Bandlimited Spectrum Estimation under the Constraint of a Minimum Inter-Sample Spacing

Technical Report No. ASD/2011/1
Dated: 28 February 2011

Radhendushka Srivastava
and
Debasis Sengupta

Indian Statistical Institute
Applied Statistics Unit
Kolkata 700 108

Bandlimited Spectrum Estimation under the Constraint of a Minimum Inter-Sample Spacing

Radhendushka Srivastava and Debasis Sengupta
Applied Statistics Unit, Indian Statistical Institute, Kolkata

Abstract

In the matter of sampling of a continuous time stationary stochastic process for the purpose of spectrum estimation, a practical constraint that comes in the way is that of a minimum separation between successive samples. It is well known that for a given threshold of separation, uniform sampling can be used to estimate only those power spectral densities which are limited to the band implied by the Nyquist theorem. It is known from a recent study [7], that no point process sampling is alias-free for non-bandlimited spectrum in the presence of this constraint. Also known estimation approaches based on stochastic sampling also do not work. In this paper, we propose a new method of bandlimited spectrum estimation based on additive random sampling, after taking into account the said constraint. It is demonstrated through theoretical studies as well as Monte Carlo simulations that, even under the constraint of a specified minimum separation between successive samples, one can accurately and precisely estimate a power spectral density as long as it is limited to an arbitrarily finite bandwidth and the sample size is sufficiently large.

Keywords: Power spectral density, spectrum estimation, renewal process sampling.

1 Introduction

Sampling of a continuous time processes and subsequent inference based on those samples are carried out in many disciplines of science and engineering. The object of inference is typically an attribute of the underlying continuous time process, such as the mean function or the power spectral density. In some applications, samples of the process are collected at uncontrolled observational epochs [citation]. When one can choose a sampling scheme for the observational time points, one typically uses uniform sampling (i.e., sampling at regular intervals) or renewal process sampling (i.e., sampling at the times of renewal of a renewal process) [citation]. In such situations, there is often a practical constraint on the minimum separation between successive samples, which can arise due to a technological constraint or from economic considerations. For example, in paleoclimatic studies based on ice core data [19], the age of a particular sample is determined from radioactive dating techniques. Samples from ice core slices at greater depths are regarded as older. Ice core slices cannot be arbitrarily thin, and this constraint induces a limit on the time separation between successive samples. In computer graphics, where the time parameter is replaced by a two-dimensional space parameter, photo-receptor arrays are designed subject to the restriction of a minimum distance between two receptor, depending on the physical dimensions and specifications of the receptor [15, 16]. The constraint arises due to similar other limitations in laser doppler velocimetry [18] as well as radar applications [4].h

Assuming that the data arise from renewal process sampling of a stationary stochastic process, we develop in this article a method for the estimation of its power spectral density (also referred to as spectrum), subject to the constraint of a minimum inter-sample spacing.

Suppose that the minimum separation between successive samples is d units of time. If one chooses to sample at regular intervals, the fastest possible sampling rate is $1/d$. Thus, only those spectra which are limited to the frequency band $[-\pi/d, \pi/d]$ can be consistently estimated from uniformly spaced samples [1]. A spectrum having larger support than $[-\pi/d, \pi/d]$ would not be distinguishable from a corresponding spectrum having a support contained in that interval, and consequently the bias cannot go to zero even for large sample size. This limitation is related to the problem of aliasing, as described by the Nyquist Theorem, in respect of reconstruction of a function from its uniformly spaced samples. For spectra having larger support than $[-\pi/d, \pi/d]$, one has to look for methods based on alternative sampling schemes consistent estimation of spectra in the presence of the above constraint.

In the absence of any constraint on the minimum inter-sample spacing, Shapiro and Silverman [8] showed that certain renewal process sampling schemes, including Poisson process sampling, are alias-free for the class of non-bandlimited process. Subsequently, consistent methods of spectrum estimation based on the covariance sequence or other aspects of stochastic samples of the process were proposed [13, 5]. These methods typically have larger variance than methods based on regularly spaced samples [15, 16].

In the presence of a constraint on the minimum separation between successive samples, many alias-free sampling schemes, including Poisson sampling, are rendered infeasible. For this reason, researchers have turned to minimum distance Poisson sampling (which discards a sampling time if it is too close to the previous one) [15, 16] or periodic non-uniform sampling [18, 17]. The latter form of sampling, ensuring a minimum separation between successive samples, has also been applied to the problem of curve estimation [2, 3].

In a recent work [7], the authors studied the effect of the inter-sample spacing constraint on spectrum estimation based on stochastically sampled data. They showed that under the said constraint, no point process sampling scheme is alias-free for the class of non-bandlimited processes. This negative result implies that two processes with different spectra cannot be distinguished from point process samples, and hence the question of consistent spectrum estimation from these samples does not arise. They also showed that there exist sampling schemes, which can identify spectra limited to a band that is larger than the Nyquist band, from the covariance sequence of the sampled values. Moreover if one uses the sampling time data together with the sampled values, then it is possible to identify spectra that are limited to any finite bandwidth.

The estimation opportunities created by these theoretical results have so far not been exploited. There are two known approaches for construction of spectrum estimators based on additive random sampling. One of these is based on the estimated covariance sequence of the sampled data [8], while the other uses the sampling times together with the sampled values [5]. It has been shown in [7] that none of the approaches work in the presence of the minimum inter sample spacing constraint.

In this paper, we develop a method of estimating the power spectral density of a stationary process limited to any finite band, based on additive random sampling. We show that the proposed estimator is consistent, under certain regularity conditions on the process and subject to the constraint of a minimum separation between successive samples. This estimator demonstrates that even if there is a restriction on how closely one can sample a continuous time process, a large number of appropriately chosen samples would permit accurate and precise estimation of the underlying spectrum, as long as it is confined to an arbitrarily large but known band.

In Section 2, we develop a spectrum estimator based on constrained additive random sampling, and establish its consistency and rate of convergence. In Section 3, we study the performance of this estimator through Monte Carlo simulations for different set of constraints on inter sample spacing. We provide some concluding remarks in Section 4 and proofs of all the theoretical results in the appendix.

2 Spectrum estimation under the constraint of minimum separation between successive samples

Let $X = \{X(t), \ -\infty < t < \infty\}$ be a real, mean square continuous and wide sense stationary stochastic process with mean zero, covariance function $C(\cdot)$, power spectral density $\phi(\cdot)$ and spectral support $[-\lambda_0, \lambda_0]$. Let $\tau = \{t_n, \ n = \ldots, -2, -1, 0, 1, 2, \ldots\}$ be a sequence of real-valued, sampling times which constitute a stationary renewal process. Assume that the renewal process τ is independent of X. The inter sample spacing density is denoted by $f(\cdot)$. The minimum separation between the successive samples, say d, implies that

$$f(u) = 0 \quad \text{for } 0 \leq u < d.$$

Without loss of generality, we assume that

$$f(u) > 0 \quad \text{for } u \geq d. \tag{1}$$

We first estimate the covariance function and subsequently using this we estimate the spectral density.

2.1 The estimator

We estimate $C(0)$ by

$$\widehat{C}(0) = \frac{1}{n} \sum_{i=1}^{n} X^2(t_i). \tag{2}$$

For $u > d$, we use the estimator

$$\widehat{C}(u) \quad = \quad \frac{1}{nH(u)} \sum_{i=1}^{n} \sum_{j=1}^{n} m_n W(m_n(u - t_i + t_j)) X(t_i) X(t_j) \quad \text{for } u > d, \tag{3}$$

where $W(\cdot)$ is a weight function, m_n is the smoothing parameter and $H(\cdot)$ is the renewal density of τ. The renewal density $H(\cdot)$ is defined as

$$H(u) = \sum_{n=1}^{\infty} f^{(n)}(u), \tag{4}$$

where $f^{(n)}(\cdot)$ is the n-fold convolution of the inter-sample spacing density $f(\cdot)$. Note that using (1), $H(u) > 0$ for $u > d$. The estimator $\widehat{C}(u)$, given by (3), is essentially a weighted average of the product of the samples of the separation at least by the amount u and the smoothing parameter m_n controls the amount of weight given if the separation between samples is larger than u. The idea is similar to the estimation of covariances at uniform lags in case of uniformly spaced samples. The factor $H(u)$ turns out to be right scaling for the estimator being asymptotically unbiased.

The main difficulty arising from the constraint lies in the lack of pairs of samples separated by lags in the range $(0, d]$. It has been shown in [7] that $C(0)$ and $C(u)$ for $u > d$ contains complete information about $C(\cdot)$ over its entire domain. A possible way of reconstructing the function over the range $(0, d]$ is to use the representation of $C(\cdot)$ in terms of its values over a grid. Note that the covariance function $C(\cdot)$ has the representation [10]

$$C(u) = \sum_{l=-\infty}^{\infty} C(lT)\text{sinc}\left(\frac{\pi}{T}(u - lT)\right), \tag{5}$$

where $T = \frac{\pi}{\lambda_0}$ and

$$\text{sinc}(x) = \begin{cases} \frac{\sin x}{x} & \text{if } x \neq 0, \\ 1 & \text{if } x = 0. \end{cases}$$

In view of (5) and the symmetry of $C(\cdot)$, one only needs to specify the sequence $\{C(lT), \, l = 0, 1, 2, \ldots\}$ in order to specify the function $C(\cdot)$ completely.

We have the estimates the values $C(lT)$ for $l = 0$ and $l > J$, where $J = [d/T]$, the integer part of d/T from (2) and (3). The remaining values, i.e., $C(T), \ldots, C(JT)$, can be expressed in terms of the left hand side of (5) and the known terms of the right hand side. Note that the left hand side of (5) can also be estimated directly for any $u > d$. Thus, the missing values satisfy the linear equations

$$\sum_{l=1}^{J} x_l(u)C(lT) = y(u) \quad \text{for } u > d, \tag{6}$$

where

$$x_l(u) = \left\{\text{sinc}\left(\frac{\pi}{T}(u - lT)\right) + \text{sinc}\left(\frac{\pi}{T}(u + lT)\right)\right\}, \text{ for } l = 1, 2, \ldots, J; \tag{7}$$

$$\begin{aligned} y(u) &= C(u) - \text{sinc}(\pi u/T)C(0) \\ &\quad - \sum_{l=J+1}^{\infty} \left\{\text{sinc}\left(\frac{\pi}{T}(u-lT)\right) + \text{sinc}\left(\frac{\pi}{T}(u+lT)\right)\right\} C(lT), \end{aligned} \tag{8}$$

for $u > d$. One can use these equations to reconstruct $C(T), \ldots, C(JT)$.

For indirect estimation of $C(T), \ldots, C(JT)$ from the linear equations (6), we define for $u > d$

$$\begin{aligned} y_n(u) &= \widehat{C}(u) - \text{sinc}(\pi u/T)\widehat{C}(0) \\ &\quad - \sum_{l=J+1}^{L_n} \left\{ \text{sinc}\left(\frac{\pi}{T}(u - lT)\right) + \text{sinc}\left(\frac{\pi}{T}(u + lT)\right) \right\} \widehat{C}(lT), \end{aligned} \tag{9}$$

where L_n is a finite integer. Note that $y_n(u)$ is an estimator of $y(u)$ defined in (8), with the infinite sum truncated at L_n. Substitution of this estimator on the right hand side of (6) gives a set of approximately linear equations in $C(T), \ldots, C(JT)$. This 'functional data' linear model [11] leads to the least squares estimator

$$\begin{pmatrix} \widehat{C}(T) \\ \widehat{C}(2T) \\ \vdots \\ \widehat{C}(JT) \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1J} \\ a_{21} & a_{22} & \cdots & a_{2J} \\ \vdots & \vdots & \ddots & \vdots \\ a_{J1} & a_{J2} & \cdots & a_{JJ} \end{pmatrix}^{-1} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_J \end{pmatrix}, \tag{10}$$

where $a_{jl} = \int_{u_1}^{u_2} x_j(u)x_l(u)du$, $b_j = \int_{u_1}^{u_2} x_j(u)y_n(u)du$, for $l, j = 1, \ldots, J$ and (u_1, u_2) is a suitable sub-interval of $[d, \infty)$. The numbers u_1 and u_2 are chosen such that $f(u)$ is less than a specified small number for $u < u_1$ and $u > u_2$, so that the $\widehat{C}(u)$ estimators *not* utilized in the linear model are those based on relatively few pairs of consecutive samples with unusually small or large separation. The description of the covariance estimator is completed by defining it in the range $0 < u \leq d$, $u \neq T, \ldots, JT$, as

$$\widehat{C}(u) = \widehat{C}(0)\text{sinc}\left(\frac{\pi u}{T}\right) + \sum_{l=1}^{L_n} \widehat{C}(lT) \left\{ \text{sinc}\left(\frac{\pi}{T}(u - lT)\right) + \text{sinc}\left(\frac{\pi}{T}(u + lT)\right) \right\}$$

Once the function $C(\cdot)$ is completely estimated, we estimate the power spectral density by a commonly used nonparametric estimator namely lag window estimator

$$\widehat{\phi}_n(\lambda) = \left\{ \frac{T}{2\pi}\widehat{C}(0) + \frac{T}{\pi} \sum_{l=1}^{n-1} \widehat{C}(lT)K(b_n l)\cos(l\lambda T) \right\} \times 1_{[-\lambda_0, \lambda_0]}(\lambda), \tag{11}$$

where $K(\cdot)$ is a covariance averaging Kernel and b_n is the kernel bandwidth.

2.2 Consistency

In order to establish the consistency of the proposed spectrum estimator, we will show that the bias and the variance of the corresponding estimator of the covariance function go to zero, as the sample size goes to infinity. For this purpose, we choose the weight function $W(\cdot)$ used in (3) has the following property.

1. The function $W(\cdot)$ is compactly supported, even, continuous, integrable and square integrable, with $\int_{-\infty}^{\infty} W(v)dv = 1$.

2. For a specified r,

$$\int_{-\infty}^{\infty} v^k W(v)dv = \begin{cases} 0 & \text{for} k = 1, 2, \ldots, r-1 \\ w_r \neq 0 & \text{for} k = r. \end{cases}$$

   The number r is termed as order of the weight function.

We make some assumptions about underlying process, inter sample spacing density for consistency and smoothing and truncation parameters.

Assumption 1. Covariance function $C(\cdot)$ is bounded by a decreasing and integrable function over $[0, \infty)$.
Assumption 1A. The function $|u|^q C(u)$, for some $q > 1$, is bounded by a decreasing and integrable function over $[0, \infty)$.

The parameter q of Assumption 1A signifies the degree of smoothness of $\phi(\cdot)$. In particular, if q is an integer, the assumption implies that $\phi(\cdot)$ is q times differentiable. Note that Assumption 1A implies Assumption 1.

Assumption 2. The inter-sample spacings density $f(\cdot)$ has a finite mean.

Assumption 3. The smoothing parameter m_n used in (3) is such that $m_n \to \infty$ as $n \to \infty$.

Assumption 4. The truncation parameter L_n used in (9) is such that $L_n \to \infty$ as $n \to \infty$.

Theorem 1 Let Assumptions 1–4 hold.

(i) The bias of the estimator $\widehat{C}(\cdot)$ converges to 0 pointwise as the sample size n goes to infinity.

(ii) Under further Assumptions 1A, $\frac{m_n \log L_n}{n} \to 0$ and $\frac{\log L_n}{m_n^r} \to 0$ as $n \to \infty$, we have

$$E[\widehat{C}(u)] - C(u) = \begin{cases} 0 & \text{if } u = 0, \\ O\left(\frac{1}{m_n^r}\right) + O\left(\frac{m_n}{n}\right) & \text{if } |u| > d, \\ O\left(\frac{1}{L_n^q}\right) + O\left(\frac{\log L_n}{m_n^r}\right) + O\left(\frac{m_n \log L_n}{n}\right) & \text{if } 0 < |u| \leq d, \end{cases}$$

where r is the order of weight function and $O(\cdot)$ is uniform in u.

Note that $L_n \to \infty$ together with $\frac{m_n \log L_n}{n} \to 0$ imply that $\frac{m_n}{n} \to 0$ as $n \to \infty$. A further assumption on fourth order moments of the process is needed to establish convergence for the variance of the estimator $\widehat{C}(u)$.

Assumption 5. The fourth moment $E|X(t)|^4$ exists for every t, and the fourth order cumulant of $X(t+t_1)$, $X(t+t_2)$, $X(t+t_3)$ and $X(t)$ does not depend on t, and this function, denoted by $Q(t_1, t_2, t_3)$, satisfies

$$|Q(t_1, t_2, t_3)| \leq \prod_{j=1}^{3} g_j(t_j),$$

where $g_j(\cdot)$, $j = 1, 2, 3$, are all continuous, even, nonnegative and integrable functions over the real line, which are non-increasing over $[0, \infty)$.

Assumption 5 holds trivially for a Gaussian process since $Q(\cdot) = 0$.

Theorem 2. Under Assumptions 1–5 and $\frac{m_n (\log L_n)^2}{n} \to 0$ as $n \to \infty$, the variance of the estimator $\widehat{C}(\cdot)$ converges uniformly to zero as follows:

$$Var(\widehat{C}(u)) = \begin{cases} O\left(\frac{1}{n}\right) & \text{if } u = 0 \\ O\left(\frac{m_n}{n}\right) & \text{if } |u| > d \\ O\left(\frac{m_n (\log L_n)^2}{n}\right) & \text{if } 0 < |u| \leq d. \end{cases}$$

Note that $L_n \to \infty$ together with $\frac{m_n (\log L_n)^2}{n} \to 0$ imply that $\frac{m_n}{n} \to 0$ as $n \to \infty$.

Under Assumptions of Theorem 1(ii) and 2, we have the rate of convergence of the mean square error (MSE) of the estimator $\widehat{C}(u)$, for $0 < |u| \leq d$, as follows:

$$\begin{aligned} E[\widehat{C}(u) - C(u)]^2 &= (E[\widehat{C}(u)] - C(u))^2 + Var(\widehat{C}(u)) \\ &= O\left(\frac{1}{L_n^{2q}}\right) + O\left(\frac{(\log L_n)^2}{m_n^{2r}}\right) + O\left(\frac{m_n^2(\log L_n)^2}{n^2}\right) \\ &\quad + O\left(\frac{m_n(\log L_n)^2}{n}\right). \end{aligned} \tag{12}$$

In view of conditions of Theorem 2, the third term in (12) can be ignored. Note that, for fixed L_n, the second term is a decreasing function of m_n while the fourth term is increasing in m_n. The fastest possible rate of convergence is achieved by equating these two rates, which yields $m_n = O\left(n^{\frac{1}{2r+1}}\right)$. Likewise, the first term is a decreasing function of L_n, while the fourth term is increasing in L_n. Again by equating the rates, we get the optimal rate of L_n satisfying the condition

$$(L_n^q \log L_n)^{\frac{2r+1}{r}} = O(n)$$

for fastest convergence of the mean squared error. Even though a closed form expression for the optimal rate of L_n is not available, use of the square root function in lieu of the log function leads to the nearly optimal rate $L_n = O(n^{\frac{2r}{(2r+1)(2q+1)}})$. Substitution of these rates of m_n and L_n in (12) shows that the optimal rate of convergence of MSE is bounded from above as follows.

$$E[\widehat{C}(u) - C(u)]^2 = o\left(n^{-\frac{1}{\left(1+\frac{1}{2r}\right)\left(1+\frac{1}{2q}\right)}}\right). \tag{13}$$

On the other hand, for $|u| > d$, the MSE of the direct estimate $\widehat{C}(u)$ is

$$E[\widehat{C}(u) - C(u)]^2 = O\left(\frac{1}{m_n^{2r}}\right) + O\left(\frac{m_n}{n}\right).$$

By using a similar argument as above it can be seen that the optimal rate of convergence of MSE for $|u| > d$ is

$$E[\widehat{C}(u) - C(u)]^2 = O\left(n^{-\frac{1}{1+\frac{1}{2r}}}\right). \tag{14}$$

We now turn to the consistency of the spectrum estimator $\widehat{\phi}_n(\lambda)$ given in (11). For this purpose, we choose the covariance averaging kernel $K(\cdot)$ used in (11) having following properties.

1. The function $K(\cdot)$ is an even, continuous, integrable and square integrable function with $K(0) = 1$, and is bounded by a nondecreasing function over $(0, \infty)$.

2. The order of the kernel $K(\cdot)$ is q, where q is as in Assumption 1A.

Assumption 6. The kernel bandwidth b_n used in (11) is such that $b_n \to 0$ as $n \to \infty$.

Theorem 3. Let Assumptions 1–4 and 6 hold.

(i) The bias of the estimator $\widehat{\phi}_n(\cdot)$ converges to 0 pointwise as the sample size n goes to infinity.

(ii) Under the conditions of Theorem 1(ii) and $\frac{1}{b_n m_n^r} \to 0$, we have

$$\begin{aligned} E[\widehat{\phi}_n(\lambda)] - \phi(\lambda) &= O\left(b_n^q\right) + O\left(\frac{1}{L_n^q}\right) + O\left(\frac{\log L_n}{m_n^r}\right) + O\left(\frac{m_n \log L_n}{n}\right) \\ &\quad + O\left(\frac{1}{b_n m_n^r}\right), \end{aligned}$$

where r is the order of the weight function $W(\cdot)$.

Theorem 4. Under conditions of Theorem 2 and Assumption 6 and $\frac{m_n}{nb_n^2} \to 0$ as $n \to \infty$, the variance of the estimator $\widehat{\phi}_n(\cdot)$ converges as follows:

$$Var[\widehat{\phi}_n(\lambda)] = O\left(\frac{m_n (\log L_n)^2}{n}\right) + O\left(\frac{m_n}{nb_n^2}\right).$$

Under conditions of Theorem 3 and 4, we have the rate of convergence of the MSE of the estimator $\widehat{\phi}_n(\cdot)$ as follows:

$$\begin{aligned} E[\widehat{\phi}_n(\lambda) - \phi(\lambda)]^2 &= O\left(b_n^{2q}\right) + O\left(\frac{1}{L_n^{2q}}\right) + O\left(\frac{(\log L_n)^2}{m_n^{2r}}\right) + O\left(\frac{m_n^2 (\log L_n)^2}{n^2}\right) \\ &\quad + O\left(\frac{1}{b_n^2 m_n^{2r}}\right) + O\left(\frac{m_n (\log L_n)^2}{n}\right) + O\left(\frac{m_n}{nb_n^2}\right). \end{aligned} \tag{15}$$

In view of conditions of Theorem 2, the fourth term in (15) is smaller than the sixth term and hence can be ignored. Note that for fixed b_n, the fifth term is a decreasing function of m_n, while the seventh term is increasing in m_n. The fastest rate of convergence is obtained by equating the two rates, which yields $m_n = O\left(n^{\frac{1}{1+2r}}\right)$. Now, for this optimal choice of m_n, the fifth term is a decreasing function of b_n, while the first term is increasing in b_n. Thus, by equating the two rates, we obtain the optimal rate of b_n as $O\left(n^{-\frac{r}{(1+2r)(1+q)}}\right)$. Substitution of these choices reduces (15) to

$$E[\widehat{\phi}_n(\lambda) - \phi(\lambda)] \quad = \quad O\left(n^{-\frac{1}{\left(1+\frac{1}{2r}\right)\left(1+\frac{1}{q}\right)}}\right) + O\left(\frac{1}{L_n^{2q}}\right) + O\left(\frac{(\log L_n)^2}{n^{\frac{2r}{1+2r}}}\right). \quad (16)$$

The requirement that the second and the third terms of the right hand side go to zero at least as fast as the first term, leads us to a range of optimal choice of L_n:

$$\log L_n = O\left(n^{\frac{1}{(1+\frac{1}{2r})(2+2q)}}\right); \qquad L_n^{-1} = O\left(n^{-\frac{1}{(1+\frac{1}{2r})(2+2q)}}\right).$$

For any choice of L_n in this optimal range, and for $m_n = O\left(n^{\frac{1}{1+2r}}\right)$ and $b_n = O\left(n^{-\frac{r}{(1+2r)(1+q)}}\right)$, the MSE of $\widehat{\phi}_n(\lambda)$ achieves the fastest rate of convergence, $O\left(n^{-\frac{1}{(1+\frac{1}{2r})(1+\frac{1}{q})}}\right)$.

We conclude this section by looking into the special case when $d \leq \frac{\pi}{\lambda_0}$, i.e., no indirect estimation of the covariance sequence is needed. In this situation, the parameter L_n is not required and the optimal rate of convergence of the MSE of $\widehat{\phi}_n(\cdot)$ is once again $O\left(n^{-\frac{1}{(1+\frac{1}{2r})(1+\frac{1}{q})}}\right)$. Observe that order r of the weight function $W(\cdot)$ can be chosen to be arbitrarily large by using appropriate weight function, and hence the optimal rate of convergence can approach $O\left(n^{-\frac{1}{1+\frac{1}{q}}}\right)$. Note that when $d \leq \frac{\pi}{\lambda_0}$, uniform sampling under the given constraint becomes alias-free, and the optimal rate of convergence of the MSE of the smoothed periodogram estimator based on uniform sampling is $O\left(n^{-\frac{1}{1+\frac{1}{2q}}}\right)$ [12]. The optimal rate of convergence of the MSE of $\widehat{\phi}_n(\cdot)$ is only marginally slower than this rate when q is large, i.e, when the underlying power spectral density is very smooth.

3 Simulation

In this section, we study the performance of the estimator $\widehat{\phi}_n(\cdot)$, and compare this performance with that of the estimator given by (**??**), which we denote here by $\widehat{\psi}_n(\cdot)$.

We consider a continuous time stationary stochastic process X with mean 0 and covariance function $C(\cdot)$ given by

$$C(u) = \begin{cases} \sigma^2 \displaystyle\sum_{j=0}^{q-|l|} \xi_j \xi_{j+|l|} & \text{if } u = \frac{l\pi}{\lambda_0}, |l| = 0, 1, \ldots, q, \\ 0 & \text{if } u = \frac{l\pi}{\lambda_0}, |n| > q, q+1, \ldots, \\ \displaystyle\sum_{l=-\infty}^{\infty} C\left(\frac{l\pi}{\lambda_0}\right) \frac{\sin(\lambda_0 u - l\pi)}{(\lambda_0 u - l\pi)} & \\ & \text{otherwise.} \end{cases}$$

This covariance function corresponds to a process limited to the frequency band $[-\lambda_0, \lambda_0]$, whose samples at regular intervals of length $\frac{\pi}{\lambda_0}$ constitute a discrete time MA(q) process with MA characteristic polynomial $\Xi(z) = \xi_0 + \xi_1 z + \xi_2 z^2 + \cdots + \xi_q z^q$ and innovation variance σ^2.

We use sampling with a stationary renewal process τ whose inter-sample spacing is distributed as $d + T$, where the random variable T has the exponential distribution with mean θ. Note that if one attempts to implement Poisson sampling by generating successive inter-sample spacings from the exponential distribution with mean θ, but is obliged to discard those inter-sample spacings which are smaller than d, then the modified sampling scheme resulting from this 'imperfect' Poisson sampling is precisely the one described as τ. We assume that n consecutive samples, denoted by $X(t_1), X(t_2), \ldots, X(t_n)$, are available for estimation.

We study the performance of the estimators $\widehat{\phi}_n(\cdot)$ and $\widehat{\psi}_n(\cdot)$ for the choices

$$\begin{aligned}
\lambda_0 &= 2\pi, \\
\Xi(z) &= (1+1.2z)^8, \\
\sigma &= 1/10^2, \\
\theta &= 1, \\
W(x) &= \begin{cases} \frac{1}{2}\{1+\cos(\pi x)\} & \text{if } -1 \le x \le 1, \\ 0 & \text{otherwise,} \end{cases} \\
\text{and } K(x) &= \begin{cases} \frac{1}{2}\{1+\cos(\pi x)\} & \text{if } -1 \le x \le 1, \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}$$

In accordance with the strategy mentioned in Section 2.1, we choose $u_1 = d$ and $u_2 = d - \theta\log(0.01\theta)$. We first assume that the minimum inter-sample spacing is $d = 0.75$. This spacing is larger than the uniform inter-sample spacing needed for sampling at the Nyquist rate (which happens to be $T = 0.5$). We run simulations for sample sizes $n = 100$ and $n = 1000$. For both $\widehat{\phi}_n(\cdot)$ and $\widehat{\psi}_n(\cdot)$ and both sample sizes, we use $b_n = 0.1$. Also, for $\widehat{\phi}_n(\cdot)$ and both sample sizes, we use $L_n = 20$. Finally, we use $m_n = 5$ for $n = 100$ and $m_n = 8$ for $n = 1000$.

Figure 1 shows the average of the estimates of $\widehat{\phi}_n(\lambda)$ and $\widehat{\psi}_n(\lambda)$ computed from 500 Monte Carlo simulation runs, along with the true density. In order to emphasize the inconsistency of $\widehat{\psi}_n(\cdot)$ in the present situation, we also include in this figure the plot of the probability limit of this estimator, which is

$$\lim_{n\to\infty} E\left[\widehat{\psi}_n(\lambda)\right] = \frac{1}{\pi}\int_d^\infty C(u)\cos(\lambda u)du.$$

The plots suggest that the $\widehat{\psi}_n(\cdot)$ converges to this wrong function, which is not even positive over its entire range. The empirical bias of the estimator $\widehat{\phi}_n(\lambda)$ becomes smaller for larger n. Figure 2, which shows the empirical mean squared errors (MSE) of the two estimators, indicates that the mean squared error of $\widehat{\phi}_n(\lambda)$ is smaller, and it reduces with sample size, while that of $\widehat{\psi}_n(\lambda)$ saturates to a non-zero level, because of the bias component.

We now turn to the more challenging case of $d = 1$. The simulations are run with the parameters chosen as described in the previous case. Figures 3 and 4 depict the empirical average and the mean squared error, respectively, for the two estimators, computed from 500 simulation runs. Figure 3 indicates convergence

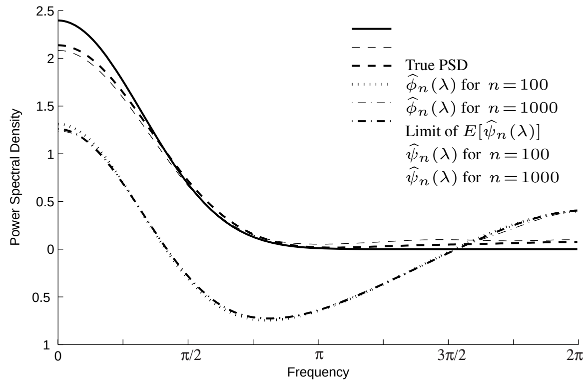

Figure 1: Average of the estimates of $\widehat{\phi}_n(\lambda)$ and $\widehat{\psi}_n(\lambda)$ for minimum separation $d = 0.75$ based on sample sizes 100 and 1000 computed from 500 Monte Carlo simulation runs, along with True PSD and $\lim_{n\to\infty} E[\widehat{\psi}_n(\lambda)]$.

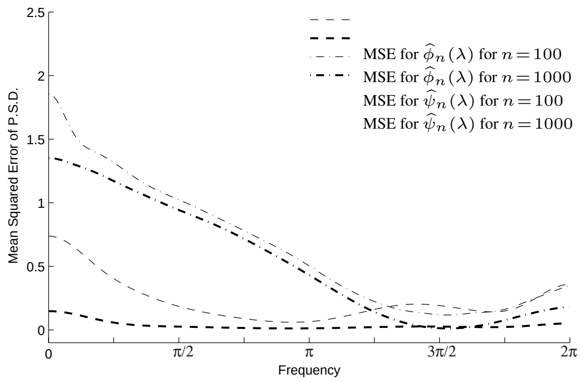

Figure 2: Empirical MSE of $\widehat{\phi}_n(\lambda)$ and $\widehat{\psi}_n(\lambda)$ for minimum separation $d = 0.75$ based on sample sizes 100 and 1000 computed from 500 Monte Carlo simulation runs.

of $\widehat{\psi}_n(\lambda)$ to its asymptotic mean, which is different from both $\phi(\lambda)$ and the asymptotic mean in the case $d = 0.75$ (see Figure 1). For $n = 100$, the estimator $\widehat{\phi}_n(\lambda)$

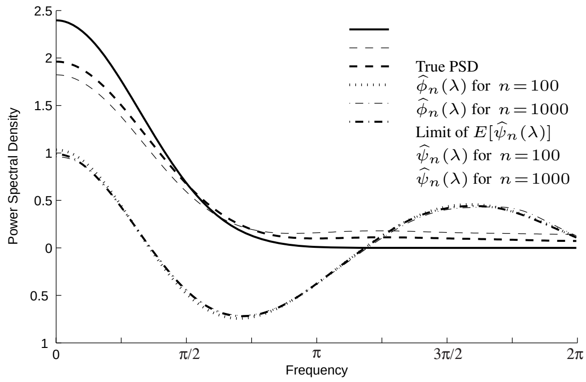

Figure 3: Average of the estimates of $\widehat{\phi}_n(\lambda)$ and $\widehat{\psi}_n(\lambda)$ for minimum separation $d = 1$ based on sample sizes 100 and 1000 computed from 500 Monte Carlo simulation runs, along with True PSD and $\lim_{n\to\infty} E[\widehat{\psi}_n(\lambda)]$.

has somewhat larger MSE than that of $\widehat{\psi}_n(\lambda)$ for some values of λ. However, for $n = 1000$, the MSE of $\widehat{\phi}_n(\lambda)$ reduces sharply. On the other hand, the MSE of $\widehat{\psi}_n(\lambda)$ saturates to a certain positive value, determined by the asymptotic bias.

The MSE of $\widehat{\phi}_n(\lambda)$ is seen to be larger in the case $d = 1$ than in the case $d = 0.75$. This finding can be explained by the fact that in the former case, the estimator $\widehat{\phi}_n(\cdot)$ would involve indirect estimation of *two* covariance parameters, $C(T)$ and $C(2T)$, as opposed to indirect estimation of $C(T)$ only in the case of $d = 0.75$. Reduction of the set of lags suitable for direct estimation is another reason why the estimator has poorer performance for larger values of d. In any case, such difficulties are made up by large sample size, as is evident from the MSE of $\widehat{\phi}_n(\lambda)$ for $n = 1000$.

We finally consider the case $d = 0$, in order to find out whether the proposed method performs reasonably well even when Poisson sampling is feasible and uniform sampling is alias-free. In this case, $\widehat{\psi}_n(\cdot)$ is consistent. We compare the performances of $\widehat{\phi}_n(\lambda)$ with $\widehat{\psi}_n(\lambda)$ as well as the smoothed periodogram estima-

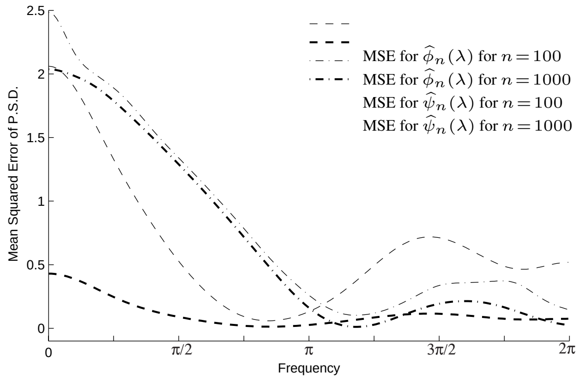

Figure 4: Empirical MSE of $\widehat{\phi}_n(\lambda)$ and $\widehat{\psi}_n(\lambda)$ for minimum separation $d = 1$ based on sample sizes 100 and 1000 computed from 500 Monte Carlo simulation runs.

tor, based on uniform sampling at Nyquist rate, defined as follows:

$$\begin{aligned}\widehat{\xi}_n(\lambda) \;&=\; \frac{T}{2\pi}\sum_{|l|<n}\left\{\frac{1}{n}\sum_{j=1}^{n-|l|} X(jT)X((j+|l|)T)\right\}K(b_n l)\\ &\quad\times e^{-il\lambda T}1_{[-\lambda_0,\lambda_0]}(\lambda),\end{aligned}$$

where $K(\cdot)$ is a covariance averaging Kernel and b_n is the kernel bandwidth. In the present case, all the three estimators are consistent, and indirect estimation of covariance is not necessary for $\widehat{\phi}_n(\lambda)$. For all the three estimators, the simulations are run with the same parameters chosen for the previous cases. The empirical averages of the three estimators based on 500 simulation runs, shown in Figure 5, indicate that bias of $\widehat{\phi}_n(\lambda)$ is in between those of $\widehat{\xi}_n(\lambda)$ and $\widehat{\psi}_n(\lambda)$, for sample sizes 100 and 1000. Its performance in terms of MSE, plotted in Figure 6, is much better than that of $\widehat{\psi}_n(\lambda)$, and marginally worse than that of $\widehat{\xi}_n(\lambda)$, for both sample sizes. These results indicate that $\widehat{\phi}_n(\lambda)$ does not perform poorly even in the special case when competing consistent estimators are available. The smaller MSE of $\widehat{\phi}_n(\lambda)$ in comparison to that of $\widehat{\psi}_n(\lambda)$, in spite of $\widehat{\phi}_n(\lambda)$ having a slower rate of convergence (see Section 2.2), may be attributed to the difference of the constants associated with the two rates.

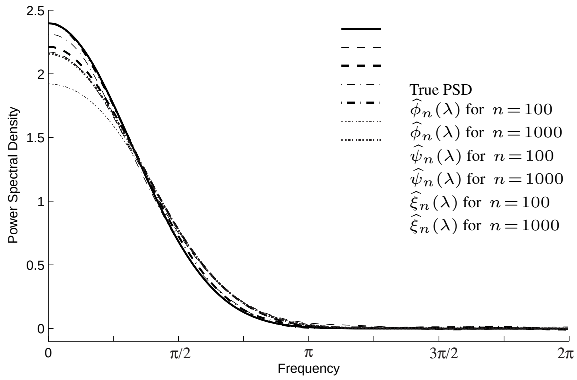

Figure 5: Average of the estimates of $\widehat{\phi}_n(\lambda)$, $\widehat{\psi}_n(\lambda)$ and $\widehat{\xi}_n(\lambda)$ for $d = 0$ based on sample sizes 100 and 1000 computed from 500 Monte Carlo simulation runs, along with True PSD.

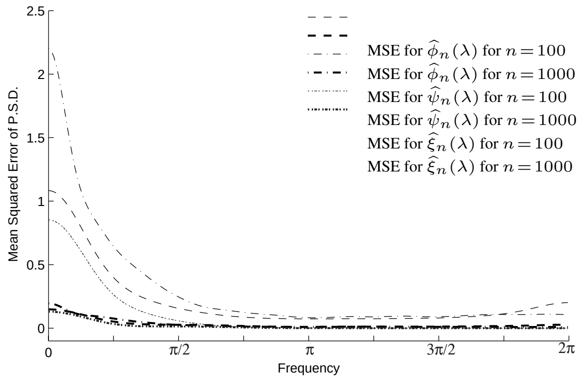

Figure 6: Empirical MSE of $\widehat{\phi}_n(\lambda)$, $\widehat{\psi}_n(\lambda)$ and $\widehat{\xi}_n(\lambda)$ for $d = 0$ based on sample sizes 100 and 1000 computed from 500 Monte Carlo simulation runs.

4 Concluding remarks

This paper provides a method of consistent estimation of an arbitrarily bandlimited power spectral density of a continuous time stationary stochastic process, under the constraint that there has to be at least a specified amount of separation between successive samples. The proposed nonparametric estimator is based on additive random sampling of the underlying process subject to this practical constraint. The estimator is the first of its kind, as its known competitors based on stochastic sampling are not consistent, and the known competitors based on uniform sampling are consistent only when the bandwidth of the underlying process is within the limit implied by the Nyquist theorem. The proposed estimator demonstrates that it is possible to judiciously use additive random sampling to surpass the Nyquist limit with the help of large sample size.

It is important to observe that the constraint of a minimum separation between successive samples makes it impossible to estimate autocovariances at small lags directly from the data. The proposed method circumvents this difficulty by expressing these autocovariances in terms of directly estimable autocovariances through the representation (5). This indirect method of estimation leads to larger variance than in the case of direct estimation. The greater the minimum separation, the larger becomes the need for indirect estimation and the larger is the variance of the resulting spectrum estimator. The simulation results reported in Section 3 confirm this fact. Thus, while it is possible to make up for the deficiency of sampling resolution through sample size, the requirement of sample size becomes large when the resolution is not good.

Some researchers have promoted the use of point process sampling, by arguing that non-bandlimited power spectral densities cannot be consistently estimated from uniformly spaced samples, no matter how large the sampling rate is [5, 8]. Srivastava and Sengupta [6] have shown that this apparent defect of uniform sampling can be rectified by increasing the sampling rate judiciously for larger sample sizes. However, their asymptotic arguments become inappropriate when inter-sample spacings are constrained to be larger than a certain threshold. On the other hand, a result of Srivastava and Sengupta [7] shows that, under this constraint, consistent estimation of a non-bandlimited power spectral density through point process sampling is also not possible. Thus, as far as estimation of a *non-bandlimited* power spectral density is concerned, uniform sampling and point process sampling have similar strengths and limitations. The spectrum estimator introduced in this paper underscores an exclusive advantage of point process sampling in the area of estimation of a *bandlimited* power spectral density.

Appendix

We denote the domain of the function $W(\cdot)$ by $[-a, a]$, the suprema of the functions $|W|(\cdot)$ and $H(\cdot)$ by M_1 and M_2, respectively, and the infimum of $H(\cdot)$ by M_3 and the function $g(u)$ is a bounding function as in Assumption 1.

Proof of Theorem 1. Part (i). Observe that

$$E[\widehat{C}(0)] = \frac{1}{n}\sum_{i=1}^{n} E[X^2(t_i)] = C(0), \tag{17}$$

i.e., $\widehat{C}(0)$ is an unbiased estimator of $C(0)$. From (3), we have for $|u| > d$

$$\begin{aligned} &E[\widehat{C}(u)] \\ &= \frac{1}{nH(u)}\sum_{i=1}^{n}\sum_{j=1}^{n} m_n E[E\{W(m_n(u - t_i + t_j))X(t_i)X(t_j)|t_i,\ i = 1, \ldots, n\}] \\ &= \frac{1}{nH(u)}\sum_{i=1}^{n}\sum_{j=1}^{n} m_n E[W(m_n(u - t_i + t_j))C(t_i - t_j)] \end{aligned}$$

After considering the case $i = j$ separately, and combining the cases $i < j$ and $i > j$, we have

$$\begin{aligned} &E[\widehat{C}(u)] \\ &= \frac{m_n W(m_n u)C(0)}{H(u)} + \frac{1}{H(u)}\int_0^\infty m_n\Bigg[\{W(m_n(u+v)) + W(m_n(u-v))\} \\ &\qquad\qquad \times C(v)\Bigg\{\frac{1}{n}\sum_{1\le i<j\le n} f^{(j-i)}(v)\Bigg\}\Bigg]dv \\ &= \frac{m_n W(m_n u)C(0)}{H(u)} + \frac{1}{H(u)}\int_0^\infty m_n\Bigg[\{W(m_n(u+v)) + W(m_n(u-v))\} \\ &\qquad\qquad \times C(v)H_n(v)\Bigg]dv, \end{aligned} \tag{18}$$

where

$$H_n(u) = \sum_{l=1}^{n-1}\left(1 - \frac{l}{n}\right) f^{(l)}(u). \tag{19}$$

By making a transformation of the variable of integration and using the symmetry of the covariance function $C(\cdot)$, we have

$$\begin{aligned} E[\widehat{C}(u)] &= \frac{m_n W(m_n u) C(0)}{H(u)} \\ &\quad + \frac{1}{H(u)} \int_{-\infty}^{\infty} W(v) C\left(u - \frac{v}{m_n}\right) \left[H_n\left(u - \frac{v}{m_n}\right) + H_n\left(-u + \frac{v}{m_n}\right)\right] dv. \quad (20) \end{aligned}$$

For sufficiently large n, we have $m_n > a/d$, and consequently $W(m_n u) = 0$ for all $|u| > d$. This implies that the first term is identically zero for large n. Thus, we have, for large n,

$$E[\widehat{C}(u)] = \frac{1}{H(u)} \int_{-\infty}^{\infty} W(v) C\left(u - \frac{v}{m_n}\right) \left[H_n\left(u - \frac{v}{m_n}\right) + H_n\left(-u + \frac{v}{m_n}\right)\right] dv \quad (21)$$

Further, by using Assumptions 1 and 2, we have the dominance

$$\begin{aligned} &\left| W(v) C\left(u - \frac{v}{m_n}\right) \left[H_n\left(u - \frac{v}{m_n}\right) + H_n\left(-u + \frac{v}{m_n}\right)\right] \right| \\ &\quad \leq |W(v)| g(0) \left[H_n\left(u - \frac{v}{m_n}\right) + H_n\left(-u + \frac{v}{m_n}\right)\right] \leq |W(v)| g(0) 2M_2, \end{aligned}$$

and Property 1 of the weight function $W(\cdot)$ ensures that the bounding function is integrable. From Assumption 3, we have the convergence

$$\lim_{n\to\infty} W(v) C\left(u - \frac{v}{m_n}\right) \left[H_n\left(u - \frac{v}{m_n}\right) + H_n\left(-u + \frac{v}{m_n}\right)\right] = W(v) C(u) H(u).$$

By applying the Dominated Convergence Theorem (DCT), we have

$$\lim_{n\to\infty} E[\widehat{C}(u)] = C(u). \quad (22)$$

In order to compute the expectation of the indirect estimator $(\widehat{C}(T), \ldots, \widehat{C}(JT))$ given in (10), we first compute $E\left[\int_{u_1}^{u_2} x_j(u) y_n(u) du\right]$. Note that for $j = 1, 2, \ldots, J$, we have

$$E\left[\int_{u_1}^{u_2} x_j(u) y_n(u) du\right] = \int_{u_1}^{u_2} x_j(u) E[y_n(u)] du,$$

where interchange of the integrals is justified by the finiteness of the double integral, which follows from arguments similar to those given below to establish the convergence. We compute

$$E[y_n(u)] = E[\widehat{C}(u)] - \text{sinc}(\pi u/T)E[\widehat{C}(0)] - \sum_{l=J+1}^{L_n} E[\hat{C}(lT)] \left\{ \text{sinc}\Big(\frac{\pi}{T}(u-lT)\Big) + \text{sinc}\Big(\frac{\pi}{T}(u+lT)\Big) \right\} . \quad (23)$$

It is seen from (17) and (22) that $E[\widehat{C}(u)] \to C(u)$ as $n \to \infty$ for $u > d$ and $E[\widehat{C}(0)] = C(0)$. By using (21) for large n, the third term of right hand side of (23) simplifies to

$$\int_{-\infty}^{\infty} W(v) \Bigg[\sum_{l=J+1}^{L_n} \frac{1}{H(lT)} C\left(lT - \frac{v}{m_n}\right) \left\{ H_n\Big(lT - \frac{v}{m_n}\Big) + H_n\Big(-lT + \frac{v}{m_n}\Big) \right\} \times \left\{ \text{sinc}\Big(\frac{\pi}{T}(u-lT)\Big) + \text{sinc}\Big(\frac{\pi}{T}(u+lT)\Big) \right\} \Bigg] dv. \quad (24)$$

By using Assumptions 1 and 2, we have the dominance

$$\begin{aligned} &\left| W(v) \sum_{l=J+1}^{L_n} \frac{1}{H(lT)} C\left(lT - \frac{v}{m_n}\right) \left\{ H_n\left(lT - \frac{v}{m_n}\right) + H_n\left(-lT + \frac{v}{m_n}\right) \right\} \times \left\{ \text{sinc}\Big(\frac{\pi}{T}(u-lT)\Big) + \text{sinc}\Big(\frac{\pi}{T}(u+lT)\Big) \right\} \right| \\ \le\ & |W(v)| \frac{2M_2}{M_3} \sum_{l=J+1}^{L_n} g\left(lT - \frac{v}{m_n}\right) \le |W(v)| \frac{2M_2}{M_3} \times 2 \sum_{l=1}^{\infty} g(lT), \end{aligned}$$

and the integrability of the bound is guaranteed by Property 1 of the weight function. From Assumptions 3 and 4, the integrand of (24) converges pointwise as

$$\begin{aligned} &\lim_{n\to\infty} W(v) \sum_{l=J+1}^{L_n} E[\widehat{C}(lT)] \left\{ \text{sinc}\Big(\frac{\pi}{T}(u-lT)\Big) + \text{sinc}\Big(\frac{\pi}{T}(u+lT)\Big) \right\} \\ =\ & W(v) \sum_{l=J+1}^{\infty} C(lT) \left\{ \text{sinc}\Big(\frac{\pi}{T}(u-lT)\Big) + \text{sinc}\Big(\frac{\pi}{T}(u+lT)\Big) \right\} \end{aligned}$$

Thus, by using the representation (5) and the DCT, we have

$$\begin{aligned}\lim_{n\to\infty} E[y_n(u)] &= C(u) - \text{sinc}(\pi u/T)C(0)\\ &\quad - \sum_{l=J+1}^{\infty} C(lT)\left\{\text{sinc}\Big(\frac{\pi}{T}(u-lT)\Big) + \text{sinc}\Big(\frac{\pi}{T}(u+lT)\Big)\right\}\\ &= y(u).\end{aligned}$$

Thus, for $j = 1, \ldots, J$, using (6) we have

$$\lim_{n\to\infty} E\left[\int_{u_1}^{u_2} x_j(u) y_n(u) du\right] = \int_{u_1}^{u_2} x_j(u)\left\{\sum_{l=1}^{J} x_l(u) C(lT)\right\} du. \quad (25)$$

By using (25) and (10), for $l = 1, \ldots, J$, we have

$$\lim_{n\to\infty} E[\widehat{C}(lT)] = C(lT). \quad (26)$$

By using a similar argument as above, for the range $0 < u \leq d$ but $u \neq T, \ldots, JT$, one can establish that

$$\lim_{n\to\infty} E[\widehat{C}(u)] = C(u).$$

This completes the proof of Part (i).

Part (ii). For sufficiently large n such that $m_n > a/d$, using Assumption 2, (18) and (19), we have

$$\begin{aligned}&E[\widehat{C}(u)]\\ &\quad= \int_0^{\infty} m_n\big[\{W(m_n(u+v)) + W(m_n(u-v))\}C(v)\big] dv\\ &\qquad - \frac{1}{H(u)}\int_0^{\infty} m_n\big[\{W(m_n(u+v)) + W(m_n(u-v))\}C(v)\sum_{l=n}^{\infty} f^{(l)}(v)\big] dv\\ &\qquad - \frac{1}{nH(u)}\int_0^{\infty} m_n\big[\{W(m_n(u+v)) + W(m_n(u-v))\}C(v)\sum_{l=1}^{n-1} l f^{(l)}(v)\big] dv.\end{aligned} \quad (27)$$

Observe that

$$\begin{aligned}
\sum_{l=1}^{\infty} l f^{(l)}(u) &= \sum_{l=1}^{\infty} l \frac{d}{du} F^{(l)}(u) = \frac{d}{du} \sum_{l=1}^{\infty} l P(N(u) \geq l) \\
&= \frac{d}{du} \sum_{l=1}^{\infty} l \sum_{j=l}^{\infty} P(N(u) = j) = \frac{d}{du} \sum_{j=1}^{\infty} P(N(u) = j) \sum_{l=1}^{j} l \\
&= \frac{1}{2} \frac{d}{du} \sum_{j=1}^{\infty} j(j-1) P(N(u) = j)
\end{aligned}$$

where $F^{(l)}(\cdot)$ is the distribution function of $f^{(l)}(\cdot)$, $N(\cdot)$ is the counting process associated with the sampling process. By using the fact that [14]

$$\sum_{j=1}^{\infty} j(j-1) P(N(u) = j) = \int_0^u \left\{ 1 + 2 \int_0^y H(v) dv \right\} dy - \int_0^u H(v) dv,$$

we have

$$\sum_{l=1}^{\infty} l f^{(l)}(u) = \frac{1}{2} + \int_0^u H(v) dv - \frac{1}{2} H(u). \tag{28}$$

In view of (28), the sum of the second and third terms of the right hand side of (27) is absolutely bounded from above by

$$\begin{aligned}
& 2 \frac{M_1}{M_3} \frac{m_n}{n} \int_0^{\infty} |C(v)| \sum_{l=1}^{\infty} l f^{(l)}(v) dv \\
\leq\ & 2 \frac{M_1}{M_3} \frac{m_n}{n} \int_0^{\infty} |C(v)| \left[\frac{1}{2} + \int_0^v H(t) dt - \frac{1}{2} H(v) \right] dv \\
\leq\ & 6 \frac{M_1 M_2}{M_3} \frac{m_n}{n} \int_0^{\infty} v |C(v)| dv.
\end{aligned}$$

Thus, for sufficiently large n, we have

$$E[\widehat{C}(u)] = \int_0^{\infty} m_n \big[\{ W(m_n(u+v)) + W(m_n(u-v)) \} C(v) \big] dv + O\left(\frac{m_n}{n} \right), \tag{29}$$

where $O(\cdot)$ is uniform. After making a suitable transformation, we have

$$E[\widehat{C}(u)] = \int_{-\infty}^{\infty} W(v) C\left(u - \frac{v}{m_n} \right) dv + O\left(\frac{m_n}{n} \right). \tag{30}$$

Note that the covariance function of a bandlimited process is infinitely differentiable. Therefore, we have

$$C\left(u-\frac{v}{m_n}\right)$$

$$= C(u)+\frac{v}{m_n}C_{(1)}(u)+\ldots+\frac{v^{r-1}}{m_n^{r-1}(r-1)!}C_{(r-1)}(u)+\frac{v^r}{r!\,m_n^r}C_{(r)}\left(u-\frac{\alpha v}{m_n}\right), \quad (31)$$

where $C_{(l)}(\cdot)$ is the lth order derivative of $C(\cdot)$ and α is an appropriate real number in the interval $(0,1)$. By using (31) and Property 2 of the weight function, we have, for large n,

$$E[\widehat{C}(u)] = C(u)+\frac{1}{m_n^r r!}\int_{-\infty}^{\infty} C_{(r)}\left(u-\frac{\alpha v}{m_n}\right)v^r W(v)dv+O\left(\frac{m_n}{n}\right).$$

From Assumption 3 and DCT, we have

$$E[\widehat{C}(u)] = C(u)+O\left(\frac{1}{m_n^r}\right)+O\left(\frac{m_n}{n}\right), \quad (32)$$

where $O(\cdot)$ is uniform. For the indirect estimation of $C(lT)$ for $l=1\ldots,J$, using (17) and (32) in (23), we have

$$E[y_n(u)]$$

$$= C(u)-\text{sinc}(\pi u/T)C(0)-\sum_{l=J+1}^{L_n}E[\hat{C}(lT)]$$

$$\times\left\{\text{sinc}\left(\frac{\pi}{T}(u-lT)\right)+\text{sinc}\left(\frac{\pi}{T}(u+lT)\right)\right\}+O\left(\frac{1}{m_n^r}\right)+O\left(\frac{m_n}{n}\right). \quad (33)$$

By using (32) and the fact that $\sum_{l=J+1}^{L_n}\left|\left\{\text{sinc}\left(\frac{\pi}{T}(u-lT)\right)+\text{sinc}\left(\frac{\pi}{T}(u+lT)\right)\right\}\right| \leq O(\log L_n)$, the third term of the right hand side of (33) simplifies to

$$\sum_{l=J+1}^{L_n}C(lT)\left\{\text{sinc}\left(\frac{\pi}{T}(u-lT)\right)+\text{sinc}\left(\frac{\pi}{T}(u+lT)\right)\right\}+O\left(\frac{\log L_n}{m_n^r}\right)+O\left(\frac{m_n\log L_n}{n}\right).$$

Now, we have

$$E(y_n(u)) = y(u)+\sum_{l=L_n+1}^{\infty}C(lT)\left\{\text{sinc}\left(\frac{\pi}{T}(u-lT)\right)+\text{sinc}\left(\frac{\pi}{T}(u+lT)\right)\right\}$$

$$+O\left(\frac{1}{m_n^r}\right)+O\left(\frac{\log L_n}{m_n^r}\right)+O\left(\frac{m_n\log L_n}{n}\right)$$

In view of Assumption 1A, we have

$$\left|\sum_{l=L_n+1}^{\infty} C(lT)\left\{\text{sinc}\left(\frac{\pi}{T}(u-lT)\right)+\text{sinc}\left(\frac{\pi}{T}(u+lT)\right)\right\}\right| \leq \sum_{l=L_n+1}^{\infty}|C(lT)|=O\left(\frac{1}{L_n^q}\right).$$

Thus, we have

$$E(y_n(u)) = y(u)+O\left(\frac{1}{L_n^q}\right)+O\left(\frac{\log L_n}{m_n^r}\right)+O\left(\frac{m_n \log L_n}{n}\right), \tag{34}$$

where the $O(\cdot)$ term is uniform.

Thus, for $j=1,\ldots,J$, using (6) and (34), we have

$$\begin{aligned} E&\left[\int_{u_1}^{u_2} x_j(u)y_n(u)du\right] \\ &= \int_{u_1}^{u_2} x_j(u)\left\{\sum_{l=1}^{J} x_l(u)C(lT)\right\} du+O\left(\frac{1}{L_n^q}\right)+O\left(\frac{\log L_n}{m_n^r}\right)+O\left(\frac{m_n \log L_n}{n}\right) \end{aligned} \tag{35}$$

By using (35) and (10), for $l=1,\ldots,J$, we have

$$E[\widehat{C}(lT)] = C(lT)+O\left(\frac{1}{L_n^q}\right)+O\left(\frac{\log L_n}{m_n^r}\right)+O\left(\frac{m_n \log L_n}{n}\right). \tag{36}$$

By using a similar argument as above, for the range $0<u\leq d$ but $u\neq T,\ldots,JT$, one can establish that

$$E[\widehat{C}(u)] = C(u)+O\left(\frac{1}{L_n^q}\right)+O\left(\frac{\log L_n}{m_n^r}\right)+O\left(\frac{m_n \log L_n}{n}\right).$$

This completes the proof. □

Proof of Theorem 2. Observe that

$$\begin{aligned} Var[\widehat{C}(0)] &= \frac{1}{n^2}\sum_{i_1=1}^{n}\sum_{i_2=1}^{n} E[X(t_{i_1})X(t_{i_2})] \\ &= \frac{C(0)}{n}+\frac{1}{n^2}\int_0^{\infty} C(v)\sum_{1\leq i_1<i_2\leq n} f^{(i_2-i_1)}(v)dv \end{aligned}$$

Thus, we have, from Assumption 2

$$Var[\widehat{C}(0)] = O\left(\frac{1}{n}\right). \tag{37}$$

By using Assumption 5, we have for $u > d$,

$$
\begin{aligned}
& E[\widehat{C}^2(u)] \\
= \; & \frac{m_n^2}{n^2 H^2(u)} \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \sum_{i_3=1}^{n} \sum_{i_4=1}^{n} E\Big[W\left(m_n(u - t_{i_1} + t_{i_2})\right) W\left(m_n(u - t_{i_3} + t_{i_4})\right) \\
& \times X(t_{i_1}) X(t_{i_2}) X(t_{i_3}) X(t_{i_4}) \Big] \\
= \; & \frac{m_n^2}{n^2 H^2(u)} \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \sum_{i_3=1}^{n} \sum_{i_4=1}^{n} E\Big[W\left(m_n(u - t_{i_1} + t_{i_2})\right) W\left(m_n(u - t_{i_3} + t_{i_4})\right) \\
& \times \Big\{ C(t_{i_1} - t_{i_2}) C(t_{i_3} - t_{i_4}) + C(t_{i_1} - t_{i_3}) C(t_{i_2} - t_{i_4}) \\
& + C(t_{i_1} - t_{i_4}) C(t_{i_2} - t_{i_3}) + Q(t_{i_1} - t_{i_4}, t_{i_2} - t_{i_4}, t_{i_3} - t_{i_4}) \Big\} \Big]
\end{aligned}
$$

Thus we have for $u > d$,

$$
\begin{aligned}
& Var[\widehat{C}(u)] \\
= \; & \frac{m_n^2}{n^2 H^2(u)} \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \sum_{i_3=1}^{n} \sum_{i_4=1}^{n} E\Big[W\left(m_n(u - t_{i_1} + t_{i_2})\right) W\left(m_n(u - t_{i_3} + t_{i_4})\right) \\
& \times \Big\{ C(t_{i_1} - t_{i_3}) C(t_{i_2} - t_{i_4}) + C(t_{i_1} - t_{i_4}) C(t_{i_2} - t_{i_3}) \\
& + Q(t_{i_1} - t_{i_4}, t_{i_2} - t_{i_4}, t_{i_3} - t_{i_4}) \Big\} \Big] \\
= \; & I_1(u) + I_2(u) + I_3(u), \qquad (38)
\end{aligned}
$$

Observe that the terms $I_1(u)$ and $I_2(u)$ are bounded from above by

$$
\begin{aligned}
I(u) \; = \; & M_1 \frac{m_n}{n} \frac{1}{n H^2(u)} \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \sum_{i_3=1}^{n} \sum_{i_4=1}^{n} \\
& E\Big[|m_n W\left(m_n(u - t_{i_1} + t_{i_2})\right) C(t_{i_1} - t_{i_3}) C(t_{i_2} - t_{i_4})| \Big] \qquad (39)
\end{aligned}
$$

It follows from Lemma 1 proved below that $I(u) = O(\frac{m_n}{n})$ where $O(\cdot)$ is uniform u. Lemma 2 proved below indicates that $I_3(u) = O(\frac{m_n}{n})$ where $O(\cdot)$ is uniform u.. Therefore, we have,

$$
Var(\widehat{C}(u)) = O\left(\frac{m_n}{n}\right) \quad \text{for } d < u < \infty \,. \qquad (40)
$$

We now consider the indirect estimators $\widehat{C}(T)$, ..., $\widehat{C}(JT)$. From (10), it is enough to consider the convergence of the variance covariance matrix of the vector $\left(\int_{u_1}^{u_2} x_1(u)y_n(u)du,\ \int_{u_1}^{u_2} x_2(u)y_n(u)du, \ldots,\ \int_{u_1}^{u_2} x_J(u)y_n(u)du\right)$. For $j, j' \in \{1, 2, \ldots J\}$, we compute

$$\begin{aligned} & Cov\left(\int_{u_1}^{u_2} x_j(u)y_n(u)du, \int_{u_1}^{u_2} x_{j'}(u)y_n(u)du\right) \\ &= \int_{u_1}^{u_2}\int_{u_1}^{u_2} x_j(u)x_{j'}(v)Cov(y_n(u), y_n(v))dudv. \end{aligned}$$

The interchange of the integrals is justified by the finiteness of the double integral, which follows from arguments similar to those given below to establish the convergence. Note that

$$\begin{aligned} & Var[y_n(u)] \\ = \ & Var\Bigg[\widehat{C}(u) - \text{sinc}(\pi u/T)\widehat{C}(0) \\ & - \sum_{l=J+1}^{L_n} \widehat{C}(lT)\left\{\text{sinc}\left(\frac{\pi}{T}(u-lT)\right) + \text{sinc}\left(\frac{\pi}{T}(u+lT)\right)\right\}\Bigg] \\ \leq \ & 3 \times Var[\widehat{C}(u)] + 3 \times Var[\text{sinc}(\pi u/T)\widehat{C}(0)] \\ & +3 \times Var\left[\sum_{l=J+1}^{L_n} \widehat{C}(lT)\left\{\text{sinc}\left(\frac{\pi}{T}(u-lT)\right) + \text{sinc}\left(\frac{\pi}{T}(u+lT)\right)\right\}\right] \\ = \ & 3 \times Var\left[\sum_{l=J+1}^{L_n} \widehat{C}(lT)\left\{\text{sinc}\left(\frac{\pi}{T}(u-lT)\right) + \text{sinc}\left(\frac{\pi}{T}(u+lT)\right)\right\}\right] \\ & +O\left(\frac{m_n}{n}\right) + O\left(\frac{1}{n}\right), \end{aligned}$$

where last inequality follows from (40) and (37). We now consider the first term. By using (40), we have

$$Var\left[\sum_{l=J+1}^{L_n} \widehat{C}(lT)\left\{\text{sinc}\left(\frac{\pi}{T}(u-lT)\right) + \text{sinc}\left(\frac{\pi}{T}(u+lT)\right)\right\}\right]$$

$$
\begin{aligned}
&= \sum_{l=J+1}^{L_n}\sum_{l'=J+1}^{L_n} Cov(\widehat{C}(lT),\widehat{C}(l'T))\left\{\text{sinc}\Big(\frac{\pi}{T}(u-lT)\Big)+\text{sinc}\Big(\frac{\pi}{T}(u+lT)\Big)\right\}\\
&\qquad\times\left\{\text{sinc}\Big(\frac{\pi}{T}(u-l'T)\Big)+\text{sinc}\Big(\frac{\pi}{T}(u+l'T)\Big)\right\}\\
&\le O\left(\frac{m_n}{n}\right)\left[\sum_{l=J+1}^{L_n}\left|\left\{\text{sinc}\Big(\frac{\pi}{T}(u-lT)\Big)+\text{sinc}\Big(\frac{\pi}{T}(u+lT)\Big)\right\}\right|\right]^2\\
&= O\left(\frac{m_n(\log L_n)^2}{n}\right),
\end{aligned}
$$

where $O(\cdot)$ is uniform.

Thus, for $j, j' \in \{1,2,\ldots,J\}$, we have

$$
\begin{aligned}
&Cov\left(\int_{u_1}^{u_2} x_j(u)y_n(u)du, \int_{u_1}^{u_2} x_{j'}(u)y_n(u)du\right)\\
&\quad = O\left(\frac{1}{n}\right)+O\left(\frac{m_n}{n}\right)+O\left(\frac{m_n(\log L_n)^2}{n}\right).
\end{aligned}
$$

This completes the proof. $\square$

Lemma 1. Under the Assumptions of Theorem 2, the function $I(u) = O\left(\frac{m_n}{n}\right)$, where $O(\cdot)$ is uniform in u.

Proof of Lemma 1. We partition the range of summation as

$$
\begin{aligned}
&\{(i_1,i_2,i_3,i_4) : 1 \le i_1,i_2,i_3,i_4 \le n\}\\
&\quad = \bigcup_{j=1}^{24} S_{1,j} \bigcup_{j=1}^{36} S_{2,j} \bigcup_{j=1}^{8} S_{3,j} \bigcup S_4,
\end{aligned}
$$

where $S_{1,j}$ for $j = 1,\ldots,24$ are sets of quadruples of indices having different types of strict order among themselves ($i_1 < i_2 < i_3 < i_4$, $i_1 < i_2 < i_4 < i_3$, and 22 other permutations) $S_{2,j}$ for $j = 1,\ldots,36$ are sets of quadruples of indices exactly two of which are equal and are in strict order with the other two indices ($i_1 < i_2 < i_3 = i_4$, $i_1 < i_3 = i_4 < i_2$, and 34 other arrangements) $S_{3,j}$ for $j = 1,\ldots,8$ are sets of quadruples of indices exactly three of which are equal and are in strict order with the fourth ($i_1 < i_2 = i_3 = i_4$, $i_2 = i_3 = i_4 < i_1$, and 6 other arrangements), and S_4 is the set $\{(i_1,i_2,i_3,i_4) : 1 \le i_1 = i_2 = i_3 = i_4 \le n\}$.

Consider $S_{1,1} = \{(i_1, i_2, i_3, i_4) : 1 \le i_1 < i_2 < i_3 < i_4 \le n\}$. By using the transformation $t_{i_2} - t_{i_1} = \vartheta_{i_2-i_1}$, $t_{i_3} - t_{i_2} = \vartheta_{i_3-i_2}$ and $t_{i_4} - t_{i_3} = \vartheta_{i_4-i_3}$, and by making use of the fact that the transformed random variables are independent, we have

$$
\begin{aligned}
& \frac{M_1}{nH^2(u)} \sum_{S_{1,1}} E\Big[|m_n W(m_n(u - t_{i_1} + t_{i_2}))\, C(t_{i_1} - t_{i_3}) C(t_{i_2} - t_{i_4})|\Big] \\
= & \frac{M_1}{nH^2(u)} \sum_{S_{1,1}} E\Big[|m_n W\left(m_n(u - \vartheta_{i_2-i_1})\right) C(\vartheta_{i_3-i_2} + \vartheta_{i_2-i_1}) C(\vartheta_{i_4-i_3} + \vartheta_{i_3-i_2})|\Big] \\
= & \frac{M_1}{nH^2(u)} \sum_{S_{1,1}} \int_0^\infty \int_0^\infty \int_0^\infty |m_n W\left(m_n(u - v_1)\right)|\, |C(v_2 + v_1) C(v_3 + v_2)| \\
& \qquad \times f^{(i_2-i_1)}(v_1) f^{(i_3-i_2)}(v_2) f^{(i_4-i_3)}(v_3) dv_1 dv_2 dv_3 \\
= & \frac{M_1}{H^2(u)} \int_0^\infty \int_0^\infty \int_0^\infty |m_n W(m_n(u - v_1))\, C(v_2 + v_1) C(v_3 + v_2)| \\
& \qquad \times \left\{ \frac{1}{n} \sum_{S_{1,1}} f^{(i_2-i_1)}(v_1) f^{(i_3-i_2)}(v_2) f^{(i_4-i_3)}(v_3) \right\} dv_1 dv_2 dv_3.
\end{aligned}
$$

Now note that

$$
\frac{1}{n} \sum_{S_{1,1}} f^{(i_2-i_1)}(v_1) f^{(i_3-i_2)}(v_2) f^{(i_4-i_3)}(v_3) \le H(v_1) H(v_2) H(v_3) \le M_2^3.
$$

Thus, we have

$$
\begin{aligned}
& \frac{M_1}{nH^2(u)} \sum_{S_{1,1}} E\Big[|m_n W(m_n(u - t_{i_1} + t_{i_2}))\, C(t_{i_1} - t_{i_3}) C(t_{i_2} - t_{i_4})|\Big] \\
\le & \frac{M_1 M_2^3}{M_3^2} \int_0^\infty \int_0^\infty \int_0^\infty |m_n\left(m_n(u - v_1)\right) C(v_2 + v_1) C(v_3 + v_2)|\, dv_1 dv_2 dv_3. \\
= & M_4 \int_0^\infty m_n |W\left(m_n(u - v_1)\right)| \int_0^\infty |C(v_2 + v_1)| \int_0^\infty |C(v_3 + v_2)| dv_3 dv_2 dv_1 \\
\le & M_5 \int_0^\infty m_n |W\left(m_n(u - v_1)\right)| \int_0^\infty |C(v_2 + v_1)| dv_2 dv_1 \\
\le & M_6 \int_0^\infty m_n |W\left(m_n(u - v_1)\right)| dv_1 \le M_7,
\end{aligned}
$$

where M_4, M_5, M_6 and M_7 are different constants.

By using similar arguments, one can establish the boundedness of

$$\frac{M_1}{nH^2(u)}\sum_{S_{1,j}}E\Big[|m_nW(m_n(u-t_{i_1}+t_{i_2}))\,C(t_{i_1}-t_{i_3})C(t_{i_2}-t_{i_4})|\Big]$$

for the partitions $S_{1,j},\ 2 \leq j \leq 22$. A slightly different argument is needed for $S_{1,23} = \{(i_1, i_2, i_3, i_4) : 1 \leq i_1 < i_4 < i_3 < i_2 \leq n\}$ and $S_{1,24} = \{(i_1, i_2, i_3, i_4) : 1 \leq i_2 < i_3 < i_4 < i_1 \leq n\}$. Consider the case of $S_{1,24}$. By using the transformation $\vartheta_{i_4-i_1} = t_{i_4} - t_{i_1}$, $\vartheta_{i_3-i_4} = t_{i_3} - t_{i_4}$ and $\vartheta_{i_2-i_3} = t_{i_2} - t_{i_3}$ and using the fact that $\vartheta_{i_4-i_1}$, $\vartheta_{i_3-i_4}$ and $\vartheta_{i_2-i_3}$ are independent random variables, we have

$$\begin{aligned}
&\frac{M_1}{nH^2(u)}\sum_{S_{1,23}}E\Big[|m_nW(m_n(u-t_{i_1}+t_{i_2}))\,C(t_{i_1}-t_{i_3})C(t_{i_2}-t_{i_4})|\Big]\\
&= \frac{M_1}{nH^2(u)}\sum_{S_{1,23}}E\Big[\,|m_nW\left(m_n(u+\vartheta_{i_4-i_1}+\vartheta_{i_3-i_4}+\vartheta_{i_2-i_3})\right)|\\
&\qquad\times|C(\vartheta_{i_4-i_1}+\vartheta_{i_3-i_4})C(\vartheta_{i_3-i_4}+\vartheta_{i_2-i_3})|\,\Big]\\
&= \frac{M_1}{H^2(u)}\int_0^\infty\int_0^\infty\int_0^\infty |m_nW\left(m_n(u+v_1+v_2+v_3)\right)|\,|C(v_1+v_2)C(v_2+v_3)|\\
&\qquad\times\left\{\frac{1}{n}\sum_{S_{1,23}}f^{(i_4-i_1)}(v_1)f^{(i_3-i_4)}(v_2)f^{(i_2-i_3)}(v_3)\right\}dv_1dv_2dv_3\\
&\leq \frac{M_1M_2^3}{M_3^2}\int_0^\infty\int_0^\infty\int_0^\infty |m_nW\left(m_n(u+v_1+v_2+v_3)\right)|\\
&\qquad\times|C(v_1+v_2)C(v_2+v_3)|\,dv_1dv_2dv_3\\
&\leq \frac{M_1M_2^3}{M_3^2}\int_0^\infty m_n|W\left(m_n(u+v_1')\right)|\\
&\qquad\times\left[\int_0^\infty |C(v_1'-v_3)|\left\{\int_0^\infty |C(v_2+v_3)|dv_2\right\}dv_3\right]dv_1'\\
&\leq M_4\int_0^\infty m_n|W\left(m_n(u+v_1')\right)|\left[\int_0^\infty |C(v_1'-v_3)|dv_3\right]dv_1'\\
&\leq M_5\int_0^\infty m_n|W\left(m_n(u+v_1')\right)|dv_1' \;\leq\; M_6,
\end{aligned}$$

where M_4, M_5, and M_6 are different constants. The boundedness of the sum over $S_{1,23}$ can be established in a similar manner.

In subsets $S_{2,j}$ for $j = 1, \ldots, 36$, the summation runs over only three indices. Consider $S_{2,1} = \{(i_1, i_2, i_3, i_4) : i_1 < i_2 < i_3 = i_4\}$, and the transformation $\vartheta_{i_2-i_1} = t_{i_2} - t_{i_1}$ and $\vartheta_{i_3-i_2} = t_{i_3} - t_{i_2}$. Then we have

$$
\begin{aligned}
&\frac{M_1}{nH^2(u)} \sum_{S_{2,1}} E\Big[|m_n W(m_n(u - t_{i_1} + t_{i_2}))\, C(t_{i_1} - t_{i_3}) C(t_{i_2} - t_{i_4})|\Big] \\
= \ & \frac{M_1}{nH^2(u)} \sum_{S_{2,1}} E\Big[\, |m_n W\left(m_n(u + \vartheta_{i_2-i_1})\right)| \\
& \quad \times |C(\vartheta_{i_2-i_1} + \vartheta_{i_3-i_2}) C(\vartheta_{i_3-i_2})|\,\Big] \\
= \ & \frac{M_1}{H^2(u)} \int_0^\infty \int_0^\infty |m_n W\left(m_n(u + v_1)\right) C(v_1 + v_2) C(v_2)| \\
& \quad \times \left\{ \frac{1}{n} \sum_{S_{2,1}} f^{(i_2-i_1)}(v_1) f^{(i_3-i_2)}(v_2) \right\} dv_1 dv_2 \\
\leq \ & \frac{M_1}{H^2(u)} \int_0^\infty \int_0^\infty |m_n W\left(m_n(u + v_1)\right) C(v_1 + v_2) C(v_2)| \\
& \quad \times H(v_1) H(v_2) dv_1 dv_2 \\
\leq \ & \frac{M_1 M_2^2 C(0)}{M_3^2} \int_0^\infty m_n |W\left(m_n(u + v_1)\right)| \int_0^\infty |C(v_2)| dv_2 dv_1 \\
\leq \ & M_4,
\end{aligned}
$$

where M_4 is a constant. A similar argument can be used to establish the boundedness of the sums over 29 other sets of quadruples of indices with $i_1 \neq i_2$. A slightly different argument is needed in the cases of the six sets with $i_1 = i_2$. We show the calculations for $S_{2,31} = \{(i_1, i_2, i_3, i_4) : i_1 = i_2 < i_3 < i_4\}$, as a representative of these six sets. By using the transformation $\vartheta_{i_3-i_2} = t_{i_3} - t_{i_2}$ and $\vartheta_{i_4-i_3} = t_{i_4} - t_{i_3}$, we have

$$
\begin{aligned}
&\frac{M_1}{nH^2(u)} \sum_{S_{2,31}} E\Big[|m_n W(m_n(u - t_{i_1} + t_{i_2}))\, C(t_{i_1} - t_{i_3}) C(t_{i_2} - t_{i_4})|\Big] \\
= \ & \frac{M_1}{nH^2(u)} \sum_{S_{2,31}} E\Big[\, |m_n W\left(m_n u\right) C(\vartheta_{i_3-i_2}) C(\vartheta_{i_3-i_2} + \vartheta_{i_4-i_3})|\,\Big]
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{M_1}{H^2(u)} \cdot m_n |W(m_n u)| \int_0^\infty \int_0^\infty |C(v_1)C(v_1+v_2)| \\
&\quad \times \left\{ \frac{1}{n} \sum_{S_{2,31}} f^{(i_3-i_2)}(v_1) f^{(i_4-i_3)}(v_2) \right\} dv_1 dv_2 \\
&\leq \frac{M_1}{H^2(u)} \cdot m_n |W(m_n u)| \int_0^\infty \int_0^\infty |C(v_1)C(v_1+v_2)| \\
&\quad \times H(v_1)H(v_2) dv_1 dv_2 \\
&\leq \frac{M_1 M_2^2}{M_3^2} \cdot m_n |W(m_n u)| \int_0^\infty |C(v_1)| \int_0^\infty |C(v_1+v_2)| dv_2 dv_1 \\
&\leq M_4 \cdot m_n |W(m_n u)|,
\end{aligned}
$$

where M_4 is a constant. For sufficiently large n (such that $m_n > a/d$), the last expression is identically zero. The threshold a/d does not depend on u. Thus, this term is identically zero for large n, uniformly for all u.

Now consider the double sums over the subsets $S_{3,1}, \ldots, S_{3,8}$. We will show the boundedness of the sums over $S_{3,1} = \{(i_1, i_2, i_3, i_4) : 1 \leq i_1 = i_2 = i_3 < i_4 \leq n\}$ and $S_{3,2} = \{(i_1, i_2, i_3, i_4) : 1 \leq i_1 < i_2 = i_3 = i_4 \leq n\}$, each case being a representative of the calculations needed in three other cases. We proceed with the sum over $S_{3,1}$ as follows.

$$
\begin{aligned}
&\frac{M_1}{nH^2(u)} \sum_{S_{3,1}} E\Big[|m_n W(m_n(u - t_{i_1} + t_{i_2}))\, C(t_{i_1} - t_{i_3}) C(t_{i_2} - t_{i_4})| \Big] \\
&= \frac{M_1}{H^2(u)} \int_0^\infty |m_n W(m_n u)\, C(0) C(v)| \left\{ \frac{1}{n} \sum_{S_{3,1}} f^{(i_4-i_3)}(v) \right\} dv \\
&\leq \frac{M_1}{H^2(u)} m_n |W(m_n u)| C(0) \int_0^\infty |C(v)|\, H(v) dv \\
&\leq \frac{M_1 M_2}{M_3^2} m_n |W(m_n u)| C(0) \int_0^\infty |C(v)|\, dv \\
&= M_4 \times m_n |W(m_n u)|,
\end{aligned}
$$

for some constant M_4. For sufficiently large n (such that $m_n > a/d$), the last expression is identically zero. The threshold a/d does not depend on u. Thus, this term is also identically zero for large n, uniformly for all u. On the other hand,

$$
\begin{aligned}
& \frac{M_1}{nH^2(u)} \sum_{S_{3,2}} E\Big[|m_n W(m_n(u - t_{i_1} + t_{i_2}))\, C(t_{i_1} - t_{i_3}) C(t_{i_2} - t_{i_4})| \Big] \\
= \;& \frac{M_1}{H^2(u)} \int_0^\infty |m_n W\left(m_n(u+v)\right) C(v) C(0)| \\
& \times \left\{ \frac{1}{n} \sum_{S_{3,2}} f^{(i_2 - i_1)}(v) \right\} dv \\
\le \;& \frac{M_1}{H^2(u)} C^2(0) \int_0^\infty m_n \left| W\left(m_n(u+v)\right)\right| H(v) dv
\end{aligned}
$$

$$
\le \frac{M_1 M_2}{M_3^2} C^2(0) \int_0^\infty |W(v)|\, dv \quad \le M_4
$$

for some constant M_4.

The sum over S_4 does not involve any random variable, and is bounded as

$$
\begin{aligned}
& \frac{M_1}{nH^2(u)} \sum_{S_4} E\Big[|m_n W(m_n(u - t_{i_1} + t_{i_2}))\, C(t_{i_1} - t_{i_3}) C(t_{i_2} - t_{i_4})| \Big] \\
\le \;& \frac{M_1}{nM_3^2} m_n W\left(m_n u\right) C^2(0).
\end{aligned}
$$

Again for large n such that $m_n > a/d$, the upper bound happens to be identically zero. This completes the proof. □

Lemma 2. Under the Assumptions of Theorem 2, the function $I_3(u) = O\left(\frac{m_n}{n}\right)$, where $O(\cdot)$ is uniform in u.

Proof of Lemma 2. It follows from Assumption 5 that $I_3(u)$ is bounded as

$$
\begin{aligned}
& \frac{n}{m_n} |I_3(u)| \\
\le \;& \frac{M_1}{nM_3^2} \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \sum_{i_3=1}^{n} \sum_{i_4=1}^{n} E\big[m_n |W\left(m_n(u - t_{i_1} + t_{i_2})\right)| \\
& \times g(t_{i_2} - t_{i_4}) g(t_{i_3} - t_{i_4}) \big].
\end{aligned}
$$

The proof follows from an argument similar to the one used in the proof of Lemma 1. □

Proof of Theorem 3. Part (i). From (11), for $\lambda \in [-\lambda_0, \lambda_0]$, we have

$$\begin{aligned} E[\widehat{\phi}_n(\lambda)] &= \frac{T}{2\pi} E[\widehat{C}(0)] + \frac{T}{\pi} \sum_{l=1}^{J} E[\widehat{C}(lT)] K(b_n l) \cos(l\lambda T) \\ &\quad + \frac{T}{\pi} \sum_{l=J+1}^{n-1} E[\widehat{C}(lT)] K(b_n l) \cos(l\lambda T) \end{aligned} \tag{41}$$

Note that the second term of right hand side of (41) is a finite sum. By using Theorem 1 Part (i) and Property 1 of kernel $K(\cdot)$ and Assumption 6, we have

$$\begin{aligned} \lim_{n\to\infty} E[\widehat{\phi}_n(\lambda)] &= \frac{T}{2\pi} C(0) + \frac{T}{\pi} \sum_{l=1}^{J} C(lT) \cos(l\lambda T) \\ &\quad + \lim_{n\to\infty} \frac{T}{\pi} \sum_{l=J+1}^{n-1} E[\widehat{C}(lT)] K(b_n l) \cos(l\lambda T). \end{aligned} \tag{42}$$

By using the exact expectation of $E(\widehat{C}(lT))$ for $l > J$, from the expression (21) for large n, we have

$$\begin{aligned} &\frac{T}{\pi} \sum_{l=J+1}^{n-1} E[\widehat{C}(lT)] K(b_n l) \cos(l\lambda T) \\ &= \frac{T}{\pi} \int_{-\infty}^{\infty} W(v) \Bigg\{ \sum_{l=J+1}^{n-1} K(b_n l) \cos(l\lambda T) \frac{1}{H(lT)} C\left(lT - \frac{v}{m_n}\right) \\ &\qquad \times \left[H_n\left(lT - \frac{v}{m_n}\right) + H_n\left(-lT + \frac{v}{m_n}\right) \right] \Bigg\} dv. \end{aligned} \tag{43}$$

From Assumptions 1, 2 and Property 1 of kernel function, the integrand of (43) is dominated as

$$\begin{aligned}
& |W(v)| \sum_{l=J+1}^{n-1} \left| K(b_n l)\cos(l\lambda T)\frac{1}{H(lT)} C\left(lT - \frac{v}{m_n}\right) \right. \\
& \qquad \left. \times \left[H_n\left(lT - \frac{v}{m_n}\right) + H_n\left(-lT + \frac{v}{m_n}\right)\right] \right| \\
\leq \; & |W(v)| \sup |K(\cdot)| \frac{2M_2}{M_3} \sum_{l=J+1}^{n-1} g\left(lT - \frac{v}{m_n}\right) \\
\leq \; & |W(v)| \sup K(\cdot) \frac{2M_2}{M_3} \times 2 \sum_{l=0}^{\infty} g\left(lT\right).
\end{aligned}$$

Integrability of the bounding function is ensured by Property 1 of weight function $W(\cdot)$. From Assumptions 3 and 6, we have pointwise convergence as

$$\begin{aligned}
& \lim_{n\to\infty} W(v) \sum_{l=J+1}^{n-1} K(b_n l)\cos(l\lambda T)\frac{1}{H(lT)} C\left(lT - \frac{v}{m_n}\right) \\
& \qquad \times \left[H_n\left(lT - \frac{v}{m_n}\right) + H_n\left(-lT + \frac{v}{m_n}\right)\right] \\
= \; & W(v) \sum_{l=J+1}^{\infty} \cos(l\lambda T) C(lT).
\end{aligned}$$

Applying DCT, we have

$$\lim_{n\to\infty} \frac{T}{\pi} \sum_{l=J+1}^{n-1} E[\widehat{C}(lT)] K(b_n l)\cos(l\lambda T) \;=\; \frac{T}{\pi} \sum_{l=J+1}^{\infty} C(lT)\cos(l\lambda T).$$

From (42), we have

$$\lim_{n\to\infty} E[\widehat{\phi}_n(\lambda)] = \frac{T}{2\pi} C(0) + \frac{T}{\pi} \sum_{l=1}^{\infty} C(lT)\cos(l\lambda T) = \phi(\lambda).$$

This completes the proof of Part (i).

Part (ii). By using Theorem 1 Part (ii), we have

$$\begin{aligned} E[\widehat{\phi}_n(\lambda)] &= \frac{T}{2\pi}C(0)+\frac{T}{\pi}\sum_{l=1}^{n-1}C(lT)K(b_nl)\cos(l\lambda T)+\left\{\sum_{l=1}^{J}K(b_nl)\cos(l\lambda T)\right\} \\ &\times\left[O\left(\frac{1}{L_n^q}\right)+O\left(\frac{\log L_n}{m_n^r}\right)+O\left(\frac{m_n\log L_n}{n}\right)\right] \\ &+\left\{\sum_{l=J+1}^{n-1}K(b_nl)\cos(l\lambda T)\right\}\times O\left(\frac{1}{m_n^r}\right) \end{aligned}$$

Note that

$$\sum_{l=J}^{n-1}K(b_nl)\cos(l\lambda T)=O\left(\frac{1}{b_n}\right). \tag{44}$$

Since

$$\phi(\lambda)=\frac{T}{2\pi}C(0)+\frac{T}{\pi}\sum_{l=1}^{\infty}C(lT)\cos(l\lambda T),$$

we have,

$$\begin{aligned} &E[\widehat{\phi}_n(\lambda)]-\phi(\lambda) \\ &= -\frac{b_n^qT}{\pi}\sum_{l=1}^{n-1}\frac{(1-K(b_nl))}{(b_nl)^q}l^qC(lT)\cos(l\lambda T)-\frac{T}{\pi}\sum_{l=n}^{\infty}C(lT)\cos(l\lambda T) \\ &\quad +O\left(\frac{1}{L_n^q}\right)+O\left(\frac{\log L_n}{m_n^r}\right)+O\left(\frac{m_n\log L_n}{n}\right)+O\left(\frac{1}{b_nm_n^r}\right) \end{aligned} \tag{45}$$

By using Assumption 1A, we have

$$\left|\sum_{l=n}^{\infty}C(lT)\cos(l\lambda T)\right|\leq\sum_{l=n}^{\infty}|C(lT)|=O\left(\frac{1}{n^q}\right).$$

By using Assumption 1A and Property 2 of the kernel function and the DCT, we have

$$\lim_{n\to\infty}\sum_{l=1}^{n-1}\frac{(1-K(b_nl))}{(b_nl)^q}l^qC(lT)\cos(l\lambda T)=\frac{k_q}{T^q}\sum_{l=1}^{\infty}(lT)^qC(lT)\cos(l\lambda T),$$

where $k_q = \lim_{n\to 0} \frac{1-K(x)}{|x|^q}$ which is non-zero for qth order kernel $K(\cdot)$. Thus, we have

$$\begin{aligned} E[\widehat{\phi}_n(\lambda)] - \phi(\lambda) &= O(b_n^q) + O\left(\frac{1}{n^q}\right) + O\left(\frac{1}{L_n^q}\right) + O\left(\frac{\log L_n}{m_n^r}\right) \\ &\quad + O\left(\frac{m_n \log L_n}{n}\right) + O\left(\frac{1}{b_n m_n^r}\right). \end{aligned}$$

The proof is completed by observing that, in view of Assumptions 1A, 4 and 5, the second term on the right hand side can be ignored in the presence of the fifth term. □

Proof of Theorem 4. From (11), we have

$$\begin{aligned} Var[\widehat{\phi}_n(\lambda)] &\leq \frac{3T^2}{(2\pi)^2} Var[\widehat{C}(0)] + \frac{3T^2}{\pi^2} Var\left[\sum_{l=1}^{J} \widehat{C}(lT)K(b_n l)\cos(l\lambda T)\right] \\ &\quad + \frac{3T^2}{\pi^2} Var\left[\sum_{l=J+1}^{n-1} \widehat{C}(lT)K(b_n l)\cos(l\lambda T)\right]. \end{aligned} \tag{46}$$

From (37), the first term on the right hand side is $O\left(\frac{1}{n}\right)$. The second term of right hand side of (46) is a finite sum, from Theorem 2, we have

$$Var\left[\sum_{l=1}^{J} \widehat{C}(lT)K(b_n l)\cos(l\lambda T)\right] = O\left(\frac{m_n(\log L_n)^2}{n}\right).$$

Now consider the third term. We have from Theorem 2

$$\begin{aligned} & Var\left[\sum_{l=J+1}^{n-1} \widehat{C}(lT)K(b_n l)\cos(l\lambda T)\right] \\ &= \sum_{l=J+1}^{n-1} \sum_{l'=J+1}^{n-1} Cov(\widehat{C}(lT), \widehat{C}(l'T))K(b_n l)\cos(l\lambda T)K(b_n l')\cos(l'\lambda T) \\ &\leq \left\{\sum_{l=J+1}^{n-1} |K(b_n l)|\right\}^2 \times O\left(\frac{m_n}{n}\right). \end{aligned}$$

Note that from Property 1 of kernel function, we have

$$\sum_{l=J+1}^{n-1} |K(b_n l)| = O\left(\frac{1}{b_n}\right).$$

Thus we have, from Assumption 3 and 6,

$$Var\left[\sum_{l=J+1}^{n-1}\widehat{C}(lT)K(b_n l)\cos(l\lambda T)\right] = O\left(\frac{m_n}{nb_n^2}\right).$$

This completes the proof. □

References

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

[2] A. Tarczynski and N. Allay, "Spectral Analysis of Randomly Sampled Signals: Suppression of Aliasing and Sampler Jitter", *IEEE Trans. Signal Processing*, vol. 52, no. 12, pp. 3324–3334, Dec. 2004.

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

[4] J. M. Nohrden, *Spectral Estimation Using Nonuniform Sampling*, Master's Thesis, Department of Electrical and Computer Engineering, Massachusetts Institute of Technology, Boston, 1995.

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

[6] 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.

[7] R. Srivastava and D. Sengupta, "Effect of inter-sample spacing constraint on spectrum estimation with irregular sampling" *IEEE Trans. Inf. Theor.*, vol. 57, no. 7, pp. 4709–4719, 2011.

[8] 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.

[9] D. R. Brillinger "The spectral analysis of stationary interval functions", *Proc. Sixth Berkeley Symp. Prob. Statist.*, pp. 483–513, 1972.

[10] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing* (Third Edition). Upper Saddle River, N. J.: Pentice-Hall, 2009.

[11] J.O. Ramsay, G. Hooker and S. Graves, *Functional data analysis with R and Matlab*. Spriger: New York , 2009.

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

[13] 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.

[14] D. J. Daley, Weakly stationary point processes and random measures, *J. Roy. Statist. Soc. Ser. B*, vol. 33, pp. 406-428, 1971.

[15] R. L. Cook, Stochastic sampling in computer graphics, *ACM Trans. Graphics*, vol. 5, No. 1, pp. 51–72, 1986.

[16] M. A. Z. Dippé, and E. H. Wold, Antialiasing through stochatic sampling, *ACM Siggraph*, vol. 19, No. 3, pp. 69–78, 1985.

[17] D. Qu, and A. Tarczynski, A novel spectral estimation method by using periodic nonuniform sampling, *Forty-First Asilomar Conference on ACSSC*, pp. 1134 - 1138, 2007.

[18] A. Ouahabi, C. Depollier, L. Simon, and D. Kouamé, Spectrum estimation from randomly sampled velocity data, *IEEE Trans. Instru. Measur.*, vol. 47, No. 4, pp. 1005–1012, 1998.

[19] J. R. Petit, J. Jouzel, D. Raynaud, N. I. Barkov, J.-M. Barnola, I. Basile, M. Bender, J. Chappellaz, M. Davis, G. Delaygue, M. Delmotte, V.M. Kotlyakov, M. Legrand, V. Y. Lipenkov, C. Lorius, L. épin, C. Ritz, E. Saltzman and M. Stievenard, Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica, *Nature*, vol. 399, pp. 429-435, 1999.