Asymptotic Confidence Intervals of Spectra and Cross-spectra of Continuous Time Processes Based on Uniformly Spaced Samples

Technical Report No. ASD/2011/2

Dated: 14 April 2011

Radhendushka Srivastava

And

Debasis Sengupta

Indian Statistical Institute

Applied Statistics Unit

Kolkata 700 108

Asymptotic Confidence Intervals of Spectra and Cross-spectra of Continuous Time Processes Based on Uniformly Spaced Samples

Radhendushka Srivastava and Debasis Sengupta *

April 14, 2011

Abstract

It is well known that if the power spectrum of a continuous time stationary stochastic process is not limited to a finite interval, data sampled from that process at any uniform sampling rate leads to biased and inconsistent spectrum estimators. As a result, depending on the range of frequency of interest, one uses anti-alias filtering and subsequently treats the filtered process as bandlimited. However, if the spectrum has a large peak in the 'stop band' of the filter, and the filter is not an ideal one, then spectrum estimation would be problematic, even if the sample size is large. In this paper, we use the well-known smoothed periodogram estimator to construct asymptotic confidence intervals shrinking to the true spectra, by taking into account the decay rate of the filtered process and allowing the sampling rate to go to infinity suitably fast as the sample size goes to infinity. The proposed method requires minimal computation, as it does not involve bootstrap or other resampling. The method is illustrated through a Monte-Carlo simulation study, and its performance is compared with conventional methods based on uniform and Poisson sampling.

Keywords: Power spectral density, spectrum estimation, smoothed periodogram, shrinking asymptotics.

*The authors are with the Applied Statistics Unit, Indian Statistical Institute, 203 B.T. Road, Kolkata 700108, India; e-mail: radhe_r@isical.ac.in, sdebasis@isical.ac.in.

1 Introduction

Estimation of power spectral density (spectrum) of a continuous time, mean square continuous, stationary stochastic process is a classical problem. Generally the estimation is based on finitely many samples of the process. It is well known that if the spectrum is compactly supported (bandlimited), then it can be estimated consistently from uniformly spaced samples, provided the sampling is done at the Nyquist rate or faster [1, 2]. Many methods for constructing estimators as well as confidence intervals of bandlimited spectra have been proposed [3, 4, 5]. For sampled non-bandlimited processes, or bandlimited processes sampled at sub-Nyquist rate, the problem of aliasing leads to biased estimation. For this reason, it is often argued that uniformly spaced samples should not be used for the purpose of estimation of a non-bandlimited spectrum [6, 7, 8].

Consequently, some researchers have turned to non-uniform sampling schemes such as stochastic sampling and periodic non-uniform sampling. Masry [8] proved the consistency of some spectrum estimators based on stochastic sampling schemes, under appropriate conditions that allow the underlying spectrum to be non-bandlimited. Poisson sampling, i.e., sampling at the jump points of the Poisson process, is a prominent example of stochastic sampling, and Masry's approach produces an explicit estimator in this case. This work has led to the construction of asymptotic confidence intervals [9] through randomly sampled data. In yet another attempt to overcome the limitations posed by the Nyquist theorem, it has been shown that for multiband processes, periodic non-uniform sampling at sub-Nyquist average rate can lead to appropriate reconstruction of the original process [10]. This work has been followed by substantial further research in the area, including consistent estimation of the underlying spectrum [11, 12].

In many applications, even if the spectrum is non-bandlimited, one is interested only in estimating the spectrum over a specific range of frequencies. In such a situation, it is common practice to use an analog low pass filter with a cut-off frequency that is higher than the highest frequency of interest, before sampling the process at a uniform rate and using those samples for spectrum estimation. One chooses the sampling rate according to the Nyquist theorem, assuming that the bandwidth of the filtered process is the cut-off frequency of the analog filter. However, it is important to note that no implementable low pass filter is ideal. On the other hand, implementable low pass filters usually have a known rate of decay. Therefore, instead of assuming that the spectrum of the filtered process is bandlimited, it is more realistic to assume that the spectrum is possibly non-bandlimited with a known rate of decay.

In this paper, we consider non-bandlimited processes with a known rate of decay and obtain asymptotic confidence intervals of the spectrum based on uniformly spaced samples. These intervals are based on a commonly used nonparametric spectrum estimator, namely the smoothed periodogram [4]. In order to ensure that the confidence intervals shrink to the true spectrum, it is necessary to arrange for the bias to go to zero with increasing sample size. For this purpose, we build on the work of Srivastava and Sengupta [13], who have shown how the bias can be controlled, and the spectrum can be estimated consistently, through progressively faster uniform sampling. The strategy of letting the sampling rate increase indefinitely along with the sample size is justified by arguing that if one has the resources to increase the sample size indefinitely, one would like to use some of those resources to sample faster, rather than being constrained by a fixed sampling rate. The asymptotic approach chosen here (referred to as 'shrinking asymptotics' by Fuentes [14]) was also adopted by other authors [15], [16], [17], [18] although asymptotic confidence intervals for the power spectral density have not been studied previously. This approach is different from the 'fixed-domain asymptotics' or 'infill asymptotics' approach [19], [20], [21], [22] which, in the present case, would have required that the time-span of the original continuous-time data (before sampling) remains fixed as the sampling rate goes to infinity.

For non-bandlimited processes with a known rate of decay, the sampling-cum-inference strategy proposed in this paper provides an alternative to renewal process sampling followed by construction of confidence intervals as proposed by Lii and Masry [9].

Let $\mathbf{X} = \{\mathbf{X}(t),\ -\infty < t < \infty\}$ be an r-dimensional mean square continuous stationary stochastic process, having zero mean. We denote the components of the process $\mathbf{X}$ by $X_a = \{X_a(t),\ -\infty < t < \infty\}$ for $a \in \{1, 2, \dots, r\}$, and the variance-covariance matrix of the process $\mathbf{X}$ at lag τ by

$$\mathbf{C}(\tau) = \begin{pmatrix} C_{11}(\tau) & C_{12}(\tau) & \dots & C_{1r}(\tau) \\ C_{21}(\tau) & C_{22}(\tau) & \dots & C_{2r}(\tau) \\ \vdots & \vdots & \ddots & \vdots \\ C_{r1}(\tau) & C_{r2}(\tau) & \dots & C_{rr}(\tau) \end{pmatrix},$$

where

$$C_{a_1 a_2}(\tau) = E\left[X_{a_1}(t+\tau)X_{a_2}(t)\right] \text{ for } a_1, a_2 \in \{1, 2, \dots, r\}.$$

The spectral and cross-spectral density matrix of the process $\mathbf{X}$ is denoted by

$$\mathbf{\Phi}(\cdot) = \begin{pmatrix} \phi_{11}(\cdot) & \phi_{12}(\cdot) & \dots & \phi_{1r}(\cdot) \\ \phi_{21}(\cdot) & \phi_{22}(\cdot) & \dots & \phi_{2r}(\cdot) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{r1}(\cdot) & \phi_{r2}(\cdot) & \dots & \phi_{rr}(\cdot) \end{pmatrix},$$

where

$$\phi_{a_1a_2}(\lambda) = \frac{1}{2\pi}\int_{-\infty}^{\infty} C_{a_1a_2}(t)e^{-it\lambda}dt, \quad \text{for } -\infty < \lambda < \infty, \quad a_1, a_2 \in \{1, 2, \dots, r\}.$$

In this paper, we construct confidence intervals of $\phi_{a_1a_2}(\lambda)$ for a_1 , $a_2 \in \{1, 2, \dots, r\}$ based on the following estimator:

$$\widehat{\phi}_{a_1a_2}(\lambda) = \frac{1}{2\pi n\rho_n}\sum_{t_1=1}^{n}\sum_{t_2=1}^{n} K(b_n(t_1-t_2))X_{a_1}\left(\frac{t_1}{\rho_n}\right)X_{a_2}\left(\frac{t_2}{\rho_n}\right)e^{-\frac{i(t_1-t_2)\lambda}{\rho_n}}1_{[-\pi\rho_n,\pi\rho_n]}(\lambda), \tag{1}$$

where $K(\cdot)$ is a covariance averaging kernel, b_n is the kernel bandwidth, ρ_n is the sampling rate and $1_A(\lambda)$ is the indicator of the event $\lambda \in A$.

Consistency and weak convergence of the estimators are crucial to the construction of confidence intervals shrinking to the true spectra and cross-spectra. In Section 2, we establish the consistency of the spectrum estimator (1) for non-bandlimited processes. This result is a generalization of a result of Srivastava and Sengupta [13] to the case of multivariate time series. It also paves the way for computation of the asymptotic distribution of the estimator. A formal description of the confidence intervals is given in Section 3, which also contains some discussion on optimal rates of shrinkage of these intervals. In Section 4, we carry out a Monte Carlo simulation study to demonstrate how the proposed approach of progressively faster sampling can improve the coverage probability of confidence intervals based on uniform sampling. We also compare the performances of the proposed sampling-cum-inference approach with that of Lii and Masry [9], in terms of empirical coverage probability and width of confidence intervals. We make some concluding remarks in Section 5. All the proofs are given in the appendix.

2 Consistency and Asymptotic Normality

2.1 Consistency

In order to establish the consistency of the estimator $\widehat{\phi}_{a_1a_2}(\cdot)$ given in (1), we make a few assumptions on the process $\mathbf{X}$, the kernel $K(\cdot)$ and the sequences b_n and ρ_n.

ASSUMPTION 1. The function $g_{a_1a_2}(\cdot)$, defined over the real line as $g_{a_1a_2}(t) = \sup_{|s|\geq|t|} |C_{a_1a_2}(s)|$ is integrable for all $a_1, a_2 \in \{1, 2, \ldots, r\}$.

ASSUMPTION 2. The covariance averaging kernel function $K(\cdot)$ is continuous, even, square integrable and bounded by a non-negative, even and integrable function having a unique maximum at 0. Further, $K(0) = 1$.

ASSUMPTION 3. The kernel bandwidth is such that $nb_n \to \infty$ as $n \to \infty$.

ASSUMPTION 4. The sampling rate is such that $\rho_n \to \infty$ and $\rho_n b_n \to 0$ as $n \to \infty$.

Note that Assumption 4 implies that $b_n \to 0$ as $n \to \infty$.

THEOREM 1. *Under Assumptions 1–4, the bias of the estimator $\widehat{\phi}_{a_1a_2}(\cdot)$ tends to zero uniformly over any closed and finite interval.*

In order to establish convergence of the variance-covariance matrix, we need a further assumption on the cumulants of the underlying process $\mathbf{X}$. Recall that the r-th order joint cumulant of the random variable $(Y_1, \ldots, Y_r)$ is given by

$$cum(Y_1, \ldots, Y_r) = \sum_{\boldsymbol{\nu}} (-1)^{p-1}(p-1)! \left(E \prod_{j\in\,\nu_1} Y_j \right) \times \cdots \times \left(E \prod_{j\in\,\nu_p} Y_j \right), \tag{2}$$

where the summation is over all partitions $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_p)$ of size $p = 1, \ldots, r$, of the index set $\{1, 2, \ldots, r\}$.

ASSUMPTION 5. The fourth moment $E\left[\left(X_{a_j}(t)\right)^4\right]$ of the process $\mathbf{X}$ is finite for all $a_j \in \{1, \ldots, r\}$, while the fourth order cumulant function defined by

$$cum\left[X_{a_1}(t+t_1), X_{a_2}(t+t_2), X_{a_3}(t+t_3), X_{a_4}(t)\right]$$

does not depend on t, and this function, denoted by $C_{a_1a_2a_3a_4}(t_1, t_2, t_3)$, satisfies

$$|C_{a_1a_2a_3a_4}(t_1, t_2, t_3)| \le \prod_{i=1}^{3} g_{a_i}(t_i),$$

where $g_{a_i}(t_i),\ i = 1, 2, 3$, are all continuous, even, nonnegative and integrable functions over the real line, which are non-increasing over $[0, \infty)$ for all $a_1, a_2, a_3, a_4 \in \{1, 2, \dots, r\}$.

Note that the cross spectral density is, in general, complex valued. Thus, the proposed estimator $\widehat{\phi}_{a_1a_2}(\cdot)$ can be represented as the vector

$$\begin{pmatrix} Re\left(\widehat{\phi}_{a_1a_2}(\lambda)\right) \\ Im\left(\widehat{\phi}_{a_1a_2}(\lambda)\right) \end{pmatrix}, \tag{3}$$

where

$$\begin{aligned} Re(\widehat{\phi}_{a_1a_2}(\lambda)) &= \frac{1}{2\pi n\rho_n} \sum_{t_1=1}^{n} \sum_{t_2=1}^{n} K(b_n(t_2 - t_1)) X_{a_1}\left(\frac{t_1}{\rho_n}\right) X_{a_2}\left(\frac{t_2}{\rho_n}\right) \\ &\quad \times \cos\left(\frac{(t_2 - t_1)\lambda}{\rho_n}\right) 1_{[-\pi\rho_n, \pi\rho_n]}(\lambda), \\ Im(\widehat{\phi}_{a_1a_2}(\lambda)) &= \frac{1}{2\pi n\rho_n} \sum_{t_1=1}^{n} \sum_{t_2=1}^{n} K(b_n(t_2 - t_1)) X_{a_1}\left(\frac{t_1}{\rho_n}\right) X_{a_2}\left(\frac{t_2}{\rho_n}\right) \\ &\quad \times \sin\left(\frac{(t_2 - t_1)\lambda}{\rho_n}\right) 1_{[-\pi\rho_n, \pi\rho_n]}(\lambda). \end{aligned}$$

THEOREM 2. *Under Assumptions 1–5, the covariance of* $\begin{pmatrix} Re\left(\widehat{\phi}_{a_1a_2}(\cdot)\right) \\ Im\left(\widehat{\phi}_{a_1a_2}(\cdot)\right) \end{pmatrix}$ *with* $\begin{pmatrix} Re\left(\widehat{\phi}_{a_3a_4}(\cdot)\right) \\ Im\left(\widehat{\phi}_{a_3a_4}(\cdot)\right) \end{pmatrix}$ *converges as follows:*

$$\lim_{n\to\infty} nb_n Cov\left[\begin{pmatrix} Re\left(\widehat{\phi}_{a_1a_2}(\lambda_1)\right) \\ Im\left(\widehat{\phi}_{a_1a_2}(\lambda_1)\right) \end{pmatrix}, \begin{pmatrix} Re\left(\widehat{\phi}_{a_3a_4}(\lambda_2)\right) \\ Im\left(\widehat{\phi}_{a_3a_4}(\lambda_2)\right) \end{pmatrix}\right] = \begin{bmatrix} \sigma_{11}(\lambda_1, \lambda_2) & \sigma_{12}(\lambda_1, \lambda_2) \\ \sigma_{21}(\lambda_1, \lambda_2) & \sigma_{22}(\lambda_1, \lambda_2) \end{bmatrix},$$

where

$$\begin{aligned}
\sigma_{11}(\lambda_1,\lambda_2) &= Re\big\{\phi^*_{a_1a_3}(\lambda_2)\phi_{a_2a_4}(\lambda_2)+\phi^*_{a_1a_4}(\lambda_2)\phi_{a_2a_3}(\lambda_2)\big\}\\
&\qquad\qquad \times B\cdot[1_{E_2}(\lambda_1,\lambda_2)+1_{E_3}(\lambda_1,\lambda_2)+2\times 1_{E_4}(\lambda_1,\lambda_2)],\\
\sigma_{12}(\lambda_1,\lambda_2) &= Im\big\{\phi_{a_1a_3}(\lambda_2)\phi^*_{a_2a_4}(\lambda_2)+\phi^*_{a_1a_4}(\lambda_2)\phi_{a_2a_3}(\lambda_2)\big\}\\
&\qquad\qquad \times B\cdot[1_{E_2}(\lambda_1,\lambda_2)+1_{E_3}(\lambda_1,\lambda_2)],\\
\sigma_{21}(\lambda_1,\lambda_2) &= Im\big\{\phi^*_{a_1a_3}(\lambda_2)\phi_{a_2a_4}(\lambda_2)+\phi^*_{a_1a_4}(\lambda_2)\phi_{a_2a_3}(\lambda_2)\big\}\\
&\qquad\qquad \times B\cdot[1_{E_2}(\lambda_1,\lambda_2)-1_{E_3}(\lambda_1,\lambda_2)],\\
\sigma_{22}(\lambda_1,\lambda_2) &= Re\big\{\phi_{a_1a_3}(\lambda_2)\phi^*_{a_2a_4}(\lambda_2)-\phi_{a_1a_4}(\lambda_2)\phi^*_{a_2a_3}(\lambda_2)\big\}\\
&\qquad\qquad \times B\cdot[1_{E_2}(\lambda_1,\lambda_2)-1_{E_3}(\lambda_1,\lambda_2)],\\
B &= \frac{1}{2}\int_{-\infty}^{\infty}K^2(x)dx,\\
E_1 &= \{(\lambda_1,\lambda_2):\lambda_1-\lambda_2\neq 0,\ \lambda_1+\lambda_2\neq 0,\ -\infty<\lambda_1,\lambda_2<\infty\},\\
E_2 &= \{(\lambda_1,\lambda_2):\lambda_1-\lambda_2=0,\ -\infty<\lambda_1,\lambda_2<\infty\}\setminus\{(0,0)\},\\
E_3 &= \{(\lambda_1,\lambda_2):\lambda_1+\lambda_2=0,\ -\infty<\lambda_1,\lambda_2<\infty\}\setminus\{(0,0)\},\\
E_4 &= \{(0,0)\}.
\end{aligned}$$

The convergence is uniform over any compact subset of E_1, E_2 or E_3. In particular, the variance-covariance matrix of the random vector $\begin{pmatrix} Re\left(\widehat{\phi}_{a_1a_2}(\cdot)\right) \\ Im\left(\widehat{\phi}_{a_1a_2}(\cdot)\right)\end{pmatrix}$ *goes to zero as $n\to\infty$, for all $a_1,a_2\in\{1,2,\ldots,r\}$.*

Theorems 1 and 2, which are generalizations of Theorems 1 and 2 of Srivastava and Sengupta [13] to the case of multivariate processes, together establish the consistency of any vector of estimators having elements of the form $\widehat{\phi}_{a_1a_2}(\cdot)$.

REMARK 1. The covariance between two complex-valued random variables is often defined as the trace of the 2×2 cross-covariance matrix of the random vectors formed by their real and imaginary parts [23]. In the case of the pair $(\widehat{\phi}_{a_1a_2}(\lambda_1),\widehat{\phi}_{a_3a_4}(\lambda_2))$, the limiting covariance according to this notion can be easily be computed from Theorem 2.

REMARK 2. Assumptions 3 and 4 imply $\rho_n/n\to 0$ as $n\to\infty$. This condition distinguishes the asymptotic approach used in this paper from 'infill asymptotics', where ρ_n/n is equal to a constant [20].

2.2 Asymptotic Normality

We will make an additional assumption about the underlying process in order to prove the asymptotic normality of the estimator.

ASSUMPTION 5A. The process $\mathbf{X}$ is strictly stationary; all moments of the process exist, i.e., $E\left[(X_a(t))^k\right] < \infty$ for each $k > 2$ and for all $a \in \{1, \ldots, r\}$; and for each $a_1, a_2, \ldots, a_k \in \{1, 2, \ldots, r\}$ and each $k > 2$, the kth order joint cumulant denoted by

$$C_{a_1 a_2 \ldots a_k}(t_1, t_2, \ldots, t_{k-1}) = cum\big(X_{a_1}(t_1+t), X_{a_2}(t_2+t), \ldots, X_{a_{k-1}}(t_{k-1}+t), X_{a_k}(t)\big),$$

satisfies

$$|C_{a_1 a_2 \ldots a_k}(t_1, t_2, \ldots, \ldots, t_{k-1})| \leq \prod_{i=1}^{k-1} g_{a_i}(t_i),$$

where $g_{a_i}(t_i),\ i = 1, \ldots, k-1$ are continuous, even, nonnegative and integrable functions over the real line, which are non-increasing over $(0, \infty)$.

Note that Assumption 5A is stronger than Assumption 5.

THEOREM 3. *Under Assumptions 1–4 and 5A, a vector of real and imaginary parts of estimated spectra or cross-spectra converges weakly as follows.*

$$\sqrt{nb_n}\left[\begin{pmatrix} Re\{\widehat{\phi}_{a_1a_2}(\lambda_1)\} \\ Im\{\widehat{\phi}_{a_1a_2}(\lambda_1)\} \\ \vdots \\ Re\{\widehat{\phi}_{a_{2J-1}a_{2J}}(\lambda_J)\} \\ Im\{\widehat{\phi}_{a_{2J-1}a_{2J}}(\lambda_J)\} \end{pmatrix} - E\begin{pmatrix} Re\{\widehat{\phi}_{a_1a_2}(\lambda_1)\} \\ Im\{\widehat{\phi}_{a_1a_2}(\lambda_1)\} \\ \vdots \\ Re\{\widehat{\phi}_{a_{2J-1}a_{2J}}(\lambda_J)\} \\ Im\{\widehat{\phi}_{a_{2J-1}a_{2J}}(\lambda_J)\} \end{pmatrix}\right] \xrightarrow{D} N_{2J}(0, \Sigma), \tag{4}$$

where $a_1, a_2, \ldots, a_{2J} \in \{1, 2, \ldots, r\}$, and the elements of Σ are defined in accordance with Theorem 2.

The foregoing theorem only shows that the vector estimator, after appropriate mean adjustment and scaling, converges weakly to a multivariate normal distribution. However, weak convergence around the true vector of spectra and cross-spectra remains to be established. Note that

$$\begin{aligned} \sqrt{nb_n}\left(\widehat{\phi}_{a_1a_2}(\lambda) - \phi_{a_1a_2}(\lambda)\right) = & \sqrt{nb_n}\left(\widehat{\phi}_{a_1a_2}(\lambda) - E[\widehat{\phi}_{a_1a_2}(\lambda)]\right) \\ & + \sqrt{nb_n}\left(E[\widehat{\phi}_{a_1a_2}(\lambda)] - \phi_{a_1a_2}(\lambda)\right). \end{aligned} \tag{5}$$

We make some further assumptions on the smoothness and the rate of decay of the spectrum and the shape of the kernel function in order to obtain the rate of convergence of the bias $E[\widehat{\phi}_{a_1a_2}(\lambda)] - \phi_{a_1a_2}(\lambda)$.

ASSUMPTION 1A. The function $g_{qa_1a_2}(\cdot)$, defined over the real line as

$$g_{qa_1a_2}(t) = \sup_{|s|\geq|t|} |s|^q |C_{a_1a_2}(s)|$$

is integrable for all $a_1, a_2 \in \{1, 2, \ldots, r\}$, for some positive number q greater than 1.

ASSUMPTION 1B. The power spectral density is such that, for all $a_1, a_2 \in \{1, 2, \ldots, r\}$ and for some $p > 1$, $\lim_{\lambda\to\infty} |\lambda^p \phi_{a_1a_2}(\lambda)| = A_{a_1a_2}$ for some non-negative number $A_{a_1a_2}$.

For any kernel $K(\cdot)$, let us define

$$k_s = \lim_{x\to 0} \frac{1 - K(x)}{|x|^s}$$

for each positive number s such that the limit exists. The characteristic exponent of the kernel is defined as the largest number s, such that the limit exists and is non-zero [24]. In other words, the characteristic exponent is the number s such that $1-K(1/y)$ is $O(y^{-s})$.

ASSUMPTION 2A. The characteristic exponent of the kernel $K(\cdot)$ is a number, for which Assumption 1A holds.

Note that Assumption 1A implies Assumption 1, and also that $\phi_{a_1a_2}(\cdot)$ is $[q]$ times differentiable, where $[q]$ is the integer part of q. Thus, the number q indicates the degree of smoothness of the spectral density. If Assumption 1A holds for a particular value of q, then it would also hold for smaller values.

The number p indicates the slowest rate of decay of the various elements of the power spectral density matrix. The following are two interesting situations, where Assumption 1B holds.

1. The real and imaginary parts of the components of the power spectral density matrix are rational functions of the form $\frac{P(\lambda)}{Q(\lambda)}$, where $P(\cdot)$ and $Q(\cdot)$ are polynomials such that the degree of $Q(\cdot)$ is more than degree of $P(\cdot)$ by at least p. Note that continuous time ARMA processes possess rational power spectral density.

2. The function $C_{a_1a_2}(\cdot)$ has the following smoothness property: $C_{a_1a_2}(\cdot)$ is p times differentiable and the p^{th} derivative of $C_{a_1a_2}(\cdot)$ is integrable.

THEOREM 4. *Under Assumptions 2–4, 1A, 1B and 2A, the bias of the estimator $\widehat{\phi}_{a_1a_2}(\lambda)$ given by (1), for $a_1, a_2 \in \{1, 2, \ldots, r\}$, is*

$$
\begin{aligned}
E[\widehat{\phi}_{a_1a_2}(\lambda) - \phi_{a_1a_2}(\lambda)] = & \left[-\frac{k_q}{2\pi}\int_{-\infty}^{\infty}|t|^q C_{a_1a_2}(t)e^{-it\lambda}dt\right](\rho_n b_n)^q + o\left((\rho_n b_n)^q\right) \\
& + \left[-\frac{1}{2\pi}\int_{-\infty}^{\infty}|t| C_{a_1a_2}(t)e^{-it\lambda}dt\right]\left(\frac{\rho_n}{n}\right) + o\left(\frac{\rho_n}{n}\right) \\
& + \left[\frac{A_{a_1a_2}}{(2\pi)^p}\sum_{|l|>0}\frac{1}{|l|^p}\right]\frac{1}{(\rho_n)^p} + o\left(\frac{1}{(\rho_n)^p}\right).
\end{aligned}
$$

The convergence is uniform over any closed and finite interval.

Theorem 4 shows that the second term on the right hand side of (5) would go to zero if the sampling rate ρ_n satisfies additional conditions.

ASSUMPTION 4A. The sampling rate is such that $\sqrt{nb_n}(\rho_n b_n)^q \to 0$ and $\sqrt{nb_n}/\rho_n^p \to 0$ as $n \to \infty$.

Note that, whenever Assumption 3 holds, Assumption 4A is stronger than Assumption 4. With this assumption, the expected values of the estimators in Theorem 3 can be replaced by their true values.

THEOREM 5. *Under Assumptions 1–3, 1A, 1B, 2A, 4A and 5A, we have the following weak convergence.*

$$
\sqrt{nb_n}\left[\begin{pmatrix} Re\{\widehat{\phi}_{a_1a_2}(\lambda_1)\} \\ Im\{\widehat{\phi}_{a_1a_2}(\lambda_1)\} \\ \vdots \\ Re\{\widehat{\phi}_{a_{2J-1}a_{2J}}(\lambda_J)\} \\ Im\{\widehat{\phi}_{a_{2J-1}a_{2J}}(\lambda_J)\} \end{pmatrix} - \begin{pmatrix} Re\{\phi_{a_1a_2}(\lambda_1)\} \\ Im\{\phi_{a_1a_2}(\lambda_1)\} \\ \vdots \\ Re\{\phi_{a_{2J-1}a_{2J}}(\lambda_J)\} \\ Im\{\phi_{a_{2J-1}a_{2J}}(\lambda_J)\} \end{pmatrix}\right] \xrightarrow{D} N_{2J}(0, \Sigma),
$$

where $a_1, a_2, \ldots, a_{2J} \in \{1, 2, \ldots, r\}$, and the elements of Σ are defined in accordance with Theorem 2.

3 Confidence Intervals of Spectra and Cross-spectra

3.1 Construction of Confidence Intervals

Using the result of Theorem 5, we construct an asymptotic confidence interval of the power spectral density $\phi_{a_1a_1}(\cdot)$ with confidence level $1-\alpha$, for any $a_1 \in \{1, 2, \ldots, r\}$, as

$$\left[\widehat{\phi}_{a_1a_1}(\lambda) - \frac{z_{\alpha/2}\hat{\sigma}_{a_1a_1}}{\sqrt{nb_n}}, \widehat{\phi}_{a_1a_1}(\lambda) + \frac{z_{\alpha/2}\hat{\sigma}_{a_1a_1}}{\sqrt{nb_n}}\right], \tag{6}$$

where $\hat{\sigma}^2_{a_1a_1} = [1+\delta_0(\lambda)]\left(\int_{-\infty}^{\infty} K^2(x)dx\right)\widehat{\phi}^2_{a_1a_1}(\lambda)$, which is a plug-in estimator obtained from a simplified expression of the variance given in Theorem 5, and $z_{\alpha/2}$ is the $(\alpha/2)^{\text{th}}$ quantile of the standard normal distribution. Theorem 5 ensures that the coverage probability of this interval would approach $1-\alpha$ as the sample size increases to infinity. The length of the interval is $(2z_{\alpha/2}\hat{\sigma}_{a_1a_1})/(nb_n)$. Note that consistency of the estimator (1) implies that $\hat{\sigma}_{a_1a_1}$ converges in probability to $\sigma_{a_1a_1}$. Therefore, it follows from Assumption 3 that the interval length shrinks to zero as n goes to infinity.

Similarly, by expressing the real and imaginary parts of the cross spectral density $\phi_{a_1a_2}(\lambda)$ as a vector, a confidence region with level $1-\alpha$, for any $a_1, a_2 \in \{1, 2, \ldots, r\}$, is constructed as

$$\left\{Z : nb_n \begin{pmatrix} Re\{\widehat{\phi}_{a_1a_2}(\lambda) - Z\} \\ Im\{\widehat{\phi}_{a_1a_2}(\lambda) - Z\} \end{pmatrix}' \widehat{\Sigma}^{-1} \begin{pmatrix} Re\{\widehat{\phi}_{a_1a_2}(\lambda) - Z\} \\ Im\{\widehat{\phi}_{a_1a_2}(\lambda) - Z\} \end{pmatrix} \le \chi^2_{\alpha,2}\right\},$$

where the constant $\chi^2_{\alpha,2}$ is the $(1-\alpha)^{\text{th}}$ quantile of the chi-square distribution with two degrees of freedom and

$$\begin{aligned}
\widehat{\Sigma} &= \begin{bmatrix} \widehat{\sigma}^2_1(\lambda) & \widehat{\sigma}_{12}(\lambda) \\ \widehat{\sigma}_{12}(\lambda) & \widehat{\sigma}^2_2(\lambda) \end{bmatrix}, \\
\widehat{\sigma}^2_1(\lambda) &= \{1 + 1_{\{0\}}(\lambda)\}B[\widehat{\phi}_{a_1a_1}(\lambda)\widehat{\phi}_{a_2a_2}(\lambda) + \{Re(\widehat{\phi}_{a_1a_2}(\lambda))\}^2 - \{Im(\widehat{\phi}_{a_1a_2}(\lambda))\}^2], \quad (7) \\
\widehat{\sigma}^2_2(\lambda) &= \{1 - 1_{\{0\}}(\lambda)\}B[\widehat{\phi}_{a_1a_1}(\lambda)\widehat{\phi}_{a_2a_2}(\lambda) - \{Re(\widehat{\phi}_{a_1a_2}(\lambda))\}^2 + \{Im(\widehat{\phi}_{a_1a_2}(\lambda))\}^2], \quad (8) \\
\widehat{\sigma}_{12}(\lambda) &= \{1_{\{0\}}(\lambda) - 1\}B\{2Re(\widehat{\phi}_{a_1a_2}(\lambda))Im(\widehat{\phi}_{a_1a_2}(\lambda))\}, \\
B &= \frac{1}{2}\int_{-\infty}^{\infty} K^2(x)dx,
\end{aligned}$$

$\widehat{\Sigma}$ being a plug-in estimator of Σ defined in Theorem 5 for $J = 1$.

One can also construct individual confidence limits for the real and imaginary parts of the cross spectral density $\phi_{a_1a_2}(\lambda)$, for any $a_1, a_2 \in \{1, 2, \ldots, r\}$, as

$$\left[Re[\widehat{\phi}_{a_1a_2}(\lambda)] - \frac{z_{\alpha/2}\widehat{\sigma}_1(\lambda)}{\sqrt{nb_n}}, Re[\widehat{\phi}_{a_1a_2}(\lambda)] + \frac{z_{\alpha/2}\widehat{\sigma}_1(\lambda)}{\sqrt{nb_n}}\right], \tag{9}$$

and

$$\left[Im[\widehat{\phi}_{a_1a_2}(\lambda)] - \frac{z_{\alpha/2}\widehat{\sigma}_2(\lambda)}{\sqrt{nb_n}}, Im[\widehat{\phi}_{a_1a_2}(\lambda)] + \frac{z_{\alpha/2}\widehat{\sigma}_2(\lambda)}{\sqrt{nb_n}}\right], \tag{10}$$

where $z_{\alpha/2}$ is the $\alpha/2$th quantile of the normal distribution. Simultaneous confidence intervals for the real and the imaginary parts can be obtained using standard techniques [4].

3.2 Optimal Rate of Shrinkage

It can be seen from the expressions of confidence intervals and confidence regions based on Theorem 5 that the size of these intervals/regions go to zero as $\frac{1}{\sqrt{nb_n}}$ goes to zero. We now seek to optimize the rates of b_n and ρ_n so that $\frac{1}{\sqrt{nb_n}}$ tends to 0 as fast as possible under the conditions of Theorem 5.

THEOREM 6. *Under Assumptions 3 and 4A, the reciprocal of the scale factor ($\sqrt{nb_n}$) used in Theorem 5 has the fastest convergence to 0 when*

$$\begin{aligned} b_n &= o\left(n^{-\frac{p+q}{p+q+2pq}}\right), \\ \rho_n &= O\left(n^{\frac{q}{p+q+2pq}}\right), \end{aligned}$$

and under these conditions, $\frac{1}{\sqrt{nb_n}} = o\left(n^{-\frac{pq}{p+q+2pq}}\right)$.

It has been shown in [13] that under the assumptions of Theorems 2 and 4, the optimal rate of convergence for mean square consistency of the estimator (1) is given as

$$E\left[\{\widehat{\phi}_{a_1a_2}(\cdot) - \phi_{a_1a_2}(\cdot)\}^2\right] = O\left(n^{-\frac{2pq}{p+q+2pq}}\right),$$

which corresponds to the choices

$$\begin{aligned} b_n &= O\left(n^{-\frac{p+q}{p+q+2pq}}\right), \\ \rho_n &= O\left(n^{\frac{q}{p+q+2pq}}\right). \end{aligned}$$

Theorem 6 shows that the optimal rate of weak convergence of the estimator $\widehat{\phi}_{a_1a_2}(\cdot)$ is slower than the square root of the optimal rate corresponding to mean square consistency.

It is important to note that for every fixed value of q, the number p, which indicates rate of decay of the spectrum, can be increased indefinitely by continuous time low pass filtering with a cut off frequency larger than the maximum frequency of interest. There are well-known filters such as the Butterworth filter, which have polynomial rate of decay of the transfer function with specified degree of the polynomial, that can be used for this purpose. For fixed q, the best rate of weak convergence given in Theorem 6, obtained by allowing p to go to infinity, happens to be $o\left(n^{-\frac{q}{1+2q}}\right)$. This rate coincides with the usual rate of weak convergence of the smoothed periodogram estimator for a bandlimited spectral density, when the sampling rate is fixed and appropriate [24].

The rate of weak convergence crucially depends on the number q, the assumed degree of smoothness of the spectrum. The stronger the assumption, the faster is the rate of convergence. The rate corresponding to $q = 1$ (weakest possible assumption) is $o\left(n^{-\frac{1}{3}}\right)$, assuming that p can be allowed to be large. For large q (strong assumption) and correspondingly large p, the rate approaches $o\left(n^{-\frac{1}{2}}\right)$.

Now, we compare the rate of optimal shrinkage of the confidence intervals of a non-bandlimited spectrum based on $\widehat{\phi}_{a_1a_2}(\cdot)$ with that of confidence intervals based on Masry's estimator for Poisson sampled data, in the special case of a univariate process. For simplicity, we consider the process $\{X_{a_1}(t), -\infty < t < \infty\}$, for some $a_1 \in \{1, \ldots, r\}$. Masry's estimator based on samples $X_{a_1}(t_1)$, $X_{a_1}(t_2)$, …, $X_{a_1}(t_n)$, $t_1, t_2, \ldots, t_n$ being the jump points a Poisson process with mean intensity β, is defined as

$$\widehat{\psi}(\lambda) \quad = \quad \frac{1}{\pi\beta n}\sum_{j=1}^{n-1}\sum_{l=1}^{n-j} X_{a_1}(t_l)X_{a_1}(t_{j+l})K\left(b_n(t_j+l-t_l)\right)\cos(\lambda(t_j+l-t_l)), \qquad (11)$$

where $K(\cdot)$ is a covariance averaging kernel and b_n is the kernel bandwidth. Masry [8] showed that, under conditions similar to those used in the case of $\widehat{\phi}_{a_1a_1}(\cdot)$, $\widehat{\psi}(\lambda)$ is consistent for $\phi_{a_1a_1}(\cdot)$. Lii and Masry [9] gave conditions to ensure

$$\sqrt{nb_n}\left[\begin{pmatrix}\widehat{\psi}_{a_1a_1}(\lambda_1)\\ \widehat{\psi}_{a_1a_1}(\lambda_2)\\ \vdots \\ \widehat{\psi}_{a_1a_1}(\lambda_J)\end{pmatrix} - \begin{pmatrix}\phi_{a_1a_1}(\lambda_1)\\ \phi_{a_1a_1}(\lambda_2)\\ \vdots \\ \phi_{a_1a_1}(\lambda_J)\end{pmatrix}\right] \xrightarrow{D} N_J(0,\Theta),$$

where the elements of Θ are defined as

$$\theta_{a_1a_1}(k,l) = 1_0(\lambda_k-\lambda_l)\beta\left[\phi_{a_1a_1}(\lambda_k)+\frac{C_{a_1a_1}(0)}{2\pi\beta}\right]^2(1+1_0(\lambda_k))\int_{-\infty}^{\infty}K^2(x)dx.$$

By using this result, one can have the confidence interval of $\phi_{a_1a_1}(\cdot)$ with level $1-\alpha$ as

$$\left[\widehat{\psi}_{a_1a_1}(\lambda)-\frac{z_{\alpha/2}\sqrt{\hat{\theta}_{a_1a_1}}}{\sqrt{nb_n}}, \widehat{\psi}_{a_1a_1}(\lambda)+\frac{z_{\alpha/2}\sqrt{\hat{\theta}_{a_1a_1}}}{\sqrt{nb_n}}\right], \qquad (12)$$

where $\hat{\theta}_{a_1a_1}$ is a plug-in estimator of the variance, obtained by replacing $C_{a_1a_1}(0)$ in the expression of $\theta_{a_1a_1}(1,1)$ by a consistent estimator. The optimal rate of shrinkage of the width of this confidence intervals is $o\left(n^{-\frac{q}{1+2q}}\right)$. This rate is faster than $o\left(n^{-\frac{q}{1+2q+q/p}}\right)$, the optimal rate of shrinkage of width of confidence intervals based on $\widehat{\phi}_{a_1a_1}(\cdot)$. However, the later rate approaches the former for large values of p.

4 Simulation

We consider the zero mean, continuous time process $X(t) = Y_1(t) + Y_2(t)$, where $Y_j(u)$, $j = 1, 2$ are independent continuous time stationary processes with spectral density

$$\begin{aligned}
\phi_{Y_1}(\lambda) &= \frac{0.05}{0.05 + \lambda^2} \\
\phi_{Y_2}(\lambda) &= 50[h(\lambda, 3.6\pi, 0.12) + h(\lambda, -3.6\pi, 0.12)] + 10[h(\lambda, 3.2\pi, 0.15) + h(\lambda, -3.2\pi, 0.15)],
\end{aligned}$$

where $h(\lambda, \alpha, \beta) = \frac{1}{\beta} e^{-\frac{1}{2}\left(\frac{\lambda-\alpha}{\beta}\right)^2}$. It follows that the spectrum of the process X has peaks at 0, $\pm 3.2\pi$ and $\pm 3.6\pi$.

Suppose one is interested in estimating the spectrum only over the frequency range $[0, \pi]$. If one wishes to use uniformly spaced samples for this purpose, then one would filter the process in continuous time at a cut-off frequency higher than π, before sampling it. We use the 8th degree Butterworth filter with cut-off frequency 1.5π. The spectral density of the filtered process is

$$\begin{aligned}
\phi(\lambda) = &\left\{ \frac{0.05}{0.05 + \lambda^2} + 50h(\lambda, 3.6, 0.12) + 50h(\lambda, -3.6, 0.12) + 10h(\lambda, 3.2, 0.15) \right. \\
&\left. + h(\lambda, -3.2, 0.15) \right\} \frac{1}{1 + \left(\frac{|\lambda|}{1.5\pi}\right)^8}
\end{aligned}$$

According to conventional wisdom, the filtered process is regarded as one with bandwidth 1.5π, and hence the sampling rate should be greater than 1.5. We choose $\rho = 2$. For the sample size $n = 100$, we compute the confidence intervals for the spectrum of X_1, as given in (6). Even though the true spectrum is infinitely differentiable, we only assume second order differentiability ($q = 2$) for the purpose of estimation. Accordingly, we choose the second order kernel

$$K(x) = \frac{1}{2}\{1 + \cos(\pi x)\} 1_{[-1,1]}(x),$$

and $b_n = 1/30$. This value of b_n is determined by the cross-validation technique [25]. Figure 1 shows the plot of the average lower and upper 95% confidence limits based on 500 simulation runs,

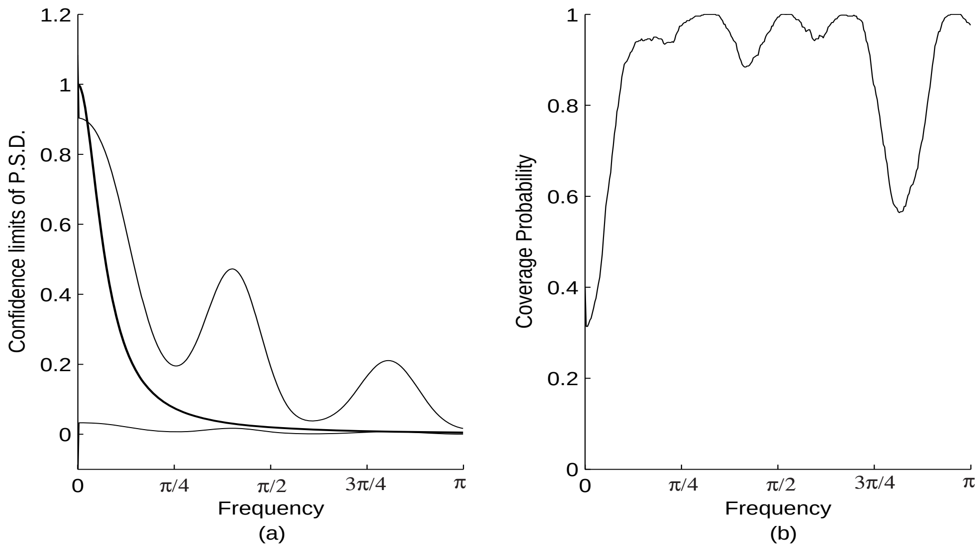


Figure 1: (a) Average lower and upper confidence limits (thin lines) together with true spectral density (thick line) and (b) empirical coverage probability of confidence intervals, for $\widehat{\phi}(\lambda)$ with $n = 100$, based on 500 simulation runs.

along with the true $\phi(\lambda)$, for $\lambda \in [0, \pi]$. Figure 1 also shows the empirical coverage probability based on 500 simulation runs. The wavy nature of the locus of the confidence intervals can be attributed to aliased versions of the peaks at frequencies -3.2π and -3.6π, which are not completely suppressed by filtering in continuous time. The empirical coverage probability is poor at zero frequency, which is due to the effect of smoothing at a location of sharp peak of the true spectrum. The coverage probability also dips at the wrapped-around versions of the peaks at -3.2π and -3.6π (more prominently at the second one).

Now suppose one has resources to increase the sample size n to 1000. Aliasing will continue to be present if the sampling rate is unchanged, and we wish to examine the effect of sampling faster as per the prescription of Section 3.2. In order to distinguish between the estimators based on fixed sampling rate and progressively faster sampling rate, we denote these two estimators by $\widehat{\xi}(\cdot)$ and $\widehat{\phi}(\cdot)$, respectively. In accordance with the chosen order of the Butterworth filter used in continuous time, we assume $p = 8$. This choice, together with the assumption $q = 2$, dictates that, as the sample size changes from 100 to 1000, the sampling rate (ρ_n) should be increased by the factor $(1000/100)^{1/21}$ and the kernel window width (b_n) should be changed by the factor $(1000/100)^{-5/21-\delta}$ for some small positive number δ. We choose $\delta = 1/6$. The optimal choice

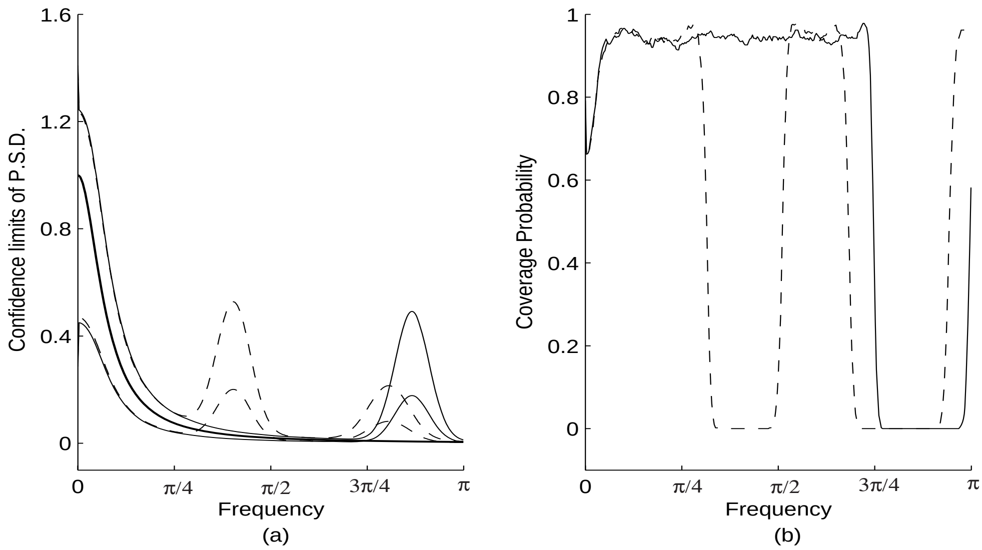


Figure 2: (a) Average lower and upper confidence limits together with true spectral density (thick line) and (b) empirical coverage probability of confidence intervals, for $\widehat{\phi}(\lambda)$ (thin solid lines) and $\widehat{\xi}(\lambda)$ (dotted lines) with $n = 1000$, based on 500 simulation runs.

of b_n for $\widehat{\xi}(\cdot)$ is $o\left(n^{-1/(1+2q)}\right)$ [4]. Therefore, in the case of $\widehat{\xi}(\cdot)$, as we switch from $n = 100$ to $n = 1000$, we change b_n by the factor $(1000/100)^{-1/5-\delta}$, for $\delta = 1/6$.

Similar further changes to ρ_n and b_n (for $\widehat{\phi}$ as well as $\widehat{\xi}$) are made when we run simulations for $n = 10,000$.

Figures 2 and 3 show the plots of the average lower and upper 95% confidence limits based on $\widehat{\phi}(\cdot)$ and $\widehat{\xi}(\cdot)$ for sample sizes 1000 and 10,000, respectively, computed from the 500 simulation runs, along with the true $\phi(\cdot)$. The figures show that effects of the spurious peaks on the confidence limits based on $\widehat{\phi}(\cdot)$ diminish for $n = 1000$ and disappear for $n = 10,000$. However, confidence limits based on $\widehat{\xi}(\cdot)$ continue to be affected by the two spurious peaks. The plot of empirical coverage probability also confirms the improvement with larger n in the case of $\widehat{\phi}(\cdot)$ and non-improvement in the case of $\widehat{\xi}$. In fact, for the fixed rate approach, which ignores the non-ideal nature of the anti-alias filtering, the coverage probabilities in the range of spurious peaks become zero for large sample size. This hazard is avoided by using progressively faster sampling.

We now compare the confidence limits given in (12) based on the Poisson sampled data with the confidence limits based on $\widehat{\phi}(\cdot)$ and $\widehat{\xi}(\cdot)$. We generate samples from the filtered process using the same specification as before, with average sampling rates $\beta = 2$ and $\beta = 0.25$. We choose

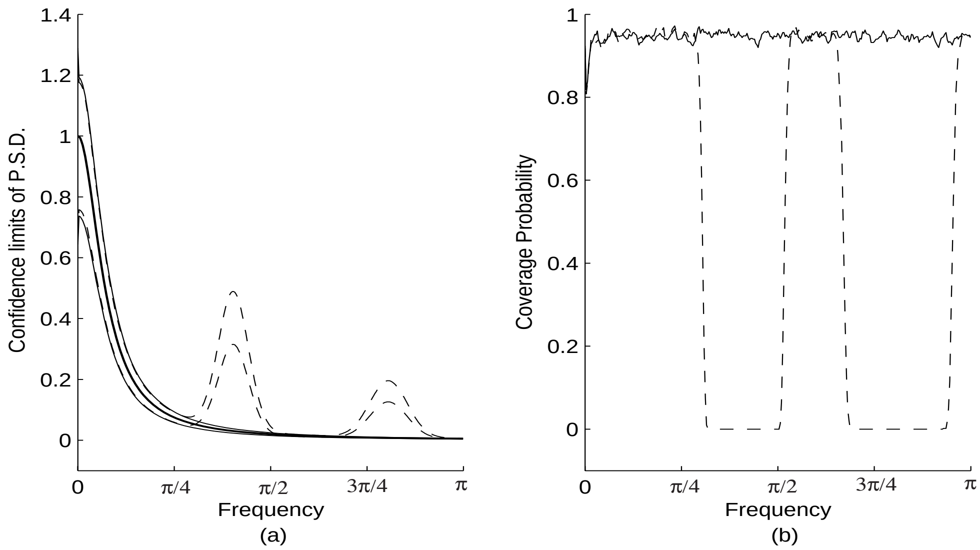
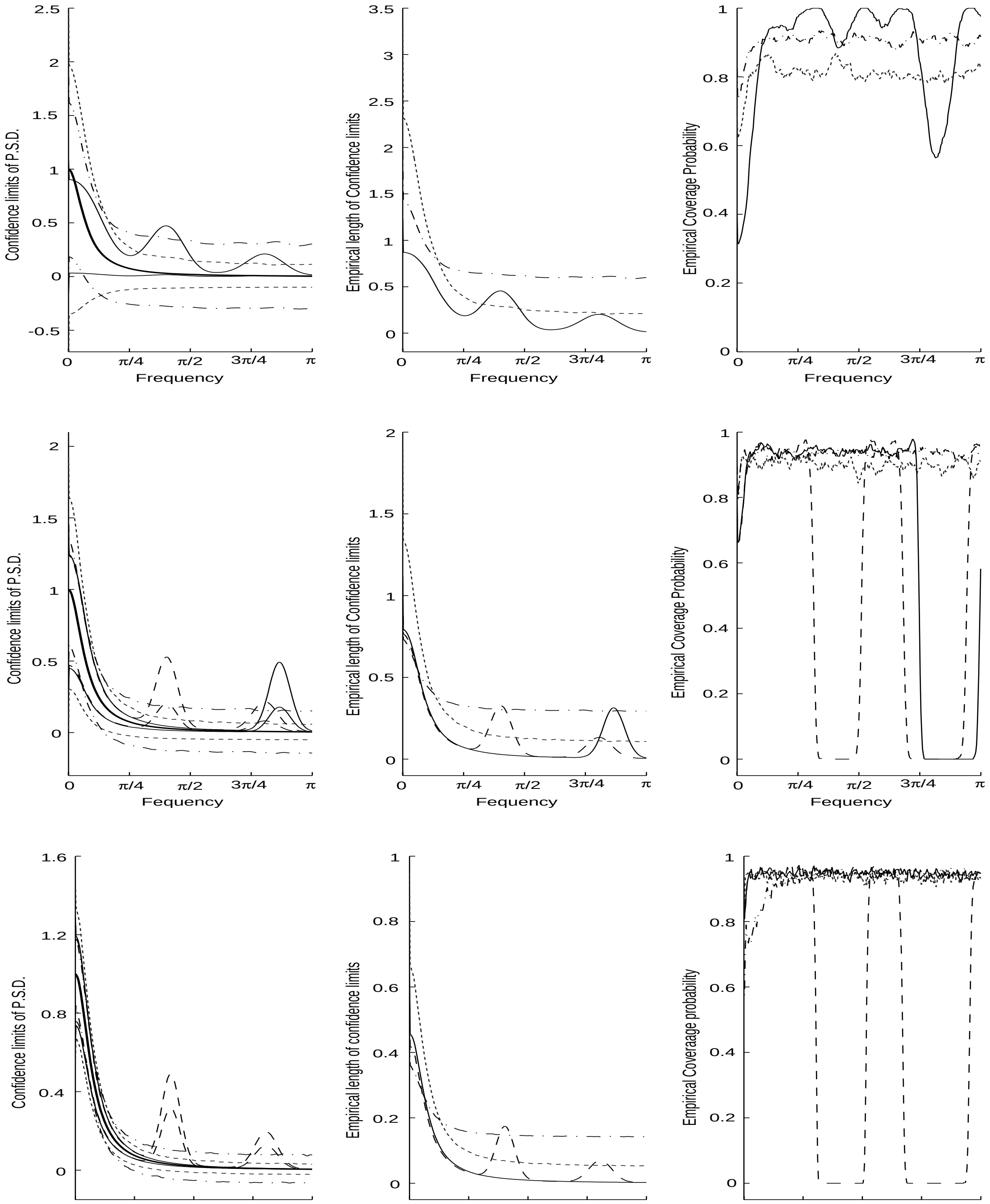


Figure 3: (a) Average lower and upper confidence limits together with true spectral density (thick line) and (b) empirical coverage probability of confidence intervals, for $\widehat{\phi}(\lambda)$ (thin solid lines) and $\widehat{\xi}(\lambda)$ (dotted lines) with $n = 10,000$, based on 500 simulation runs.

$b_n = 1/30$ (as in the case of confidence intervals based on uniform sampling) for sample size $n = 100$. For larger sample sizes, we change b_n by the factor $(n/100)^{-1/5-\delta}$ for $\delta = 1/6$, in accordance with the finding [9] that the optimal choice of b_n for $\widehat{\psi}(\cdot)$ is $o\left(n^{-1/(1+2q)}\right)$.

Figure 4 shows the plot of average lower and upper 95% confidence limits based on 500 simulation runs based on $\widehat{\phi}(\cdot)$, $\widehat{\xi}(\cdot)$ and $\widehat{\psi}(\cdot)$ along with the true spectral density. The average lengths of the confidence intervals and the empirical coverage probabilities are also plotted. The plots indicate that the confidence intervals based on Poisson sampled data are generally wider than those based on $\widehat{\phi}(\cdot)$ and progressively faster uniform sampling. The latter intervals achieve the nominal coverage probability across frequencies at smaller sample size in comparison with the cases of the other two sets of confidence intervals. The average lower confidence limits based on Poisson sampled data are negative over much of the frequency range of interest even for the sample size $n = 10,000$, while the tighter limits based on $\widehat{\phi}(\cdot)$ do not have this problem. The generally wider confidence limits based on Poisson sampled data with average sampling rate $\beta = 0.25$ than with $\beta = 2$, show how these confidence limits depend on the average sampling rate for finite sample size, even though the estimator is known to be consistent for *any* choice of β [9].



P.S.D.	$\widehat{\phi}(\cdot)$	$\widehat{\xi}(\cdot)$	$\widehat{\psi}(\cdot)$, $\beta = 2$	$\widehat{\psi}(\cdot)$, $\beta = 0.25$
━━━━	────	— — —	- - - - - - - -	· — · — ·

Figure 4: Plots of average upper and lower confidence limits (left), average length of confidence intervals (center) and empirical coverage probability (right) based on 500 simulation runs for different spectrum estimators, for sample sizes 100 (top row), 1000 (middle row) and 10,000 (bottom row).

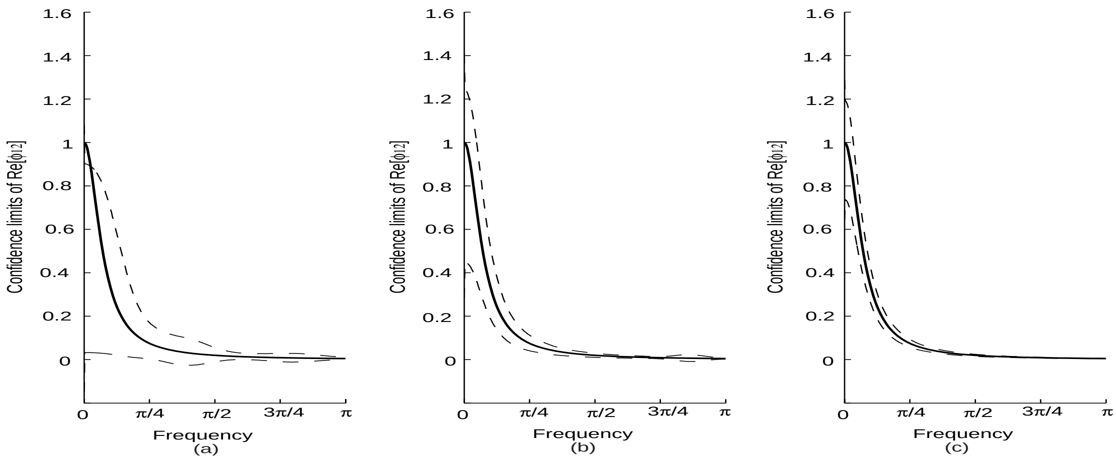


Figure 5: Average lower and upper confidence limits (dotted lines) for $Re[\widehat{\phi}_{12}](\lambda)$, based on 500 simulation runs, along with true $Re[\phi_{12}](\lambda)$ (solid line), for sample size (a) $n = 100$, (b) $n = 1000$ and (c) $n = 10,000$.

Now we consider the bivariate process,

$$\begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} = \begin{pmatrix} Y_1 + Y_2 \\ Y_1 \end{pmatrix},$$

where Y_1 and Y_2 are described at the beginning of this section. We again use the 8th degree Butterworth filter with cut-off frequency 1.5π. The real and imaginary parts of the cross spectral density of the filtered process are

$$\begin{aligned} Re[\phi_{12}(\lambda)] &= \left\{\frac{0.05}{0.05+\lambda^2}\right\}\frac{1}{1+\left(\frac{|\lambda|}{1.5\pi}\right)^8}, \\ Im[\phi_{12}(\lambda)] &= 0. \end{aligned}$$

Figure 5 and 6 show the plots of average confidence limits for $Re[\phi_{11}](\cdot)$ based on progressively faster uniform sampling as given in (9) and the empirical coverage probability, respectively, based on 500 simulation runs and choices of the parameters as in case of $\phi(\cdot)$. The plots show that the confidence limits get sharper and the empirical coverage probability across frequencies approaches the nominal value, as the sample size increases from 100 to 10,000. The corresponding plots for $Im[\phi_{11}](\cdot)$, given in Figures 7 and 8, depict a similar story. The spurious peaks in the plot of confidence limits (Figure 7) show the effect of aliasing in the spectral density of the first component of the process, which diminishes at $n = 1000$ and disappears at $n = 10,000$.

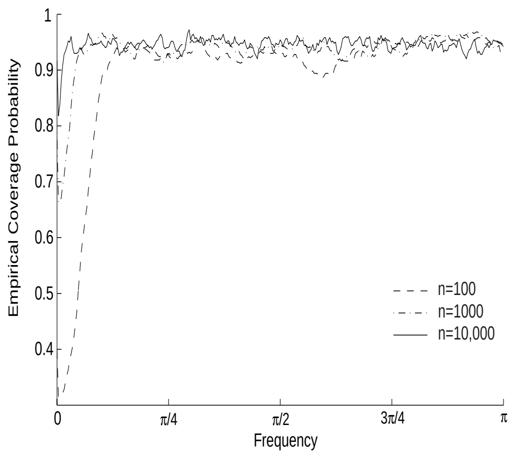


Figure 6: Empirical coverage probability of the confidence intervals of $Re[\widehat{\phi_{12}}](\lambda)$, based on 500 simulation runs.

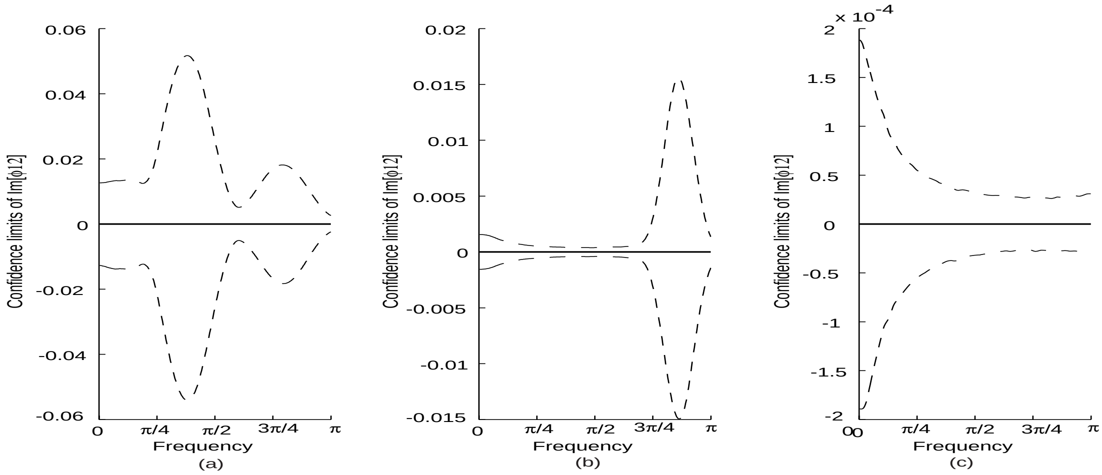


Figure 7: Average lower and upper confidence limits (dotted lines) for $Im[\widehat{\phi_{12}}](\lambda)$, based on 500 simulation runs, along with true $Im[\phi_{12}](\lambda)$ (solid line), for sample size (a) $n = 100$, (b) $n = 1000$ and (c) $n = 10,000$.

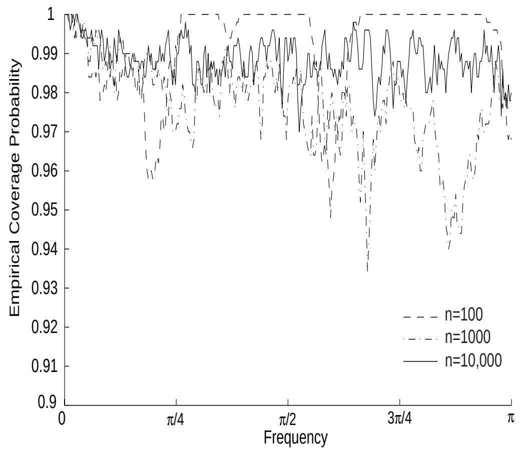


Figure 8: Empirical coverage probability of the confidence intervals of $Im[\widehat{\phi}_{12}](\lambda)$, based on 500 simulation runs.

5 Concluding Remarks

The confidence intervals of power spectral densities presented in this paper are based on a well-known spectrum estimator, namely, the smoothed periodogram. It is argued that when greater sample size is available, one can make judicious use of the same by sampling suitably faster, and constructing confidence intervals accordingly. The proposed intervals are the first ones based on uniform sampling, that do not require an assumption of finite bandwidth of the underlying process. No parametric model has been assumed for the power spectral density or the underlying probability distribution of the samples. The method presented here is computationally simple (the order of computation being the same as that for the computation of the spectrum estimate), as no resampling is needed. Another advantage is that it produces confidence intervals with width shrinking to zero as the sample size goes to infinity.

The proposed method of spectrum estimation utilizes the decay rate of the spectrum using progressively faster uniform sampling. The simulation results of Section 4 indicate that these confidence intervals have adequate coverage probability at moderate sample size. These also have a better coverage probability with smaller width than those based on Poisson sampling, when the sample size is large. The results derived here can pave the way for the construction of asymptotic confidence *bands* that shrink to the true power spectral density, which may possibly be non-bandlimited.

Appendix : Proofs

We denote by $K_1(\cdot)$ a function that bounds the covariance averaging kernel $K(\cdot)$ as in Assumption 2. Further, we denote $K_1(0)$ by M.

PROOF OF THEOREM 1. We shall show that the bias of the estimator $\widehat{\phi}_{a_1a_2}(\lambda)$ given by (1) converges to 0 uniformly over $[\lambda_l, \lambda_u]$ for any λ_l, λ_u such that $\lambda_l < \lambda_u$. Note that

$$E[\widehat{\phi}_{a_1a_2}(\lambda)] = \frac{1}{2\pi\rho_n} \sum_{u=-(n-1)}^{n-1} \left(1 - \frac{|u|}{\rho_n}\right) K(b_n u) C_{a_1a_2}\left(\frac{u}{\rho_n}\right) e^{-\frac{iu\lambda}{\rho_n}} 1_{[-\pi\rho_n, \pi\rho_n]}(\lambda).$$

Consider the simple function $S_n(\cdot)$, defined over $[\lambda_l, \lambda_u] \times \mathbb{R}$, by

$$S_n(\lambda, x) = \frac{1}{2\pi} \sum_{u=-(n-1)}^{n-1} \left(1 - \frac{|u|}{\rho_n}\right) K(b_n u) C_{a_1a_2}\left(\frac{u}{\rho_n}\right) e^{-\frac{iu\lambda}{\rho_n}} 1_{[-\pi\rho_n, \pi\rho_n]}(\lambda) 1_{\left(\frac{u-1}{\rho_n}, \frac{u}{\rho_n}\right]}(x).$$

Observe that $\int_{-\infty}^{\infty} S_n(\lambda, x)dx = E[\widehat{\phi}_{a_1a_2}(\lambda)]$. Define the function $S(\cdot)$, over $[\lambda_l, \lambda_u] \times \mathbb{R}$, by $S(\lambda, x) = \frac{1}{2\pi} C_{a_1a_2}(x) e^{-ix\lambda}$.

For any $x \in \mathbb{R}$, let $u_n(x)$ be the smallest integer greater than or equal to $\rho_n x$. Note that the interval $\left(\frac{u_{n-1}(x)}{\rho_n}, \frac{u_n(x)}{\rho_n}\right]$ contains the point x and $\lim_{n\to\infty} \frac{u_n(x)}{\rho_n} = x$. For sufficiently large n, we have from Assumptions 3 and 4,

$$S_n(\lambda, x) = \frac{1}{2\pi} \left(1 - \frac{|u_n(x)|}{\rho_n} \frac{\rho_n}{n}\right) K\left(b_n \rho_n \frac{u_n(x)}{\rho_n}\right) C_{a_1a_2}\left(\frac{u_n(x)}{\rho_n}\right) e^{-\frac{iu_n(x)\lambda}{\rho_n}} 1_{[-\pi\rho_n, \pi\rho_n]}(\lambda).$$

Proving the uniform convergence of $Bias[\widehat{\phi}_{a_1a_2}(\lambda)]$ over the finite interval $[\lambda_l,\ \lambda_u]$ amounts to proving

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} S_n(\lambda, x)dx = \int_{-\infty}^{\infty} S(\lambda, x)dx, \tag{A.1}$$

uniformly over $[\lambda_l,\ \lambda_u]$.

Observe that $\int_{-\infty}^{\infty} S(\lambda, t)dt = \phi_{a_1a_2}(\lambda)$, which is continuous. By virtue of the continuity of the limiting function, (A.1) is equivalent to proving that $\int_{-\infty}^{\infty} S_n(\lambda, x)dx$ converges continuously over this interval [26], i.e., for any sequence $\lambda_n \to \lambda$,

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} S_n(\lambda_n, x)dx = \int_{-\infty}^{\infty} S(\lambda, x)dx, \tag{A.2}$$

where $\lambda_n, \lambda \in [\lambda_l,\ \lambda_u]$.

By continuity of the function $S_n(\lambda, x)$ with respect to x and λ, we have from Assumptions 3 and 4, for any fixed x,

$$\lim_{n\to\infty} |S_n(\lambda_n, x) - S(\lambda, x)| = 0.$$

Note that from Assumptions 1 and 2, we have the dominance

$$|S_n(\lambda_n, x)| \leq M \sum_{|u|<n} \left| C_{a_1a_2}\left(\frac{u}{\rho_n}\right) \right| 1_{\left(\frac{u-1}{\rho n}, \frac{u}{\rho n}\right]}(x) \leq M g_{a_1a_2}(x),$$

where $g_{a_1a_2}(\cdot)$ is the function described in Assumption 1. Thus, by applying the dominated convergence theorem (DCT), we have (A.2).

Hence, $E[\widehat{\phi}_{a_1a_2}(\lambda)] \to \phi(\lambda)$ uniformly on $[\lambda_l,\ \lambda_u]$. □

PROOF OF THEOREM 2. We begin by calculating the covariance between the estimators $Re(\widehat{\phi}_{a_1a_2}(\cdot))$ and $Re(\widehat{\phi}_{a_3a_4}(\cdot))$.

$$\begin{aligned}
&Cov\left[Re(\widehat{\phi}_{a_1a_2}(\lambda_1)), Re(\widehat{\phi}_{a_3a_4}(\lambda_2))\right] \\
=&\frac{1}{(2\pi)^2(n\rho_n)^2}\sum_{t_1=1}^{n}\sum_{t_2=1}^{n}\sum_{t_3=1}^{n}\sum_{t_4=1}^{n} K(b_n(t_2-t_1))K(b_n(t_4-t_3)) \\
&\times Cov\left[X_{a_1}\left(\frac{t_1}{\rho_n}\right)X_{a_2}\left(\frac{t_2}{\rho_n}\right), X_{a_3}\left(\frac{t_3}{\rho_n}\right)X_{a_4}\left(\frac{t_4}{\rho_n}\right)\right]\cos\left(\frac{(t_2-t_1)\lambda_1}{\rho_n}\right)\cos\left(\frac{(t_4-t_3)\lambda_2}{\rho_n}\right) \\
=&\frac{1}{(2\pi)^2(n\rho_n)^2}\sum_{t_1=1}^{n}\sum_{t_2=1}^{n}\sum_{t_3=1}^{n}\sum_{t_4=1}^{n} K(b_n(t_2-t_1))K(b_n(t_4-t_3))\left[C_{a_1a_3}\left(\frac{t_1-t_3}{\rho_n}\right)C_{a_2a_4}\left(\frac{t_2-t_4}{\rho_n}\right)\right. \\
&\left.+C_{a_1a_4}\left(\frac{t_1-t_4}{\rho_n}\right)C_{a_2a_3}\left(\frac{t_2-t_3}{\rho_n}\right)+C_{a_1a_2a_3a_4}\left(\frac{t_1-t_4}{\rho_n}, \frac{t_2-t_4}{\rho_n}, \frac{t_3-t_4}{\rho_n}\right)\right] \\
&\times\cos\left(\frac{(t_1-t_2)\lambda_1}{\rho_n}\right)\cos\left(\frac{(t_3-t_4)\lambda_2}{\rho_n}\right) \\
=&T_1(\lambda_1,\lambda_2)+T_2(\lambda_1,\lambda_2)+T_3(\lambda_1,\lambda_2),
\end{aligned}$$

where the three terms correspond to the three summands appearing inside square brackets in the previous step.

Now consider the function $T_1(\lambda_1, \lambda_2)$. By using the transformations $u_1 = t_1 - t_2$, $u_2 = t_1 - t_3$ and $u_3 = t_2 - t_4$, we have

$$\begin{aligned}
&T_1(\lambda_1,\lambda_2) \\
&=\frac{1}{(2\pi)^2(n\rho_n)^2}\sum_{t_1=1}^{n}\sum_{u_1=t_1-1}^{t_1-n}\sum_{u_2=t_1-1}^{n-t_1}\sum_{u_3=t_1-1-u_1}^{t_1-n-u_1} K(b_nu_1)K(b_n(u_1-u_2+u_3))C_{a_1a_3}\left(\frac{u_2}{\rho_n}\right) \\
&\quad\times C_{a_2a_4}\left(\frac{u_3}{\rho_n}\right)\cos\left(\frac{u_1\lambda_1}{\rho_n}\right)\cos\left(\frac{(u_1-u_2+u_3)\lambda_2}{\rho_n}\right).
\end{aligned}$$

The range of the four summations on the right hand side is described by the set of inequalities $1 \leq t_1 \leq n$ and $t_1 - n \leq u_1, u_2, u_1 + u_3 \leq t_1 - 1$, which is equivalent to the inequalities $-(n-1) \leq u_1, u_2, u_1+u_3 \leq (n-1)$ and $\max\{u_1, u_2, u_1+u_3\}+1 \leq t_1 \leq \min\{u_1, u_2, u_1+u_3\}$.

Therefore, the expression for $T_1(\lambda_1, \lambda_2)$ simplifies to

$$\frac{1}{(2\pi)^2 n\rho_n^2} \sum_{u_1=-(n-1)}^{(n-1)} \sum_{u_2=-(n-1)}^{(n-1)} \sum_{u_3=-(n-1)-u_1}^{(n-1)-u_1} U_n(u_1, u_2, u_3) K(b_n u_1) K(b_n(u_1 - u_2 + u_3))$$
$$\times C_{a_1a_3}\left(\frac{u_2}{\rho_n}\right) C_{a_2a_4}\left(\frac{u_3}{\rho_n}\right) \cos\left(\frac{u_1\lambda_1}{\rho_n}\right) \cos\left(\frac{(u_1 - u_2 + u_3)\lambda_2}{\rho_n}\right),$$

where

$$U_n(u_1, u_2, u_3) = \left(1 + \frac{\min(u_1, u_2, u_1+u_3)}{n} - \frac{\max(u_1, u_2, u_1+u_3)}{n}\right).$$

By writing the cosine functions in terms of complex exponentials, we have

$$\begin{aligned}
& T_1(\lambda_1, \lambda_2) \\
=& \frac{1}{(4\pi)^2 n\rho_n^2} \sum_{u_1=-(n-1)}^{(n-1)} \sum_{u_2=-(n-1)}^{(n-1)} \sum_{u_3=-(n-1)-u_1}^{(n-1)-u_1} U_n(u_1, u_2, u_3) K(b_n u_1) K(b_n(u_1 - u_2 + u_3)) \\
& \times C_{a_1a_3}\left(\frac{u_2}{\rho_n}\right) C_{a_2a_4}\left(\frac{u_3}{\rho_n}\right) \left\{ e^{-i\frac{(\lambda_1-\lambda_2)u_1}{\rho_n}} e^{-i\frac{\lambda_2 u_2}{\rho_n}} e^{i\frac{\lambda_2 u_3}{\rho_n}} + e^{i\frac{(\lambda_1-\lambda_2)u_1}{\rho_n}} e^{i\frac{\lambda_2 u_2}{\rho_n}} e^{-i\frac{\lambda_2 u_3}{\rho_n}} \right. \\
& \left. + e^{i\frac{(\lambda_1+\lambda_2)u_1}{\rho_n}} e^{-i\frac{\lambda_2 u_2}{\rho_n}} e^{i\frac{\lambda_2 u_3}{\rho_n}} + e^{-i\frac{(\lambda_1+\lambda_2)u_1}{\rho_n}} e^{i\frac{\lambda_2 u_2}{\rho_n}} e^{-i\frac{\lambda_2 u_3}{\rho_n}} \right\} \\
=& T_{11}(\lambda_1, \lambda_2) + T_{12}(\lambda_1, \lambda_2) + T_{13}(\lambda_1, \lambda_2) + T_{14}(\lambda_1, \lambda_2),
\end{aligned} \tag{A.3}$$

where the four terms correspond to the four summands appearing within braces in the last factor on the right hand side of (A.3).

By using the results of Lemmas 1 and 2 given below, we have the convergence

$$\lim_{n\to\infty} nb_n T_{11}(\lambda_1, \lambda_2) = \frac{1}{2} B\phi_{a_1a_3}(\lambda_2)\phi^*_{a_2a_4}(\lambda_2) 1_{E_2\cup E_4}(\lambda_1, \lambda_2)$$

and similar arguments show that

$$\begin{aligned}
\lim_{n\to\infty} nb_n T_{12}(\lambda_1, \lambda_2) &= \frac{1}{2} B\phi^*_{a_1a_3}(\lambda_2)\phi_{a_2a_4}(\lambda_2) 1_{E_2\cup E_4}(\lambda_1, \lambda_2), \\
\lim_{n\to\infty} nb_n T_{13}(\lambda_1, \lambda_2) &= \frac{1}{2} B\phi_{a_1a_3}(\lambda_2)\phi^*_{a_2a_4}(\lambda_2) 1_{E_3\cup E_4}(\lambda_1, \lambda_2) \\
& \text{and} \\
\lim_{n\to\infty} nb_n T_{14}(\lambda_1, \lambda_2) &= \frac{1}{2} B\phi^*_{a_1a_3}(\lambda_2)\phi_{a_2a_4}(\lambda_2) 1_{E_3\cup E_4}(\lambda_1, \lambda_2).
\end{aligned}$$

For the function $T_2(\lambda_1, \lambda_2)$, one can similarly use the transformations $u_1 = t_1 - t_2$, $u_2 = t_1 - t_4$ and $u_3 = t_2 - t_3$, interchange the order of summation and expand the cosine functions in terms of complex exponentials to obtain

$$
\begin{aligned}
&T_2(\lambda_1,\lambda_2)\\
=&\frac{1}{(2\pi)^2 n\rho_n^2}\sum_{u_1=-(n-1)}^{(n-1)}\sum_{u_2=-(n-1)}^{(n-1)}\sum_{u_3=-(n-1)-u_1}^{(n-1)-u_1} U_n(u_1,u_2,u_3)K(b_nu_1)K(b_n(-u_1+u_2-u_3))\\
&\times C_{a_1a_4}\left(\frac{u_2}{\rho_n}\right)C_{a_2a_3}\left(\frac{u_3}{\rho_n}\right)\cos\left(\frac{u_1\lambda_1}{\rho_n}\right)\cos\left(\frac{(-u_1+u_2-u_3)\lambda_2}{\rho_n}\right)\\
=&\frac{1}{(4\pi)^2 n\rho_n^2}\sum_{u_1=-(n-1)}^{(n-1)}\sum_{u_2=-(n-1)}^{(n-1)}\sum_{u_3=-(n-1)-u_1}^{(n-1)-u_1} U_n(u_1,u_2,u_3)K(b_nu_1)K(b_n(-u_1+u_2-u_3))\\
&\times C_{a_1a_4}\left(\frac{u_2}{\rho_n}\right)C_{a_2a_3}\left(\frac{u_3}{\rho_n}\right)\left\{e^{-i\frac{(\lambda_1-\lambda_2)u_1}{\rho_n}}e^{-i\frac{\lambda_2u_2}{\rho_n}}e^{i\frac{\lambda_2u_3}{\rho_n}}+e^{i\frac{(\lambda_1-\lambda_2)u_1}{\rho_n}}e^{i\frac{\lambda_2u_2}{\rho_n}}e^{-i\frac{\lambda_2u_3}{\rho_n}}\right.\\
&\left.+e^{i\frac{(\lambda_1+\lambda_2)u_1}{\rho_n}}e^{-i\frac{\lambda_2u_2}{\rho_n}}e^{i\frac{\lambda_2u_3}{\rho_n}}+e^{-i\frac{(\lambda_1+\lambda_2)u_1}{\rho_n}}e^{i\frac{\lambda_2u_2}{\rho_n}}e^{-i\frac{\lambda_2u_3}{\rho_n}}\right\}\\
&=T_{21}(\lambda_1,\lambda_2)+T_{22}(\lambda_1,\lambda_2)+T_{23}(\lambda_1,\lambda_2)+T_{24}(\lambda_1,\lambda_2).
\end{aligned}
$$

By using similar arguments as in the case of $nb_nT_{11}(\lambda_1,\lambda_2)$, it can be shown that

$$
\begin{aligned}
\lim_{n\to\infty} nb_nT_{21}(\lambda_1,\lambda_2) &= \frac{1}{2}B\phi_{a_1a_4}(\lambda_2)\phi^*_{a_2a_3}(\lambda_2)1_{E_2\cup E_4}(\lambda_1,\lambda_2),\\
\lim_{n\to\infty} nb_nT_{22}(\lambda_1,\lambda_2) &= \frac{1}{2}B\phi^*_{a_1a_4}(\lambda_2)\phi_{a_2a_3}(\lambda_2)1_{E_2\cup E_4}(\lambda_1,\lambda_2),\\
\lim_{n\to\infty} nb_nT_{23}(\lambda_1,\lambda_2) &= \frac{1}{2}B\phi_{a_1a_4}(\lambda_2)\phi^*_{a_2a_3}(\lambda_2)1_{E_3\cup E_4}(\lambda_1,\lambda_2)\\
&\text{and}\\
\lim_{n\to\infty} nb_nT_{24}(\lambda_1,\lambda_2) &= \frac{1}{2}B\phi^*_{a_1a_4}(\lambda_2)\phi_{a_2a_3}(\lambda_2)1_{E_3\cup E_4}(\lambda_1,\lambda_2).
\end{aligned}
$$

Finally, for the term $T_3(\lambda_1,\lambda_2)$, we use the transformations $u_1=t_1-t_4$, $u_2=t_2-t_4$ and $u_3=t_3-t_4$ and interchange the order of summations to have

$$
\begin{aligned}
&T_3(\lambda_1,\lambda_2)\\
=&\frac{1}{(2\pi)^2(n\rho_n)^2}\sum_{u_1=-(n-1)}^{(n-1)}\sum_{u_2=-(n-1)}^{n-1}\sum_{u_3=-(n-1)}^{(n-1)}\{n-\min_i(u_i)+\max_i(u_i)\}K(b_n(u_1-u_2))K(b_nu_3)\\
&\times C_{a_1a_2a_3a_4}\left(\frac{u_1}{\rho_n},\frac{u_2}{\rho_n},\frac{u_3}{\rho_n}\right)\cos\left(\frac{(u_1-u_2)\lambda_1}{\rho_n}\right)\cos\left(\frac{u_3\lambda_2}{\rho_n}\right).
\end{aligned}
$$

From Assumptions 2 and 5, we have

$$
nb_n|T_3(\lambda_1,\lambda_2)|\le\rho_nb_nM^2\sum_{u_1=-(n-1)}^{n-1}\sum_{u_2=-(n-1)}^{n-1}\sum_{u_2=-(n-1)}^{n-1}g_{a_1}\left(\frac{u_1}{\rho_n}\right)g_{a_2}\left(\frac{u_2}{\rho_n}\right)g_{a_3}\left(\frac{u_3}{\rho_n}\right)\frac{1}{\rho_n^3}. \tag{A.4}
$$

Now consider the function $S_n(\cdot)$ defined over $\mathbb{R}$ as

$$S_n(x) = \sum_{u_1=-(n-1)}^{n-1} g_{a_1}\left(\frac{u_1}{\rho_n}\right) 1_{(\frac{u_1-1}{\rho_n}, \frac{u_1}{\rho_n}]}(x).$$

Observe that $\lim_{n\to\infty} S_n(x) = g_{a_1}(x)$ and $|S_n(x)|$ is dominated by $g_{a_1}(\cdot)$. By applying DCT, we have

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} S_n(x)dx = \lim_{n\to\infty} \sum_{u_1=-(n-1)}^{n-1} g_{a_1}\left(\frac{u_1}{\rho_n}\right)\frac{1}{\rho_n} = \int_{-\infty}^{\infty} g_{a_1}(x)dx.$$

Thus, the upper bound of $nb_nT_3(\lambda_1,\lambda_2)$ given by (A.4) is $O(\rho_n b_n)$. Assumption 4 ensures that $nb_nT_3(\lambda_1,\lambda_2)$ converges to zero uniformly.

By combining all these terms, we have the convergence of $nb_nCov\left[Re(\widehat{\phi}_{a_1a_2}(\lambda_1)), Re(\widehat{\phi}_{a_3a_4}(\lambda_2))\right]$ as given in the theorem. Convergence of the other three covariances follow from a similar argument.

□

LEMMA 1. *For $\lambda_1 - \lambda_2 = 0$, the function $T_{11}(\lambda_1,\lambda_2)$ converges as follows.*

$$\lim_{n\to\infty} nb_nT_{11}(\lambda_1,\lambda_2) = \frac{1}{2}B\phi_{a_1a_3}(\lambda_2)\phi^*_{a_2a_4}(\lambda_2).$$

The convergence is uniform on any compact subset of the set

$$E = \{(\lambda_1,\lambda_2) : \lambda_1 - \lambda_2 = 0, \ -\infty < \lambda_1,\lambda_2 < \infty\}.$$

PROOF OF LEMMA 1. Consider a compact subset E' of the set E. Consider the simple function $S_n(\cdot)$, defined over $E' \times \mathbb{R}^3$ by

$$\begin{aligned}
&S_n(\lambda_1,\lambda_2,x_1,x_2,x_3) \\
&= \sum_{u_1=-(n-1)}^{(n-1)} \sum_{u_2=-(n-1)}^{(n-1)} \sum_{u_3=-(n-1)-u_1}^{(n-1)-u_1} U_n(u_1,u_2,u_3)K(b_nu_1)K(b_n(u_1-u_2+u_3))C_{a_1a_3}\left(\frac{u_2}{\rho_n}\right) \\
&\quad \times e^{-i\frac{u_2\lambda_2}{\rho_n}} C_{a_2a_4}\left(\frac{u_3}{\rho_n}\right) e^{i\frac{u_3\lambda_2}{\rho_n}} 1_{((u_1-1)b_n, u_1b_n]}(x_1) 1_{\left(\frac{u_2-1}{\rho_n}, \frac{u_2}{\rho_n}\right]}(x_2) 1_{\left(\frac{u_3-1}{\rho_n}, \frac{u_3}{\rho_n}\right]}(x_3).
\end{aligned}$$

So that

$$nb_nT_{11}(\lambda_1,\lambda_2) = \frac{1}{(4\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S_n(\lambda_1,\lambda_2,x_1,x_2,x_3)dx_1dx_2dx_3.$$

Define $u_{1n}(x_1)$, $u_{2n}(x_2)$ and $u_{3n}(x_3)$ as the smallest integers greater than or equal to x_1/b_n, ρ_nx_2 and ρ_nx_3, respectively. Thus, $(x_1,x_2,x_3) \in (b_nu_{1n-1}(x_1), b_nu_{1n}(x_1)] \times \left(\frac{u_{2n-1}(x_2)}{\rho_n}, \frac{u_{2n}(x_2)}{\rho_n}\right] \times$

$\left(\frac{u_{3n-1}(x_3)}{\rho_n}, \frac{u_{3n}(x_3)}{\rho_n}\right]$ and $b_n u_{1n}(x_1) \to x_1$, $\frac{u_{2n}(x_2)}{\rho_n} \to x_2$, $\frac{u_{3n}(x_3)}{\rho_n} \to x_3$ as $n \to \infty$. Since $nb_n \to \infty$ and $b_n\rho_n \to 0$ as $n \to \infty$, we have, for any point $(x_1, x_2, x_3) \in \mathbb{R}^3$ and large enough n, the inequalities $-\frac{nb_n - x_1}{b_n\rho_n} < x_3 < \frac{nb_n - x_1}{b_n\rho_n}$, i.e., $-n+1-u_{1n}(x_1) < u_{3n}(x_3) < n-1-u_{1n}(x_1)$. Thus, for sufficiently large n, we have

$$\begin{aligned}
&S_n(\lambda_1, \lambda_2, x_1, x_2, x_3)\\
=&U_n(u_{1n}(x_1), u_{2n}(x_2), u_{3n}(x_3))K(b_n u_{1n}(x_1))K(b_n(u_{1n}(x_1) - u_{2n}(x_2) + u_{3n}(x_3)))\\
&\times C_{a_1a_3}\left(\frac{u_{2n}(x_2)}{\rho_n}\right) e^{-i\frac{u_{2n}(x_2)\lambda_2}{\rho_n}} C_{a_2a_4}\left(\frac{u_{3n}(x_3)}{\rho_n}\right) e^{i\frac{u_{3n}(x_3)\lambda_2}{\rho_n}}.
\end{aligned} \tag{A.5}$$

Observe that, under Assumptions 1,3 and 4, the function $S_n(\lambda_1, \lambda_2, x_1, x_2, x_3)$ converges to the function $S(\cdot)$, defined over $E' \times \mathbb{R}^3$ by

$$S(\lambda_1, \lambda_2, x_1, x_2, x_3) = K^2(x_1)C_{a_1a_3}(x_2)\, e^{-ix_2\lambda_2} C_{a_2a_4}(x_3)\, e^{ix_3\lambda_2}.$$

Note that $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S(\lambda_1, \lambda_2, x_1, x_2, x_3)dx_1dx_2dx_3$ is a continuous function in (λ_1, λ_2). As in the proof of Theorem 1, we prove the convergence of the left hand side of (A.5) uniformly on E', by showing that for any sequence $(\lambda_{1n}, \lambda_{2n}) \to (\lambda_1, \lambda_2)$,

$$\begin{aligned}
&\lim_{n\to\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S_n(\lambda_{1n}, \lambda_{2n}, x_1, x_2, x_3)dx_1dx_2dx_3\\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S(\lambda_1, \lambda_2, x_1, x_2, x_3)dx_1dx_2dx_3.
\end{aligned}$$

for $(\lambda_{1n}, \lambda_{2n}), (\lambda_1, \lambda_2) \in E'$. The latter convergence follows, through Assumption 1 and 2 and the DCT, from the dominance

$$|S_n(\lambda_{1n}, \lambda_{2n}, x_1, x_2, x_3)| \le MK_1(x_1)g_{a_1a_3}(x_2)\, g_{a_2a_4}(x_3)\,.$$

and the convergence of the integrand, which holds because of the continuity of $C_{a_1a_3}(\cdot)$, $C_{a_2a_4}(\cdot)$ and the kernel and the exponential functions. Hence, $nb_nT_{11}(\cdot)$ converges as stated uniformly on the compact set E'. □

LEMMA 2. *For $\lambda_1 - \lambda_2 \neq 0$, the function $nb_nT_{11}(\lambda_1, \lambda_2)$ converges to zero. The convergence is uniform on any compact subset of the set E_1 given by*

$$E = \{(\lambda_1, \lambda_2) : \lambda_1 - \lambda_2 \neq 0,\ -\infty < \lambda_1, \lambda_2 < \infty\}.$$

PROOF OF LEMMA 2. Let E' be any compact subset of the set E. Consider the simple function $S_n(\cdot)$, defined over $E' \times \mathbb{R}^3$ by

$$
\begin{aligned}
&S_n(\lambda_1,\lambda_2,x_1,x_2,x_3)\\
&=\sum_{u_1=-(n-1)}^{(n-1)}\sum_{u_2=-(n-1)}^{(n-1)}\sum_{u_3=-(n-1)-u_1}^{(n-1)-u_1} U_n(u_1,u_2,u_3)K(b_nu_1)K(b_n(u_1-u_2+u_3))e^{-i\frac{u_1(\lambda_1-\lambda_2)}{\rho_n}}\\
&\quad\times C_{a_1a_3}\left(\frac{u_2}{\rho_n}\right)e^{-i\frac{u_2\lambda_2}{\rho_n}}C_{a_2a_4}\left(\frac{u_3}{\rho_n}\right)e^{i\frac{u_3\lambda_2}{\rho_n}}1_{((u_1-1)b_n,u_1b_n)]}(x_1)1_{\left(\frac{(u_2-1)}{\rho_n},\frac{u_2}{\rho_n}\right]}(x_2)1_{\left(\frac{(u_3-1)}{\rho_n},\frac{u_3}{\rho_n}\right]}(x_3).
\end{aligned}
$$

So that

$$
nb_nT_{11}(\lambda_1,\lambda_2)=\frac{1}{(4\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}S_n(\lambda_1,\lambda_2,x_1,x_2,x_3)dx_1dx_2dx_3.
$$

An argument similar to that used in the proof of Lemma 1 shows that for $(x_1,x_2,x_3)\in\mathbb{R}^3$ and sufficiently large n,

$$
\begin{aligned}
&S_n(\lambda_1,\lambda_2,x_1,x_2,x_3)\\
&=U_n(u_{1n}(x_1),u_{2n}(x_2),u_{3n}(x_3))K(b_nu_{1n}(x_1))K(b_n(u_{1n}(x_1)-u_{2n}(x_2)+u_{3n}(x_3)))e^{-i\frac{u_{1n}(x_1)(\lambda_1-\lambda_2)}{\rho_n}}\\
&\quad\times C_{a_1a_3}\left(\frac{u_{2n}(x_2)}{\rho_n}\right)e^{-i\frac{u_{2n}(x_2)\lambda_2}{\rho_n}}C_{a_2a_4}\left(\frac{u_{3n}(x_3)}{\rho_n}\right)e^{i\frac{u_{3n}(x_3)\lambda_2}{\rho_n}}.
\end{aligned}
$$

where $u_{1n}(x_1)$, $u_{2n}(x_2)$ and $u_{3n}(x_3)$ are the smallest integers greater than or equal to x_1/b_n, ρ_nx_2 and ρ_nx_3, respectively.

For obtaining the uniform convergence of $nb_nT_{11}(\lambda_1,\lambda_2)$, consider

$$
\begin{aligned}
&\sup_{(\lambda_1,\lambda_2)\in E'}\left|\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}S_n(\lambda_1,\lambda_2,x_1,x_2,x_3)dx_1dx_2dx_3\right|\\
&\le\sup_{(\lambda_1,\lambda_2)\in E'}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\Big|\,S_n(\lambda_1,\lambda_2,x_1,x_2,x_3)-g_n(\lambda_1,\lambda_2,x_1,x_2,x_3)\,\Big|\,dx_1dx_2dx_3\\
&\quad+\sup_{(\lambda_1,\lambda_2)\in E'}\left|\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}g_n(\lambda_1,\lambda_2,x_1,x_2,x_3)dx_1dx_2dx_3\right|,
\end{aligned}
\tag{A.6}
$$

where the function $g_n(\cdot)$ is defined over $E'\times\mathbb{R}^3$ by

$$
g_n(\lambda_1,\lambda_2,x_1,x_2,x_3)=K^2(x_1)e^{-i\frac{x_1(\lambda_1-\lambda_2)}{b_n\rho_n}}C_{a_1a_3}(x_2)\,e^{-ix_2\lambda_2}C_{a_2a_4}(x_3)\,e^{ix_3\lambda_2}.
$$

We will show the uniform convergence of the right hand side of (A.6) by considering the two terms separately. For the first term, we follow the route taken in the proof of Theorem 1, i.e., show that for any sequence $(\lambda_{1n},\lambda_{2n})\to(\lambda_1,\lambda_2)$,

$$
\lim_{n\to\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\Big|\,S_n(\lambda_{1n},\lambda_{2n},x_1,x_2,x_3)-g_n(\lambda_{1n},\lambda_{2n},x_1,x_2,x_3)\,\Big|\,dx_1dx_2dx_3=0,
$$

for $(\lambda_{1n}, \lambda_{2n}), (\lambda_1, \lambda_2) \in E'$. For this purpose, we write the above integral as

$$\begin{aligned}
&\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Big| S_n(\lambda_{1n}, \lambda_{2n}, x_1, x_2, x_3) - g_n(\lambda_{1n}, \lambda_{2n}, x_1, x_2, x_3) \Big| \, dx_1 dx_2 dx_3 \\
&\leq \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Big| S_n(\lambda_{1n}, \lambda_{2n}, x_1, x_2, x_3) - G_n(\lambda_{1n}, \lambda_{2n}, x_1, x_2, x_3) \Big| \, dx_1 dx_2 dx_3 \\
&\quad + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Big| G_n(\lambda_{1n}, \lambda_{2n}, x_1, x_2, x_3) - g_n(\lambda_{1n}, \lambda_{2n}, x_1, x_2, x_3) \Big| \, dx_1 dx_2 dx_3, \quad \text{(A.7)}
\end{aligned}$$

where the function $G_n(\cdot)$ is defined over $E' \times \mathbb{R}^3$ by

$$G_n(\lambda_1, \lambda_2, x_1, x_2, x_3) = K^2(x_1) e^{-i\frac{u_{1n}(x_1) b_n (\lambda_1 - \lambda_2)}{b_n \rho_n}} C_{a_1 a_3}(x_2)\, e^{-ix_2\lambda_2} C_{a_2 a_4}(x_3)\, e^{ix_3\lambda_2}.$$

Now observe that

$$\begin{aligned}
&|S_n(\lambda_{1n}, \lambda_{2n}, x_1, x_2, x_3) - G_n(\lambda_{1n}, \lambda_{2n}, x_1, x_2, x_3)| \\
&\qquad \leq M \left| e^{-i\frac{u_{1n}(x_1) b_n (\lambda_{1n} - \lambda_{2n})}{b_n \rho_n}} \alpha_n(\lambda_{1n}, \lambda_{2n}, x_1, x_2, x_3) \right|,
\end{aligned}$$

where

$$\begin{aligned}
&\alpha_n(\lambda_{1n}, \lambda_{2n}, x_1, x_2, x_3) \\
=&U_n(u_{1n}(x_1), u_{2n}(x_2), u_{3n}(x_3)) K(b_n u_{1n}(x_1)) K(b_n(u_{1n}(x_1) - u_{2n}(x_2) + u_{3n}(x_3))) C_{a_1 a_3}\left(\frac{u_{2n}(x_2)}{\rho_n}\right) \\
&\times e^{-i\frac{u_{2n}(x_2)\lambda_{2n}}{\rho_n}} C_{a_2 a_4}\left(\frac{u_{3n}(x_3)}{\rho_n}\right) e^{i\frac{u_{3n}(x_3)\lambda_{2n}}{\rho_n}} - K^2(x_1) C_{a_1 a_3}(x_2)\, e^{-ix_2\lambda_{2n}} C_{a_2 a_4}(x_3)\, e^{ix_3\lambda_{2n}}.
\end{aligned}$$

Since $\alpha_n(\lambda_n, x, t, t') \to 0$ as $n \to \infty$, we have

$$\lim_{n\to\infty} |S_n(\lambda_{1n}, \lambda_{2n}, x_1, x_2, x_3) - G_n(\lambda_{1n}, \lambda_{2n}, x_1, x_2, x_3)| = 0$$

Since from Assumption 1 and 2, we have the dominance

$$|S_n(\lambda_{1n}, \lambda_{2n}, x_1, x_2, x_3) - G_n(\lambda_{1n}, \lambda_{2n}, x_1, x_2, x_3)| \leq 2MK_1(x_1) g_{a_1 a_3}(x_2) g_{a_2 a_4}(x_2).$$

By applying the DCT, we have

$$\lim_{n\to\infty} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Big| S_n(\lambda_{1n}, \lambda_{2n}, x_1, x_2, x_3) - G_n(\lambda_{1n}, \lambda_{2n}, x_1, x_2, x_3) \Big| \, dx_1 dx_2 dx_3 = 0.$$

Turning to the second term on the right hand side of (A.7), observe that for any fixed x_1,

$$\left| e^{-i\frac{u_{1n}(x_1) b_n (\lambda_{1n} - \lambda_{2n})}{b_n \rho_n}} - e^{-i\frac{x_1(\lambda_{1n} - \lambda_{2n})}{b_n \rho_n}} \right| \leq \frac{\lambda_{1n} - \lambda_{2n}}{\rho_n}.$$

Thus,

$$
\begin{aligned}
&|G_n(\lambda_{1n}, \lambda_{2n}, x_1, x_2, x_3) - g_n(\lambda_{1n}, \lambda_{2n}, x_1, x_2, x_3)| \\
\leq & M^2 g_{a_1a_3}(0) g_{a_2a_4}(0) \left| e^{-i\frac{u_{1n}(x_1)b_n(\lambda_{1n}-\lambda_{2n})}{b_n\rho_n}} - e^{-i\frac{x_1(\lambda_{1n}-\lambda_{2n})}{b_n\rho_n}} \right| \\
\leq & M^2 g_{a_1a_3}(0) g_{a_2a_4}(0) \frac{\lambda_{1n} - \lambda_{2n}}{\rho_n},
\end{aligned}
$$

and so

$$
\lim_{n\to\infty} |G_n(\lambda_{1n}, \lambda_{2n}, x_1, x_2, x_3) - g_n(\lambda_{1n}, \lambda_{2n}, x_1, x_2, x_3)| = 0.
$$

From Assumption 1 and 2, we have the dominance

$$
|G_n(\lambda_{1n}, \lambda_{2n}, x_1, x_2, x_3) - g_n(\lambda_{1n}, \lambda_{2n}, x_1, x_2, x_3)| \leq 2MK_1(x_1) g_{a_1a_3}(x_2) g_{a_2a_4}(x_2).
$$

which leads us, through another use of the DCT, to the convergence of the second integral of (A.7). This establishes that the first term on the right hand side of (A.6) converges to 0. We only have to deal with the second term. Let

$$
s_n(\lambda_1, \lambda_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_n(\lambda_1, \lambda_2, x_1, x_2, x_3) dx_1 dx_2 dx_3.
$$

In order to establish the uniform convergence of $s_n(\cdot)$ over E', it is enough to show that $s_n(\lambda_{1n}, \lambda_{2n}) \to 0$ for any sequence $(\lambda_{1n}, \lambda_{2n}) \to (\lambda_1, \lambda_2)$, where $(\lambda_{1n}, \lambda_{2n}), (\lambda_1, \lambda_2) \in E'$. By using the Reimann-Lebesgue lemma, we have $s_n(\lambda_1, \lambda_2) \to 0$. Thus, the second term on the right hand side of (A.6) also converges to 0. Hence, $nb_nT_{11}(\lambda_1, \lambda_2)$ converges to 0 uniformly on E' as $n \to \infty$. □

In order to prove Theorem 3, we will need the following lemma, which describes the asymptotic behaviour of the joint cumulants of the estimators $\widehat{\phi}_{a_1a_2}(\cdot)$ for $a_1, a_2 \in \{1, 2, \ldots, r\}$. In the present case, a cumulant defined as in (2) may be complex-valued.

LEMMA 3. *Under the Assumptions 1–4 and 5A, for $L > 2$, the Lth order joint cumulant of the vector $\left(\widehat{\phi}_{a_1a_2}(\lambda_1), \ \ldots, \widehat{\phi}_{a_{2L-1}a_{2L}}(\lambda_L)\right)$ for $a_1, a_2, \ldots, a_{2L} \in \{1, 2, \ldots, r\}$ is bounded from above as follows.*

$$
\left| cum\left(\widehat{\phi}_{a_1a_2}(\lambda_1), \ldots, \widehat{\phi}_{a_{2L-1}a_{2L}}(\lambda_L)\right)\right| \leq Q \cdot (nb_n)^{-(L-1)}, \tag{A.8}
$$

where the constant Q does not depend on $\lambda_1, \ldots, \lambda_L$.

PROOF. $cum(\widehat{\phi}_{a_1a_2}(\lambda_1), \widehat{\phi}_{a_3a_4}(\lambda_2), \ldots, \widehat{\phi}_{a_{2L-1}a_{2L}}(\lambda_L))$ can be written as

$$\begin{aligned}
&cum(\widehat{\phi}_{a_1a_2}(\lambda_1), \widehat{\phi}_{a_3a_4}(\lambda_2), \ldots, \widehat{\phi}_{a_{2L-1}a_{2L}}(\lambda_L)) \\
&= \frac{1}{(\pi n\rho_n)^L} \sum_{t_1=1}^{n}\sum_{t_2=1}^{n} \cdots \sum_{t_{2L-1}=1}^{n}\sum_{t_{2L}=1}^{n} K(b_n(t_1-t_2)) \ldots K(b_n(t_{2L-1}-t_{2L})) \\
&\times e^{-\frac{i(t_1-t_2)\lambda_1}{\rho_n}} \times \cdots e^{-\frac{i(t_{2L-1}-t_{2L})\lambda_L}{\rho_n}} cum\left(X_{a_1}\left(\frac{t_1}{\rho_n}\right)X_{a_2}\left(\frac{t_2}{\rho_n}\right), \ldots, X_{a_{2L-1}}\left(\frac{t_{2L-1}}{\rho_n}\right)X_{a_{2L}}\left(\frac{t_{2L}}{\rho_n}\right)\right)
\end{aligned} \tag{A.9}$$

It follows that

$$\begin{aligned}
&|cum(\widehat{\phi}_{a_1a_2}(\lambda_1), \widehat{\phi}_{a_3a_4}(\lambda_2), \ldots, \widehat{\phi}_{a_{2L-1}a_{2L}}(\lambda_L))| \\
&\leq \frac{1}{(n\rho_n)^L} \sum_{t_1=1}^{n}\sum_{t_2=1}^{n} \cdots \sum_{t_{2L-1}=1}^{n}\sum_{t_{2L}=1}^{n} |K(b_n(t_1-t_2)) \ldots K(b_n(t_{2L-1}-t_{2L}))| \\
&\times\left|cum\left(X_{a_1}\left(\frac{t_1}{\rho_n}\right)X_{a_2}\left(\frac{t_2}{\rho_n}\right), \ldots, X_{a_{2L-1}}\left(\frac{t_{2L-1}}{\rho_n}\right)X_{a_{2L}}\left(\frac{t_{2L}}{\rho_n}\right)\right)\right|
\end{aligned}$$

Now

$$\begin{aligned}
&cum\left(X_{a_1}\left(\frac{t_1}{\rho_n}\right)X_{a_2}\left(\frac{t_2}{\rho_n}\right), \ldots, X_{a_{2L-1}}\left(\frac{t_{2L-1}}{\rho_n}\right)X_{a_{2L}}\left(\frac{t_{2L}}{\rho_n}\right)\right) \\
&= \sum_{\boldsymbol{\nu}} C_{a_{j_{11}}a_{j_{12}}\ldots a_{j_{1k_1}}}\left(\frac{t_{j_{11}}-t'_1}{\rho_n}, \ldots, \frac{t_{j_{1,k_1-1}}-t'_1}{\rho_n}\right) \cdots C_{a_{j_{P1}}a_{j_{P2}}\ldots a_{j_{Pk_P}}}\left(\frac{t_{j_{P1}}-t'_P}{\rho_n}, \ldots, \frac{t_{j_{P,k_P-1}}-t'_P}{\rho_n}\right)
\end{aligned}$$

where the summation is over all *indecomposable* ([4], [27]) partitions $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_P)$, such that $\nu_p = (j_{p1}, \ldots, j_{pk_p})$, $p = 1, \ldots, P$, of the table

1	2
3	4
⋮	⋮
2L-1	2L

and $t'_p = t_{j_{pk_p}}$, $p = 1, \ldots, P$. Since the partition $\boldsymbol{\nu}$ is indecomposable, we have $t_{j_{pl}} - t'_p \neq t_{2m} - t_{2m-1}$, for $l = 1, \ldots, k_p;\ p = 1, \ldots, P;\ m = 1, \ldots, L$.

Define

$$u_{j_{pl}} = t_{j_{pl}} - t'_p;\ l = 1, \ldots, k_p; p = 1, \ldots, P.$$

Note that $u_{j_{pk_p}} = 0$ for $p = 1, \ldots, P$. Then the joint cumulant of $(\widehat{\phi}_{a_1a_2}(\lambda_1)$, $\widehat{\phi}_{a_3a_4}(\lambda_2)$, …,

$\widehat{\phi}_{a_{2L-1}a_{2L}}(\lambda_L))$ given by (A.9) is absolutely bounded by

$$\frac{1}{(n\rho_n)^L}\sum_{\boldsymbol{\nu}}\sum_{t'_1=1}^{n}\sum_{u_{j_{11}}=-(t'_1-1)}^{n-t'_1}\cdots\sum_{u_{j_{1,k_1-1}}=-(t'_1-1)}^{n-t'_1}\cdots\sum_{t'_P=1}^{n}\sum_{u_{j_{P1}}=-(t'_P-1)}^{n-t'_P}\cdots\sum_{u_{j_{P,k_P-1}}=-(t'_P-1)}^{n-t'_P}$$
$$\left| K[b_n(u_1+t'_{p_1}-u_2-t'_{p_2})]\cdots K[b_n(u_{2L-1}+t'_{p_{2L-1}}-u_{2L}-t'_{p_{2L}})]\right|$$
$$\times\left|C_{a_{j_{11}}a_{j_{12}}\dots a_{j_{1k_1}}}\left(\frac{u_{j_{11}}}{\rho_n},\dots,\frac{u_{j_{1,k_1-1}}}{\rho_n}\right)\dots C_{a_{j_{P1}}a_{j_{P2}}\dots a_{j_{Pk_P}}}\left(\frac{u_{j_{P1}}}{\rho_n},\dots,\frac{u_{j_{P,k_P-1}}}{\rho_n}\right)\right|, \tag{A.10}$$

where p_m is that member of the set $\{1,2,\dots,P\}$ which satisfies $t_m\in\nu_{p_m}$ for $m=1,\dots,L$.

We will now show that the set $A=\{t'_{p_1}-t'_{p_2},\dots,t'_{p_{2L-1}}-t'_{p_{2L}}\}$ has $P-1$ linearly independent elements. Note that the set A consists of differences of pairs of elements of the set $\{t'_1,t'_2,\dots,t'_P\}$. So the set A can have at most $P-1$ linearly independent differences. Suppose that the set A has exactly $P-j$ linearly independent differences for some $j\geq 1$. Denote the $P-j$ independent differences of the set A by

$$A_1=\left\{t'_{p_{2k_1-1}}-t'_{p_{2k_1}},t'_{p_{2k_2-1}}-t'_{p_{2k_2}},\dots,t'_{p_{2k_{P-j}-1}}-t'_{p_{2k_{P-j}}}\right\},$$

where $k_1,\dots,k_{P-j}\in\{1,2,\dots,L\}$. Let, if possible, $j>1$, and consider a difference $t'_{l_1}-t'_{l_2}$ for $l_1,l_2\in\{1,2,\dots,P\}$ which is linearly independent of the elements of the set A_1. Since the partition $\boldsymbol{\nu}$ is indecomposable, the sets ν_{l_1} and ν_{l_2} *communicate* [27]. Therefore, there exists an index set $\{s_1,s_2,\dots,s_r\}$ with $r\geq 2$, which is a proper subset of $\{1,2,\dots,P\}$, such that $s_1=l_1$, $s_r=l_2$ and the pairs $(\nu_{s_1},\nu_{s_2}),(\nu_{s_2},\nu_{s_3}),\dots,(\nu_{s_{r-1}},\nu_{s_r})$ are *hook* [27]. Consequently, there exist indices $j_1,\dots,j_{r-1}\in\{1,\dots,L\}$ such that for $m=1,\dots,r-1$, one of the points t_{2j_m-1} and t_{2j_m} belongs to ν_{s_m} and the other belongs to $\nu_{s_{m+1}}$. It follows that for $m=1,\dots,r-1$, $(t'_{p_{2j_m-1}}-t'_{p_{2j_m}})$ is in A, and hence, they can be written as linear combinations of the members of A_1. Note that for $m=1,\dots,r-1$, $(t'_{s_{m-1}}-t'_{s_m})$ is equal to either $(t'_{p_{2j_m-1}}-t'_{p_{2j_m}})$ or $-(t'_{p_{2j_m-1}}-t'_{p_{2j_m}})$. Thus,

$$t'_{l_1}-t'_{l_2}=t'_{s_1}-t'_{s_r}=(t'_{s_1}-t'_{s_2})+(t'_{s_2}-t'_{s_3})+\cdots+(t'_{s_{r-1}}-t'_{s_r})$$

can be written as a linear combination of the members of A_1. This fact contradicts the assumption that $t'_{l_1}-t'_{l_2}$ is linearly independent of the elements of the set A_1. Therefore, j cannot be larger than 1. This proves that the set A cannot contain fewer than $P-1$ linearly independent differences.

Consider the $P-1$ linearly independent elements of the set A_1, where $j=1$, and define

$$\begin{aligned} v_1 &= u_{2k_1-1}+t'_{p_{2k_1-1}}-u_{2k_1}-t'_{p_{2k_1}}, \\ &\vdots \\ v_{P-1} &= u_{2k_{P-1}-1}+t'_{p_{2k_{P-1}-1}}-u_{2k_{P-1}}-t'_{p_{2k_{P-1}}}. \end{aligned}$$

Using the above transformation, and by replacing the P sums over indices $t'_1,\dots,t'_P$ by $P-1$ sums overs the indices $v_1,\dots,v_{P-1}$, we find that the joint cumulant given in (A.10) is bounded from above by

$$\begin{aligned} &\frac{1}{n^{L-1}\rho_n^L}\sum_{\boldsymbol{\nu}} M^{L-P+1}\sum_{u_{j_{11}}=-(n-1)}^{n-1}\cdots\sum_{u_{j_{1,k_1-1}}=-(n-1)}^{n-1}\cdots\sum_{u_{j_{P1}}=-(n-1)}^{n-1}\cdots\sum_{u_{j_{P,k_P-1}}=-(n-1)}^{n-1} \\ &\sum_{v_1=-3n}^{3n}\cdots\sum_{v_{P-1}=-3n}^{3n}|K(b_nv_1)|\times\cdots\times|K(b_nv_{P-1})| \\ &\times\left|C_{a_{j_{11}}a_{j_{12}}\dots a_{j_{1k_1}}}\left(\frac{u_{j_{11}}}{\rho_n},\dots,\frac{u_{j_{1,k_1-1}}}{\rho_n}\right)\right|\times\cdots\left|C_{a_{j_{P1}}a_{j_{P2}}\dots a_{j_{Pk_P}}}\left(\frac{u_{j_{P1}}}{\rho_n},\dots,\frac{u_{j_{P,k_P-1}}}{\rho_n}\right)\right|. \end{aligned} \tag{A.11}$$

The above simplification has been made by taking into account the upper bound for $L-P+1$ copies of $K(\cdot)$ and conservative estimates of the ranges of summation of $v_1,\dots,v_{P-1}$. Now one can rewrite the expression in (A.11) as follows.

$$\begin{aligned} &\sum_{\boldsymbol{\nu}} M^{L-P+1}\frac{(\rho_nb_n)^{L-P}}{(nb_n)^{L-1}}\left[\sum_{v_1=-3n}^{3n}K(b_nv_1)b_n\right]\times\cdots\times\left[\sum_{v_{P-1}=-3n}^{3n}K(b_nv_{P-1})b_n\right] \\ &\times\left\{\frac{1}{\rho_n^{k_1-1}}\sum_{u_{j_{11}}=-(n-1)}^{n-1}\cdots\sum_{u_{j_{1,k_1-1}}=-(n-1)}^{n-1}\left|C_{a_{j_{11}}a_{j_{12}}\dots a_{j_{1k_1}}}\left(\frac{u_{j_{11}}}{\rho_n},\dots,\frac{u_{j_{1,k_1-1}}}{\rho_n}\right)\right|\right\}\times\cdots \\ &\times\left\{\frac{1}{\rho_n^{k_P-1}}\sum_{u_{j_{P1}}=-(n-1)}^{n-1}\cdots\sum_{u_{j_{P,k_P-1}}=-(n-1)}^{n-1}\left|C_{a_{j_{P1}}a_{j_{P2}}\dots a_{j_{Pk_P}}}\left(\frac{u_{j_{P1}}}{\rho_n},\dots,\frac{u_{j_{P,k_P-1}}}{\rho_n}\right)\right|\right\}. \end{aligned} \tag{A.12}$$

Consider the simple function $S_n(\cdot)$ defined over $\mathbb{R}$ by

$$S_n(x)=\sum_{v_1=-3n}^{3n}K(b_nv_1)1_{(b_nv_1-1,b_nv_1]}(x).$$

Note that $\int_{-\infty}^{\infty}S_n(x)dx=\sum_{v_1=-3n}^{3n}K(b_nv_1)b_n$, and from Assumption 2 we have the dominance $S_n(x)\le K_1(x)$. By applying the DCT, we have

$$\sum_{v_1=-3n}^{3n}K(b_nv_1)b_n\to\int_{-\infty}^{\infty}|K(x)|dx.$$

This fact establishes the convergence of the sums over $v_1, \ldots, v_{P-1}$.

Consider the simple function $T_n(\cdot)$ defined over $\mathbb{R}^{k_1-1}$ by

$$\begin{aligned}
&T_n(x_1, x_2, \ldots, x_{k_1-1}) \\
&= \sum_{u_{j_{11}}=-(n-1)}^{n-1} \cdots \sum_{u_{j_{1,k_1-1}}=-(n-1)}^{n-1} C_{a_{j_{11}} a_{j_{12}} \ldots a_{j_{1k_1}}} \left(\frac{u_{j_{11}}}{\rho_n}, \ldots, \frac{u_{j_{1,k_1-1}}}{\rho_n} \right) \\
&\quad \times 1_{\left(\frac{u_{j_{11}}-1}{\rho_n}, \frac{u_{j_{11}}}{\rho_n} \right]}(x_1) \ldots 1_{\left(\frac{u_{j_{1,k_1-1}}-1}{\rho_n}, \frac{u_{j_{1,k_1-1}}}{\rho_n} \right]}(x_{k_1-1}).
\end{aligned}$$

Note that

$$\begin{aligned}
&\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} T_n(x_1, \ldots, x_{k_1-1}) dx_1 \ldots dx_{k_1-1} \\
&= \frac{1}{\rho_n^{k_1-1}} \sum_{u_{j_{11}}=-(n-1)}^{n-1} \cdots \sum_{u_{j_{1,k_1-1}}=-(n-1)}^{n-1} \left| C_{a_{j_{11}} a_{j_{12}} \ldots a_{j_{1k_1}}} \left(\frac{u_{j_{11}}}{\rho_n}, \ldots, \frac{u_{j_{1,k_1-1}}}{\rho_n} \right) \right|.
\end{aligned}$$

From Assumption 5A, we have that the function $T_n(\cdot)$ is bounded by an integrable function. Thus, by applying the DCT, we have

$$\begin{aligned}
&\lim_{n\to\infty} \frac{1}{\rho_n^{k_1-1}} \sum_{u_{j_{11}}=-(n-1)}^{n-1} \cdots \sum_{u_{j_{1,k_1-1}}=-(n-1)}^{n-1} \left| C_{a_{j_{11}} a_{j_{12}} \ldots a_{j_{1k_1}}} \left(\frac{u_{j_{11}}}{\rho_n}, \ldots, \frac{u_{j_{1,k_1-1}}}{\rho_n} \right) \right| \\
&= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left| C_{a_{j_{11}} a_{j_{12}} \ldots a_{j_{1k_1}}} (x_1, \ldots, x_{k_1-1}) \right| dx_1 \ldots dx_{k_1-1}.
\end{aligned}$$

Likewise, we have the convergence for the remaining $P-1$ sets of summations. Using these above convergence results, the upper bound of (A.10) given in (A.12) can be written as

$$\sum_{\boldsymbol{\nu}} \frac{(\rho_n b_n)^{L-P}}{(nb_n)^{L-1}} d_{\boldsymbol{\nu}},$$

where $d_{\boldsymbol{\nu}}$ are appropriate constants. The summation is over the finite number of indecomposable partitions, and the worst-case value of the partition size P is L. Therefore, the upper bound is $O\left((nb_n)^{-(L-1)}\right)$. This Completes the proof. □

PROOF OF THEOREM 3. Note that the first moment of the random vector on the left hand side of (4) is zero and the second moment converges in accordance with Theorem 2. Further,

$$\begin{aligned}
&cum(c_1(Y_1 - d_1), c_1(Y_2 - d_2), \ldots, c_J(Y_J - d_J)) \\
&= c_1 c_2 \cdots c_J \times cum(Y_1, Y_2, \ldots, Y_J),
\end{aligned}$$

for any set of constants $c_1, \ldots, c_J, d_1, \ldots, d_J$. From the above fact and Lemma 3, for all $k > 2$, the absolute value of the kth order joint cumulant of the random vector on the left hand side of (4) is bounded from above by an $O((nb_n)^{k/2-k+1})$ term. According to Assumption 3, this upper bound tends to 0 as n tends to infinity. This completes the proof. $\square$

PROOF OF THEOREM 4. The result can be proved along the lines of the proof of Theorem 3 of [13]. $\square$

PROOF OF THEOREM 5. The weak convergence of the first term on the right hand side of (5) follows from Theorem 3. On the other hand, the second term can be written, in view of Theorem 4, as

$$\sqrt{nb_n}\left(E[\widehat{\phi}_{a_1a_2}(\lambda_1)] - \phi_{a_1a_2}(\lambda_1)\right) = \sqrt{nb_n}\left(O\left((\rho_n b_n)^q\right) + O\left(\frac{\rho_n}{n}\right) + O\left(\frac{1}{\rho_n^p}\right)\right). \quad \text{(A.13)}$$

Under Assumption 3,

$$\lim_{n\to\infty}\sqrt{nb_n}\rho_n^q b_n^q = 0 \Rightarrow \lim_{n\to\infty}\sqrt{nb_n}\frac{\rho_n}{n} = 0.$$

Therefore, under Assumptions 3 and 4A, the right hand side of (A.13) goes to zero as $n \to \infty$. This completes the proof. $\square$

PROOF OF THEOREM 6. Note that under Assumption 4A, we have

$$\lim_{n\to\infty}\sqrt{nb_n}\frac{1}{\rho_n^p} = 0 \quad \text{(A.14)}$$

and

$$\begin{aligned}
\lim_{n\to\infty}\sqrt{nb_n}\rho_n^q b_n^q = 0 &\Leftrightarrow \lim_{n\to\infty}(nb_n)^{\frac{1}{2q}}\, b_n\rho_n = 0 \\
&\Leftrightarrow \lim_{n\to\infty}(nb_n)^{1+\frac{1}{2q}}\,\frac{\rho_n}{n} = 0 \\
&\Leftrightarrow \lim_{n\to\infty}\sqrt{nb_n}\left(\frac{\rho_n}{n}\right)^{\frac{q}{1+2q}} = 0. \quad \text{(A.15)}
\end{aligned}$$

From (A.14) and (A.15), we have

$$\begin{aligned}
\frac{1}{\sqrt{nb_n}} &= o\left(\left(\frac{\rho_n}{n}\right)^{\frac{q}{1+2q}}\right), && \text{(A.16)} \\
\text{and } \frac{1}{\sqrt{nb_n}} &= o\left(\left(\frac{1}{\rho_n}\right)^{p}\right). && \text{(A.17)}
\end{aligned}$$

The right hand sides of (A.16) and (A.17) are increasing and decreasing functions, respectively, of ρ_n. Assumption 3, together with (A.14), indicate that ρ_n goes to infinity as n goes to infinity. The rate given by (A.16) will be unduly slow if ρ_n goes to infinity too slowly, while the rate given by (A.17) will be unduly slow if ρ_n goes to infinity too fast. At either event, $1/\sqrt{nb_n}$ will have a sub-optimal rate of convergence to zero. It follows that $1/\sqrt{nb_n}$ has the fastest convergence to zero if

$$O\left(\left(\frac{n}{\rho_n}\right)^{\frac{q}{1+2q}}\right) = O\left(\rho_n^p\right).$$

This condition requires that $\rho_n = O\left(n^{\frac{q}{p+q+2pq}}\right)$. For this rate of ρ_n, (A.16) implies that

$$b_n = o\left(n^{-\frac{p+q}{p+q+2pq}}\right) \quad \text{and } \frac{1}{\sqrt{nb_n}} = o\left(n^{-\frac{pq}{p+q+2pq}}\right).$$

This completes the proof. □

References

[1] C. E. Shannon, "Comunication in presence of noise" *Proc. IRE*, vol. 37, pp. 10–21, 1949.

[2] S. M. Kay, *Modern Spectral Estimation: Theory and Application*. Englewood Cliffs, New Jersey: Prentice Hall, 1999.

[3] J. W. H. Swanepoel and J. W. J. van Wyk, "The bootstrap applied to power spectral density function estimation", *Biometrika*, vol. 73, no. 1, pp. 135–141, 1986.

[4] D. R. Brillinger *Time Series Data Analysis and Theory*, Philadelphia: SIAM, 2001.

[5] D. N. Politis, J. P. Romano, T. L. Lai, "Bootstrap confidence bands for spectra and cross-spectra", *IEEE Trans. Signal Processing*, vol. 40 pp. 1206–1215, 1992.

[6] H. S. Shapiro and R. A. Silverman, "Alias-free sampling of random noise", *J. Soc. Indust. Appl. Math*, vol. 8, no. 2, pp. 225–248, 1960.

[7] F. J. Beutler, "Alias-free randomly timed sampling of stochastic processes", *IEEE Trans. Inf. Theor.*, vol. IT-16, no. 2, pp. 147–152, 1970.

[8] E. Masry, "Alias-free sampling: An alternative conceptualization and its applications", *IEEE Trans. Inf. Theor.*, vol. IT-24, no. 3, pp. 173–183, 1978.

[9] K. S. Lii and E. Masry, "Spectral estimation of continuous-time stationary processes from random sampling", *Stochastic Process. Appl.*, vol. 52, issue 1, pp. 39–64, 1994.

[10] H. J. Landau, "Sampling, data transmission, and the Nyquist rate", *Proc. IEEE*, vol. 55, no. 10, pp. 1701–1706, 1967.

[11] F. A. Marvasti, *Nonuniform Sampling*, New York: Kluwer Plenum, 2001.

[12] R. Venkataramani and Y. Bresler, "Optimal sub-Nyquist nonuniform sampling and reconstruction for multiband signals", *IEEE Trans. Signal Processing*, vol. 49, no. 10, pp. 2301–2313, 2001.

[13] R. Srivastava, and D. Sengupta, "Consistent estimation of non-bandlimited spectral density from uniformly spaced samples" *IEEE Trans. Inf. Theor.*, vol. 56, No. 8, pp. 3642–3659, 2010.

[14] M. Fuentes, "Spectral methods for nonstationary spatial processes", *Biometrika*, vol. 89, pp. 197–210, 2002.

[15] A. G. Constantine and P. Hall, "Characterizing surface smoothness via estimation of effective fractal dimension", *J. Roy. Statist. Soc. Ser. B*, vol. 56, pp. 96–113, 1994..

[16] P. Hall, N. I. Fisher and B. Hoffmann, "On the nonparametric estimation of covariance functions", *Ann. Statist.*, vol. 22, pp. 2115–2134, 1994.

[17] S. N. Lahiri, "Asymptotic distribution of the empirical spatial cumulative distribution function predictor and prediction bands based on a subsampling method", *Probab. Theory Related Fields*, vol. 114, pp. 55–84, 1999.

[18] M. Pawlak and U. Stadtmller "Signal sampling and recovery under dependent errors " *IEEE Trans. Inf. Theor.*, vol. 53, No. 7, pp. 2526–2541, 2007.

[19] H. Chen, D. G. Simpson and Z. ying, "Infill asymptotics for a stochastic process model with measurement error", *Statist. Sinica*, vol. 10, pp. 141–156, 2000.

[20] M. L. Stein, "Fixed-domain asymptotics for spatial periodograms", *J. Amer. Statist. Assoc.*, vol. 90, pp. 1277–1288, 1995.

[21] H. Zhang, and D. L. Zimmerman, "Towards reconciling two asymptotic frameworks in spatial statistics", *Biometrika*, vol. 92, pp. 921–936, 2005.

[22] C. Y. Lim, and M. Stein, "Properties of spatial cross-periodograms using fixed-domain asymptotics", *J. Multivariate Anal.*, vol. 99, pp. 1962–1984, 2008.

[23] P. J. Brockwell and R. A. Davis *Time Series: Theory and Methods*, New York: Springer-Verlag, 1991.

[24] E. Parzen, “On consistent estimates of the spectrum of a stationary time series”, *Ann. Math. Statist.*, vol. 28, no. 2, pp. 329–348, 1957.

[25] T. C. M. Lee, “A simple span selector for periodogram smoothing”, *Biometrika*, vol. 84, no. 4, pp. 965–969, 1997.

[26] S. I. Resnick, *Extreme Values, Regular Variation, and Point Processes*. New York: Springer-Verlag, pp. 1–2, 1987.

[27] V. P. Leonov, and A. N. Shiryayev, “On a method of calculation of semi-invariants”, *Theory Probab. Appl.*, vol. 4, pp. 319–329, 1959.