0)}{\partial B\partial \mu} &= 2\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right)\big\{-X(m,n)(m\cos(m\mu^0+n\lambda^0))-\\
&\quad \sin(m\mu^0+n\lambda^0)(mA^0\sin(m\mu^0+n\lambda^0)-mB^0\cos(m\mu^0+n\lambda^0))\big\}\\
\frac{\partial^2 S(\boldsymbol{\theta}^0)}{\partial B\partial \lambda} &= 2\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right)\big\{-X(m,n)(n\cos(m\mu^0+n\lambda^0))-\\
&\quad \sin(m\mu^0+n\lambda^0)(nA^0\sin(m\mu^0+n\lambda^0)-nB^0\cos(m\mu^0+n\lambda^0))\big\}\\
\frac{\partial^2 S(\boldsymbol{\theta}^0)}{\partial \mu\partial \lambda} &= 2\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right)\big\{X(m,n)mn(A^0\cos(m\mu^0+n\lambda^0)+B^0\sin(m\mu^0+n\lambda^0))+\\
&\quad mn(A^0\sin(m\mu^0+n\lambda^0)-B^0\cos(m\mu^0+n\lambda^0))^2\big\}
\end{aligned}
$$

It follows that

$$S'(\boldsymbol{\theta}^0)\boldsymbol{D} \to N_4(\mathbf{0}, 2\sigma^2\boldsymbol{\Sigma}).$$

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

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

hence the result follows. ∎

Proof of Theorem 4: It may be seen that if we can establish (5), (6) and (7) for a general weight function $w(s,t)$ satisfying Assumption 4, then we are done. We will show (5), the other two can be shown along the same way. We need to show that if $w(s,t)$ satisfies Assumption 4, then

$$\lim_{M,N\rightarrow\infty} \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} w\left(\frac{m}{M}, \frac{n}{N}\right) \cos^2(m\mu + n\lambda) = \frac{1}{2} \int_0^1 \int_0^1 w(s,t) ds dt. \tag{20}$$

To show (20) note that for a given $\epsilon > 0$, there exists a two-dimensional polynomial $p_\epsilon(x,y)$ such that $|w(s,t) - p_\epsilon(x,y)| \leq \epsilon$, for all $x, y \in [0,1]$. Hence, we have

$$\int_0^1 \int_0^1 p_\epsilon(s,t) ds dt - \epsilon \leq \int_0^1 \int_0^1 w(s,t) ds dt \leq \int_0^1 \int_0^1 p_\epsilon(s,t) ds dt + \epsilon, \tag{21}$$

$$\begin{aligned} &\frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} p_\epsilon\left(\frac{m}{M}, \frac{n}{N}\right) \cos^2(m\mu + n\lambda) - \epsilon \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} \cos^2(m\mu + n\lambda) \leq \\ &\frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} w\left(\frac{m}{M}, \frac{n}{N}\right) \cos^2(m\mu + n\lambda) \leq \\ &\frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} p_\epsilon\left(\frac{m}{M}, \frac{n}{N}\right) \cos^2(m\mu + n\lambda) + \epsilon \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} \cos^2(m\mu + n\lambda) \end{aligned} \tag{22}$$

Taking limits $M, N \rightarrow \infty$ in (22), and using (5) and (21), (20) can be obtained. ∎

Proof of Theorem 6: Note that for each component the asymptotic distribution follows exactly along the same line as the proof of Theorem 2 or Theorem 4. The only difference in this case is that we need to show that the asymptotic covariance matrix of the two different components is a zero matrix. Note that it mainly follows from the following results for

$1 \leq j \neq k \leq p$,

$$\lim_{M,N\to\infty} \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} w\left(\frac{m}{M}, \frac{n}{N}\right) \cos(m\mu_k + n\lambda_k) \cos(m\mu_j + n\lambda_j) = 0$$
$$\lim_{M,N\to\infty} \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} w\left(\frac{m}{M}, \frac{n}{N}\right) \sin(m\mu_k + n\lambda_k) \sin(m\mu_j + n\lambda_j) = 0$$
$$\lim_{M,N\to\infty} \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} w\left(\frac{m}{M}, \frac{n}{N}\right) \sin(m\mu_k + n\lambda_k) \cos(m\mu_j + n\lambda_j) = 0.$$

Appendix B

We need the following Lemma to prove Theorem 7.

Lemma 4: Let us denote

$$S_{1c} = \{\theta : \theta = (A, B, \mu, \lambda)^\top, |\theta - \theta_1^0| \geq 4c\}.$$

If there exists a $c > 0$, such that

$$\underline{\lim} \inf_{\theta \in S_{1c}} [Q_1(\theta) - Q_1(\theta_1^0)] > 0 \qquad a.s, \tag{23}$$

then $\widetilde{\theta}_1$ that minimizes $Q_1(\theta)$, is a strongly consistent estimator of θ_1^0.

Proof: It follows by contradiction using simple arguments. ∎

Proof of Theorem 7: Let us consider the quantity $[Q_1(\theta) - Q_1(\theta^0)]$. Since

$$\begin{aligned} Q_1(\theta) &= \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} w\left(\frac{m}{M}, \frac{n}{N}\right) (y(m,n) - g_1(m,n;\theta))^2 \\ Q_1(\theta^0) &= \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} w\left(\frac{m}{M}, \frac{n}{N}\right) (X(m,n) + \sum_{k=2}^{p} g_k(m,n;\theta))^2, \end{aligned}$$

$$Q_1(\theta) - Q_1(\theta^0) = f_{11}(\theta) + f_{21}(\theta) + f_{31}(\theta).$$

Here

$$\begin{aligned} f_{11}(\theta) &= \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N} w\left(\frac{m}{M},\frac{n}{N}\right)(g_1(m,n;\theta^0)-g_1(m,n;\theta))^2 + \\ &\quad \frac{2}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\left[w\left(\frac{m}{M},\frac{n}{N}\right)(g_1(m,n;\theta^0)-g_1(m,n;\theta))\sum_{k=2}^{p} g_k(m,n;\theta)\right] \\ f_{21}(\theta) &= \frac{2}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\left[w\left(\frac{m}{M},\frac{n}{N}\right)X(m,n)(g_1(m,n;\theta^0)-g_1(m,n;\theta))\right] \end{aligned}$$

Using Lemma 1 and Lemma 2, it follows that

$$\sup_{\theta\in S_{1c}} |f_{21}(\theta)| \stackrel{a.s.}{\rightarrow} 0.$$

Using a length and straight forward calculations, splitting the set S_{1c} as in Theorem 1, it can be shown that

$$\underline{\lim} \inf_{\theta\in S_{1c}} f_{11}(\theta) > 0 \qquad a.s.$$

Hence, (23) is satisfied. This proves Lemma 7. ∎

Proof of Theorem 8: This proof can be obtained following similar line as in Theorem 2. In this case using the Taylor series expansion of $S_1(\theta) = MNQ_1(\theta)$ around the point θ^0 and following similar line along the proof of Theorem 2, the required result can be obtained. ∎

References

[1] Fang-Jiong C., Carrson C. F., Chi-Wah K., Sam K. (2007), "Estimation of Two-Dimensional Frequencies Using Modified Matrix Pencil Method", *IEEE Transactions on Signal Processing*, vol. 55, 2, 718 - 724.

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

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

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

[5] Hou, Y., Cuyt, A., Lee, W-S and Bhowmik, D. (2021), "Decomposing Textures using Exponential Analysis", *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Toronto, ON, Canada, pp. 1920–1924.

[6] Ivanov, A.V. and Lymar, O.V. (2020a), "Asymptotic normality of the least squares estimator of two-dimensional sinusoidal observation model parameters ", *Theory of Probability and Mathematical Statistics*, No. 100, 107 – 131.

[7] Ivanov, A. V. and Lymar, O.V. (2020b) "Asymptotic properties of periodogram parameter estimators for a trigonometric observation model on the plane", *Theory of Probability and Mathematical Statistics*, No. 101, 129 – 151.

[8] Ivanov, A.V. and Savych, I.M. (2021), "On the least squares estimator asymptotic normality of the multivariate symmetric textured surface parameters", *. Theory of Probability and Mathematical Statistics*, No. 105, 151-169.

[9] Kim, T.S., Kim, H.K. and Choi, H.C. (2000), "Asymptotic properties of LAD estimators of a nonlinear time series regression model", *Journal of Korean Statistical Society*, vol. 29, No. 2, 187–199.

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

[11] Marchant, R., Jackway, P. (2017), "Sinusoidal Image Model Derived from the Circular Harmonic Vector", *Journal of Mathematical Imaging and Vision*, vol. 57, 138–163.

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

[13] Porwal, A., Mitra, S. and Mitra, A. (2019), "Order estimation of 2-dimensional complex superimposed exponential signal model using exponentially embedded-family (EEF) rule: large sample consistency properties", *Multidimensional Systems and Signal Processing*, vol. 30, 1293–1308.

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

[15] Rao, C.R., Zhao, L. and Zhou, B., (1994), "Maximum likelihood estimation of 2-D superimposed exponential signals", *IEEE Transactions on Signal Processing*, Vol. 42, 7, 1795–1802.

[16] Zhang, H. and Mandrekar, V. (2001), "Estimation of hidden frequencies for 2D stationary processes", *Journal of Time Series Analysis*, vol. 22, no. 5, 613–629.