Spectrum Estimation with Uniformly and Stochastically Sampled Data: Some Challenges and Strategies

Radhendushka Srivastava

Indian Statistical Institute, Kolkata

April, 2011

Spectrum Estimation with Uniformly and Stochastically Sampled Data: Some Challenges and Strategies

Radhendushka Srivastava

Thesis submitted to the Indian Statistical Institute
in partial fulfillment of the requirements
for the award of the degree of
Doctor of Philosophy

April, 2011

Indian Statistical Institute
203, B.T. Road, Kolkata, India.

To my parents and my sister Anshu

Acknowledgements

I started learning statistics by chance. Statistics was one of the subjects in my under-graduate course. But very soon I started enjoying the subject by realizing its wide applications. From the moment I came to know about the Indian Statistical Institute, it was my dream to be a part of this great institution. I am extremely lucky that I got the opportunity to pursue research in statistics in ISI. Looking back into the four years since I became a member of the family of research scholars, it was not easy to complete my dream without the help and support from my teachers and friends. I would like to take this opportunity to thank them.

First and foremost, I want to express my gratitude to my supervisor Prof. Debasis Sengupta, whose encouragement and careful supervision from the preliminary to the concluding level made this thesis possible. It was a wonderful experience working with such a caring and affectionate researcher like him. Throughout my research period, he provided encouragement, sound advice, excellent teaching, good company, and lots of good ideas. I would have been lost without him.

I would like to express my indebtedness to Prof. B.V. Rao, who taught me basic and advanced probability theory. His advices on using various mathematical tools helped me to resolve many issues arising from my research work. I am grateful to Prof. Arup Bose and Prof. Probal Chaudhari for their suggestions during my course work, which helped me to improve my basic understanding of statistics and mathematical analysis and my knowledge of various aspects of time series analysis and multivariate analysis. I would like to thank Prof. Somesh C. Bagchi, Prof. Krishanu Maulik, Prof. Amites Dasgupta, Prof. Sashi M. Srivastava, Prof. Gopal K. Basak, Prof. Atanu Biswas and Prof. Anup Dewanji, who had been my teachers in various courses. I would also like to thank Prof. Anil K. Ghosh for various academic discussions. I further express my thanks to Prof. Girija K. Shukla and Prof. Abhijit Bhattacharya of Indian Institute of Management Lucknow for encouraging me to pursue a PhD degree.

I dedicate this thesis to my sister late Anshu. She was an excellent scholar and gave many advices that proved to be useful later. She taught me basic statistics and guided me to qualify the CSIR research fellowship. Her emotional support and care gave me a lot of confidence. Initially, academic life was very difficult in ISI but support from and discussions with my friends Abhijit Mandal, Rajat Subhra Hazra, Biswarup Das, Kaushik Saha, Arunanghsu Biswas, Subhra Sankar Dhar, Soumen Sarkar, Ashis Mandal and Abhijit Pal made it easier. I would like to express my special thanks to all of them. It is my pleasure to thank my friends Subhajit, Palash and Buddhananda, whose support and various academic discussions made life enjoyable in ISI. I would like to thank all my seniors and juniors. I would also like to thank my friends Pankaj, Abhishek, Hasan, Sachin and Siddhartha, who supported me throughout my college life.

Finally, I would like to express my gratitude to my parents, whose hard work, emotional support, love and care made it possible to fulfill my dreams. I would also like to acknowledge my younger brother for his love and support.

22 April, 2011 — Radhendushka Srivastava

Notations and abbreviations

Notation	Meaning	Page number of first use
i	$\sqrt{-1}$	1
X	Stationary stochastic process $\{X(t), -\infty < t < \infty\}$	1
$C(\cdot)$	Covariance function of X	1
$\phi(\cdot)$	Power spectral density of X	1
$\Phi(\cdot)$	Power spectral distribution of X	1
f	Frequency	1
f_0	Bandwidth of X when it is a bandlimited process	1
τ	$\{t_j,\ j = \ldots, -2, -1, 0, 1, 2, \ldots\}$ (set of sampling times)	2
X_τ	$\{X(t_j) :\ j = \ldots, -2, -1, 0, 1, 2, \ldots\}$ (sampled process)	2
c	Covariance sequence of X_τ	2
$\text{sinc}(u)$	$\begin{cases} \frac{\sin(u)}{u} & \text{if } u \neq 0, \\ 1 & \text{if } u = 0. \end{cases}$	3
T	$\frac{1}{2f_0}$, inter-sample spacing when bandlimited process X is sampled at Nyquist rate	3
$1_E(\cdot)$	Indicator function of the set E	4
n	Sample size	4
$K(\cdot)$	Covariance averaging kernel	4
b_n	Kernel bandwidth	4
β	Mean intensity of τ when it is a stationary point process	6

Notation	Meaning	Page number of first use
$\widehat{\psi}_n(f)$	$\frac{2}{\beta n}\sum_{l=1}^{n-1}\sum_{j=1}^{n-l}X(t_j)X(t_{j+l})K(b_n(t_{j+l}-t_j))\cos(2\pi f(t_{j+l}-t_j))$, an estimator of $\phi(f)$ based on n samples	6
d	Minimum permissible inter-sample spacing	7
ρ_n	Sampling rate when τ is a grid and sample size is n	13
$\widehat{c}_{ln}$	$\frac{1}{n}\sum_{j=1}^{n-\|l\|}X\left(\frac{j}{\rho_n}\right)X\left(\frac{j+\|l\|}{\rho_n}\right)$, an estimator of c_l	13
$\widehat{\phi}_n(f)$	$\frac{1}{\rho_n}\sum_{\|j\|<n}\widehat{c}_{jn}K(b_n j)e^{-\frac{i2\pi f j}{\rho_n}}1_{[-1/2\rho_n,1/2\rho_n]}(f)$, an estimator of $\phi(f)$ based on n samples	13
MSE	Mean squared error	14
$Q(\cdot)$	Fourth order cumulant function of X	15
p	Decay parameter of $\phi(\cdot)$	16
q	Smoothness parameter of $\phi(\cdot)$	16
DCT	Dominated Convergence Theorem	24
μ_τ	Reduced covariance measure of τ when it is a stationary point process	83
$\gamma(\cdot)$	Probability density function of inter-sample spacing when τ is a stationary renewal process	94
$H(\cdot)$	Renewal density function corresponding to $\gamma(\cdot)$	94
$\widehat{C}_{nd}(u)$	An estimator of $C(u)$ defined in (5.15)	109
$\widehat{\phi}_{nd}(f)$	$T\widehat{C}_{nd}(0)+2T\sum_{l=1}^{n-1}\widehat{C}_{nd}(lT)K(b_n l)\cos(2\pi f lT)1_{[-f_0,f_0]}(f)$, an estimator of $\phi(f)$ based on n samples with minimum inter-sample spacing d	109

Contents

Chapter 1

Introduction

1.1 Power spectral density and sampling

Let $X = \{X(t),\ -\infty < t < \infty\}$ be a real valued, continuous time, mean square continuous stationary stochastic process with mean 0 and integrable autocovariance function (also referred to as covariance function) $C(\cdot)$ defined as

$$C(t) = E\left[X(t+u)X(u)\right], \quad -\infty < t < \infty. \tag{1.1}$$

The power spectral density (also referred to as spectral density) of the process, $\phi(\cdot)$, is the Fourier transform of $C(\cdot)$, defined as

$$\phi(f) = \int_{-\infty}^{\infty} C(t)e^{-i2\pi ft}dt, \quad -\infty < f < \infty. \tag{1.2}$$

It can be shown that the power spectral density is not only real valued, but is also non-negative, even and integrable (Brillinger, 2001). In the sequel, we shall refer to the variable f as the frequency. The power spectral distribution function, $\Phi(\cdot)$, is defined as

$$\Phi(f) = \int_{-\infty}^{f} \phi(u)du, \quad -\infty < f < \infty. \tag{1.3}$$

The process X is said to be bandlimited if $\phi(\cdot)$ assumes the value zero outside a finite interval. If the smallest such interval is $[-f_0, f_0]$, then this interval is called the spectral support, and the frequency f_0 is called the bandwidth of the process X. If there is no finite interval for which this condition holds, then the process is said to be non-bandlimited.

One often attempts to gather information about $\phi(\cdot)$ through a set of sample values of the process X. Let τ be a countable set of real numbers and X_τ be the sampled process, defined as

$$X_\tau = \{X(t),\ t \in \tau\}. \tag{1.4}$$

Note that X_τ can be regarded as a discrete time process. The most common example of a sampled process is a uniformly (regularly) sampled process, which corresponds to the choice $\tau = \{j/\rho : j =, \ldots, -2, -1, 0, 1, 2, \ldots\}$, where ρ is the (fixed) sampling rate. In this case, the autocovariance sequence $c = \{\ldots, c_{-2}, c_{-1}, c_0, c_1, c_2, \ldots\}$ of the process X_τ is defined as

$$c_j = E\left[X\left(\frac{j+k}{\rho}\right) X\left(\frac{k}{\rho}\right)\right] = C\left(\frac{j}{\rho}\right), \qquad j = \ldots, -2, -1, 0, 1, 2, \ldots. \tag{1.5}$$

The power spectral density of X_τ is the Fourier transform $\phi_\rho(\cdot)$ of the sequence c, given by

$$\phi_\rho(\xi) = \sum_{j=-\infty}^{\infty} c_j e^{-i2\pi\xi j}, \quad -\frac{1}{2} < \xi < \frac{1}{2}, \tag{1.6}$$

provided $\sum_{j=-\infty}^{\infty} |c_j| < \infty$. This periodic function is related to the power spectral density of the underlying continuous time process by the relation

$$\phi_\rho(\xi) = \rho \sum_{l=-\infty}^{\infty} \phi\left(\rho(\xi + l)\right), \quad -\frac{1}{2} < \xi < \frac{1}{2}. \tag{1.7}$$

If $\phi(\cdot)$ has bandwidth f_0 and $\rho \geq 2f_0$, then the above relation simplifies to

$$\phi_\rho(\xi) = \rho\ \phi(\rho\xi), \quad -\frac{1}{2} < \xi < \frac{1}{2},$$

i.e.,

$$\phi(f) = \frac{1}{\rho}\ \phi_\rho\left(\frac{f}{\rho}\right), \quad -\frac{\rho}{2} < f < \frac{\rho}{2}. \tag{1.8}$$

Thus, the power spectral density of the continuous time process can be obtained from that of the sampled process. The corresponding representation of the covariance function $C(\cdot)$ in terms of $\left\{C\left(\frac{j}{\rho}\right), j = \ldots, -2, -1, 0, 1, 2, \ldots\right\}$, which had been noticed by Whittaker (1915) and is currently found in textbooks (Oppenheim and Schafer, 2009), is given by

$$C(t) = \sum_{m=-\infty}^{\infty} C\left(\frac{m}{\rho}\right) \operatorname{sinc}\left(\pi\left(t\rho - m\right)\right), \tag{1.9}$$

where

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

While the reconstruction (1.8) is possible if $\rho \geq 2f_0$, it is not possible if $\rho < 2f_0$ (i.e., if the inter-sample spacing is more than $T = \frac{1}{2f_0}$). This result is known as the Nyquist theorem, and the threshold rate $2f_0$ is known as the Nyquist sampling rate (Shannon, 1949).

When the process X is not bandlimited, or when $\rho < 2f_0$, the shifted versions of $\phi(\cdot)$ appearing on the right hand side of (1.7) overlap with one another, making it impossible to recover $\phi(\cdot)$ from $\phi_\rho(\cdot)$. This type of non-identifiability, where different forms of $\phi(\cdot)$ can lead to the same $\phi_\rho(\cdot)$, is called aliasing.

The sampled process X_τ can also be studied for some other forms of sampling times τ, as we shall see later.

1.2 Spectrum estimation from different types of sampled data

The classical problem of spectrum estimation refers to the estimation of the spectral density $\phi(\cdot)$, or of the spectral distribution $\Phi(\cdot)$. Spectrum estimation has been found to be useful in communication theory (Oppenheim and Schafer, 2009), seismology (Hung, 2002; Costain and Çoruh, 2004), oceanography (Hasselmann et al., 1963), physics (MacDonald and Ness, 1961), signal and image processing (Eldar et al., 1997), analysis of internet traffic data (Roughan, 2006), ecological studies (Matson et al., 1994), medical sciences (French and Holden, 1971) and so on. A finite set of observations of the process, sampled at either uniform or non-uniform time intervals, is generally used for the purpose of spectrum estimation.

Uniform sampling is the most common form of sampling, mainly because it is easy to implement and analyse (Higgins, 1996; Benedetto and Ferreira, 2001). A simple estimator of the spectral density $\phi(\cdot)$, based on n uniformly spaced samples $X(1/\rho), X(2/\rho), \ldots, X(n/\rho)$, is the periodogram defined as

$$I_n(f) = \frac{1}{n\rho}\left|\sum_{j=1}^{n} X\left(\frac{j}{\rho}\right) e^{-\frac{i2\pi fj}{\rho}}\right|^2 1_{[-\rho/2,\rho/2]}(f) = \frac{1}{\rho}\sum_{|j|<n} \widehat{c}_j e^{-\frac{i2\pi jf}{\rho}} 1_{[-\rho/2,\rho/2]}(f), \quad (1.10)$$

where $\widehat{c}_j$ is an estimator of covariance c_j, defined as

$$\widehat{c}_j = \frac{1}{n}\sum_{l=1}^{n-|j|} X\left(\frac{l}{\rho}\right) X\left(\frac{l+|j|}{\rho}\right), \quad j = -(n-1), -(n-2), \ldots, n-2, n-1, \quad (1.11)$$

and $1_E(\cdot)$ is the indicator function of the set E, which takes the value 1 when its argument is in E and the value 0 otherwise. The periodogram is known to be an inconsistent estimator of the spectral density (Brillinger, 2001). When the bandwidth of the process X is smaller than $\rho/2$, the periodogram is asymptotically unbiased, but its variance does not go to zero as the sample size goes to infinity. A common nonparametric spectrum estimator that overcomes this defect is

$$\widehat{\phi}_\rho(f) = \frac{1}{\rho}\sum_{|j|<n} \widehat{c}_j K(b_n j) e^{-\frac{i2\pi fj}{\rho}} 1_{[-\rho/2,\rho/2]}(f), \quad (1.12)$$

where $K(\cdot)$ is a covariance averaging kernel and b_n is the kernel bandwidth. An alternative representation of this estimator is

$$\widehat{\phi}_\rho(f) = \frac{1}{\rho b_n}\int_{-\infty}^{\infty} I_n(\varphi)\kappa\left(\frac{f-\varphi}{\rho b_n}\right) d\varphi, \quad (1.13)$$

where

$$\kappa(f) = \sum_{j=-\infty}^{\infty}\int_{-\infty}^{\infty} K(t) e^{-i2\pi(f+j\rho)t} dt.$$

The convolution with $\kappa(\cdot)$ indicates smoothing of $I_n(\cdot)$. For this reason, the estimator $\widehat{\phi}_\rho(\cdot)$ is known as the smoothed periodogram. If $b_n \to 0$ and $nb_n \to \infty$ as $n \to \infty$, then, assuming that X has bandwidth $f_0 \leq \rho/2$, the estimator $\widehat{\phi}_\rho(\cdot)$ is known to be consistent for the spectral density $\phi(\cdot)$ (Parzen, 1957). However, if $f_0 > \rho/2$, or if X is non-bandlimited, then it follows from the definition of $\widehat{\phi}_\rho(\cdot)$ that this estimator has bias that cannot go to zero, and hence it is inconsistent.

Apart from these nonparametric methods, the literature on spectrum estimation based on uniform sampling contains a large collection of parametric and nonparametric methods, and these have been studied in detail in several books (Pillai and Shim, 1993; Kay, 1999; Naidu, 1999).

Even though uniform sampling is easy to implement, it suffers from the problem of aliasing, as discussed in the foregoing section. It follows from the Nyquist theorem that the class of spectral densities that can be consistently estimated from samples collected at the uniform rate ρ consists of spectral densities having bandwidth smaller than or equal to $\rho/2$. This drawback of uniform sampling has prompted researchers to look for alternative sampling strategies.

The scope of periodic sampling can be expanded somewhat in the case of processes whose spectral support contain sub-intervals of zero spectral density. Prominent examples of such a process are bandpass processes (having spectral support of the form $[-f_2, -f_1] \cup [f_1, f_2]$ for some $f_2 > f_1 > 0$) and multiband processes (having spectral support in the form of union of more than two intervals). It has been shown that for such processes, periodic non-uniform sampling (sampling at irregular intervals that follow a periodic pattern) at sub-Nyquist average rate can lead to appropriate reconstruction of the original process (Landau, 1967). This minimum rate, called the Landau rate, is the Lebesgue measure of the support of the spectrum (Landau, 1967; Marvasti, 2001). This work has been followed by substantial further research in the area, including consistent estimation of the underlying spectrum (Marvasti, 2001; Venkataramani and Bresler, 2001).

While periodic non-uniform sampling nicely exploits the gaps in the spectral support, it does not offer any advantage where such gaps do not exist. For a general process, Shapiro and Silverman (1960) considered alias-free stochastic sampling schemes (τ) in the sense that two continuous time processes with different spectral densities do not produce the same spectral density of the sampled process. The most general form of stochastic sampling that has been considered is one where the sampling times constitute a point process. A special case of this is additive random sampling, where the samples are drawn at the renewal epochs of a renewal process, i.e., the inter-sampling spacings are independent and identically distributed (iid). A further special case is Poisson sampling, where the sampling times coincide with the jumps of a homogeneous Poisson process, and the inter-sample spacings have an exponential distribution. Shapiro and Silverman (1960) proved that additive random sampling through a class of renewal processes, including Poisson sampling, is alias free. This was followed by much research in the area of stochastic sampling. Masry (1978b) proposed an estimator based on Poisson sampled data, similar to the smoothed periodogram defined in (1.12), and proved its consistency under some conditions, without the restriction of any limited bandwidth of the underlying spectral density.

Let τ be the set of jump points of a homogeneous Poisson process with rate β. Consider the estimator $\widehat{\psi}_n(\cdot)$, based on n samples of the process X at $t_1, t_2, \ldots, t_n \in \tau$, defined as

$$\widehat{\psi}_n(f) = \frac{2}{n\beta} \sum_{l=1}^{n-1} \sum_{j=1}^{n-l} X(t_j) X(t_{j+l}) K(b_n(t_{j+l} - t_j)) \cos(2\pi f(t_{j+l} - t_j)) 1_{(-\infty,\infty)}(f), \quad (1.14)$$

where $K(\cdot)$ is a covariance averaging kernel and b_n its bandwidth. Masry (1978b) proved that for any arbitrary average sampling rate β, under some regularity conditions on the process X that does not require X to be bandlimited, this estimator is consistent for the spectral density $\phi(\cdot)$ provided $b_n \to 0$ and $nb_n \to \infty$.

Subsequently, the possibility of breaking free from the nuisance of aliasing with the help of non-uniform sampling enthused many researchers and practitioners. Several non-uniform sampling based methodologies of estimation of the spectral density of a non-bandlimited process have been proposed (Mitchel, 1987; Lehr and Lii, 1997; Tarczynski and Allay, 2004; Stoica and Sandgren, 2006; Stoica et al., 2009). Some of these methods are analogous to methods developed for uniformly sampled data.

The estimation strategies mentioned above are meant for situations where one can design the sampling epochs. Examples of such applications include internet traffic data (Roughan, 2006), seismology (Hung, 2002; Costain and Çoruh, 2004), image processing (Eldar et al., 1997) and so on. However, irregularly spaced data can also occur naturally in many practical situations like seismic studies (Ozbek and Ferber, 2005), turbulent velocity fluctuation (Tummers and Passchier, 2000), laser doppler anemometer studies (Nobach et al., 1998), wide-band antenna arrays (Ishimaru and Chen, 1965), computer aided tomography, spotlight-mode synthetic aperture radar (Munson et al., 1983) and so on. For spectrum estimation from such data, there have been attempts to use standard methods based on uniform sampling, after suitable weighting (Bronez, 1988) or interpolation (Tummers and Passchier, 2000) of the data.

1.3 Matters investigated in this thesis

Whenever one has control over the sampling mechanism, selection of the sampling scheme (uniform or non-uniform) is a serious issue in relation to the problem of spectrum estimation. Two major arguments in favour of uniform sampling are: (a) it is logistically easy to implement, and (b) it provides the basis for a vast collection of estimators, most of which are computationally simple and theoretically understood well. On the other hand, the possibility of aliasing and the proven inconsistency of spectral estimators based on uniform sampling are arguments against it. In this thesis, we reexamine the strengths and weaknesses of uniform and stochastic sampling through new theoretical results as well as a series of carefully designed simulation studies.

We begin with the question of consistency of the estimator (1.12) based on uniform sampling. The consistency of any estimator concerns its behaviour as the sample size goes to infinity. However, it does not make practical sense to let the sample size tend to infinity while keeping the uniform sampling rate fixed. If one gathers more and more resources to increase the sample size, one can use some of those resources to sample faster. In this thesis, we examine whether the estimator (1.12) can be consistent for a non-bandlimited spectral density *when the sampling rate increases suitably as the sample size goes to infinity*, how it performs in comparison to the estimator (1.14) based on Poisson sampled data, and whether one can use this estimator to construct asymptotic confidence intervals that shrink to the true spectral density.

Even if one has control over the sampling mechanism, there may be a practical constraint on the minimum separation between successive samples. Let d be the minimum allowable separation between successive samples. Under this constraint, the fastest possible rate of uniform sampling is $1/d$. It has already been observed that sampling at this rate is alias-free only for the class of spectral densities that are restricted to the bandwidth $1/2d$. Interestingly, non-uniform sampling schemes also pose difficulties in this situation. Many stochastic sampling schemes, including Poisson sampling, are not even feasible in the presence of this constraint, and some implementable schemes are no longer alias-free. In this thesis, we investigate whether there exists a stochastic sampling scheme which is alias-free for the class of all spectra, and if so, how one can use it to estimate the power spectral density consistently.

In this thesis, we have dealt exclusively with continuous time, mean square continuous stationary stochastic process with zero mean. Processes with a non-zero mean are usually handled in the same manner as zero mean processes after centering, and this adjustment generally does not pose any technical difficulty. We do not attempt to work with non-stationary processes. Since such processes are often modeled in terms of stationary processes (see for example, Priestley (1983)), it may be expected that clearer understanding and enrichment of the latter area would be beneficial for the former.

1.4 Organization of the thesis

In Chapter 2, it is shown that under certain regularity conditions on the underlying process, the smoothing mechanism and the sampling rate, the estimator given by (1.12) is consistent for the spectral density, even when the latter is not bandlimited. These results necessitate new asymptotic arguments, since the expressions for asymptotic bias and variance of the estimators for fixed sampling rate, given in textbooks (Brillinger, 2001), become inappropriate when the sampling rate is allowed to go to infinity together with the sample size. The consistency of the estimator is proved under a set of assumptions that are similar to those used to establish the consistency of the estimator (1.14) based on Poisson sampled data. Under additional conditions on the smoothness and the decay of the underlying spectral density, the optimal rates of convergence of the estimators based on uniformly sampled and Poisson-sampled data are compared. Applicability of the theoretical results to finite sample sizes is examined through Monte-Carlo simulations.

In Chapter 3, we extend the above asymptotic approach to the case of multivariate non-bandlimited stationary processes. After establishing consistency of the vector of smoothed periodogram estimators of the spectral and cross-spectral densities, we show that, with appropriate scaling, the limiting distribution of this estimator is multivariate normal. The rate of convergence is studied and compared with the rate for mean square consistency. The results are used to construct asymptotic confidence intervals for spectral and cross spectral densities. Appropriateness of the theoretical results and empirical coverage probabilities at moderate sample sizes are studied through Monte-Carlo simulations.

In Chapter 4, we consider stochastic sampling schemes that are alias-free, under the constraint of a minimum separation between successive samples. This study is based on two definitions of alias-free sampling given by Shapiro and Silverman (1960) (formalized by Beutler (1970)) and Masry (1978a), respectively. We show that, subject to this constraint, *no point process sampling scheme is alias-free for the class of all spectra* under either definition. Subsequently, we restrict attention to the class of spectral densities having arbitrary but finite spectral support, and look for alias-free sampling schemes under both definitions.

In Chapter 5, we build on the positive answers emerging from the investigations of Chapter 4, and look for a consistent estimator of a power spectral density that has arbitrary but finite spectral support, using appropriately sampled data under the constraint of a minimum separation between successive samples. It emerges that, the approaches previously considered in the absence of this constraint would not work in the present case. Subsequently, we develop a consistent estimator, and study its small sample performance through Monte Carlo simulations.

In Chapter 6, we explain how the findings of this thesis improve the understanding of the relative merits of uniform and stochastic sampling. For the sake of completion of this discussion, we also compile and collate some previously known results. This summary helps one to rationalize one's expectations from uniform sampling, Poisson sampling and other stochastic sampling schemes. We also outline a few potential areas of further research.

Strategies for simulation from different continuous time stochastic processes, which were found to be useful for the simulation results reported in this thesis, are given in the appendix.

The contents of Chapters 2–5 are based on Srivastava and Sengupta (2010), Srivastava and Sengupta (2011a), Srivastava and Sengupta (2011c) and Srivastava and Sengupta (2011b), respectively.

Chapter 2

Uniformly sampled non-bandlimited processes: Consistent estimation of spectral density

2.1 Introduction

A common nonparametric spectrum estimator, based on n uniformly spaced samples at rate ρ, is the smoothed periodogram defined in (1.12) as follows.

$$\widehat{\phi}_\rho(f) = \frac{1}{\rho} \sum_{|j|<n} \widehat{c}_j K(b_n j) e^{-\frac{i2\pi f j}{\rho}} 1_{[-\rho/2,\rho/2]}(f),$$

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

$$\widehat{c}_j = \frac{1}{n} \sum_{l=1}^{n-|j|} X\left(\frac{l}{\rho}\right) X\left(\frac{l+|j|}{\rho}\right), \quad j = -(n-1), -(n-2), \ldots, n-2, n-1,$$

as defined in (1.11). If $b_n \to 0$ and $nb_n \to \infty$ as $n \to \infty$, then, assuming that X has bandwidth $f_0 \leq \rho/2$, the bias and the variance of the estimator $\widehat{\phi}_\rho(\cdot)$ go to zero as n goes to infinity, implying that the estimator is consistent for the spectral density $\phi(\cdot)$ (Parzen, 1957). When the process X is non-bandlimited, the variance continues to go to zero, but the bias does not, as we have seen in Section 1.2. Thus, the estimator is inconsistent whenever X is a non-bandlimited process.

As mentioned in Section 1.2, stochastic sampling can lead to a consistent spectrum estimator, such as (1.14), even for a non-bandlimited process. One might question whether, for finite sample size, the estimation error of (1.14) is substantially smaller than that of the estimator (1.12). An empirical study by Roughan (2006) in the special case of active measurements for network performance produced mixed results, which led the author to conclude that, while spectral estimators based on Poisson sampling have less efficiency (i.e., high variance), such techniques could be used to detect periodicities in the system, and to determine which rate of uniform sampling would be inadequate. Moreover, it has been observed that generally, spectral estimators based on non-uniformly sampled data have higher variances than those based on uniformly sampled data (Roberts and Gaster, 1980; Moore et al., 2008). In the absence of a comprehensive empirical study, it is not possible to claim superiority of estimators based on one type of sampling over the other. However, it appears that uniform sampling is sometimes avoided because of the stigma of inconsistency attached to spectral estimators based on uniformly sampled data (Wolf et al., 2007).

Consistency of an estimator is a large sample property of an estimator. It is reasonable to expect that, as one has more and more resources to gather information, the precision as well as the accuracy of the estimator should improve. However, the estimator defined in (1.12) is designed to be biased, no matter how large the sample size is. Upon closer examination though, it transpires that the phenomenon of non-diminishing bias is largely due to a peculiarity of the asymptotic arguments, and not so much an inherent limitation of the estimator. Specifically, if one has the resources to increase the sample size, there is no reason why one should not use some of these resources to sample faster, so that better justice can be done to higher frequencies. Realizing this fact, practitioners fix the intended range of spectrum estimation, and then sample an appropriately filtered process at a sufficiently high frequency to avoid aliasing. Sometimes one goes for successively higher rates of uniform sampling to determine an appropriate rate of sampling (Eldar et al., 1997). However, these common sense approaches are yet to be backed up by appropriate asymptotic calculations. There is a need to bridge this gap by working out the large sample properties of estimators *when the sampling rate increases suitably as the sample size goes to infinity.*

It may be noted that this asymptotic approach had been adopted by other authors (Constantine and Hall, 1994; Hall et al., 1994; Lahiri, 1999) in other time series problems, and was referred to by Fuentes (2002) as 'shrinking asymptotics'. This approach is different from the 'fixed-domain asymptotics' or 'infill asymptotics' approach (Chen et al., 2000; Stein, 1995; Zhang and Zimmerman, 2005; Lim and Stein, 2008) 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.

Without assuming that the underlying process is bandlimited, we examine the asymptotic properties of the estimator given in (1.12) by letting ρ go to infinity at an appropriate rate, as n goes to infinity. In the sequel, we shall use the notation ρ_n instead of ρ, in order to explicitly indicate the dependence of the sampling rate on the sample size. The estimator of (1.12) is accordingly denoted as

$$\widehat{\phi}_n(f) = \frac{1}{\rho_n} \sum_{|j|<n} \widehat{c}_{jn} K(b_n j) e^{-\frac{i2\pi f j}{\rho_n}} 1_{[-\rho_n/2,\ \rho_n/2]}(f), \tag{2.1}$$

where

$$\widehat{c}_{jn} = \frac{1}{n} \sum_{l=1}^{n-|j|} X\left(\frac{l}{\rho_n}\right) X\left(\frac{l+|j|}{\rho_n}\right), \quad j = -(n-1), -(n-2), \ldots, n-2, n-1, \tag{2.2}$$

and the alternative representation given in (1.13) is written as

$$\widehat{\phi}_n(f) = \frac{1}{\rho_n b_n} \int_{-\infty}^{\infty} I_n(\varphi) \kappa_n\left(\frac{f-\varphi}{\rho_n b_n}\right) d\varphi, \tag{2.3}$$

where

$$\begin{aligned} I_n(f) &= \frac{1}{n\rho_n} \left| \sum_{j=1}^{n} X\left(\frac{j}{\rho_n}\right) e^{-\frac{i2\pi f j}{\rho_n}} \right|^2 1_{[-\rho_n/2,\rho_n/2]}(f), \\ \kappa_n(f) &= \sum_{j=-\infty}^{\infty} \int_{-\infty}^{\infty} K(t) e^{-i2\pi(f+j\rho_n)t} dt. \end{aligned}$$

The degree of smoothness of the smoothed periodogram, $\widehat{\phi}_n(\cdot)$, is controlled by the parameter $\rho_n b_n$ of the frequency domain window $\kappa_n(\cdot)$.

In Section 2.2, we prove the consistency of the estimator $\widehat{\phi}_n(\cdot)$ under some general conditions. In Section 2.3, we calculate the rate of convergence of the bias and the variance of this estimator, and determine the optimal rates at which ρ_n and nb_n should go to infinity so that the mean squared error (MSE) has the fastest possible rate of convergence. Subsequently, we compare the rates of convergence of the bias and the variance of this estimator with those of the estimator (1.14), based on Poisson sampling. We provide proofs of all the theoretical results in section 2.4. We present the results of a simulation study in Section 2.5 and provide some concluding remarks in Section 2.6.

2.2 Consistency of the smoothed periodogram estimator

Consider the mean square continuous, wide sense stationary stochastic process $\{X(t), -\infty < t < \infty\}$ with zero mean, (auto-)covariance function $C(\cdot)$ and spectral density $\phi(\cdot)$. In order to prove that the estimator (2.1) is consistent, it is sufficient to show that the bias and the variance of the estimator tend to zero as the sample size (n) tends to infinity.

We assume the following condition on the covariance function $C(\cdot)$.

Assumption 2.1. *The function $h_0(\cdot)$, defined over the real line as $h_0(t) = \sup_{s\geq|t|} |C(s)|$ is integrable.*

Remark 2.1. Assumption 2.1 is equivalent to saying that the covariance function $C(\cdot)$ is bounded over $[0,\infty)$ by a non-negative, non-increasing and integrable function.

We assume the following conditions on the choice of the kernel $K(\cdot)$, the kernel bandwidth b_n and the sampling rate ρ_n.

Assumption 2.2. *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 2.3. *The kernel bandwidth is such that $nb_n \to \infty$ as $n \to \infty$.*

Assumption 2.4. *The sampling rate is such that $\rho_n \to \infty$ and $\rho_n b_n \to 0$ as $n \to \infty$.*

Note that Assumption 2.4 implies that $b_n \to 0$ as $n \to \infty$.

Remark 2.2. Assumption 2.4 says that the smoothing parameter of smoothed periodogram (see (2.3)) goes to zero, and the sampling rate goes to infinity, as the sample size goes to infinity.

Theorem 2.1. *Under Assumptions 2.1–2.4, the bias of the estimator $\widehat{\phi}_n(\cdot)$ given by (2.1) tends to zero uniformly over any closed and finite interval.*

Before examining the variance of the estimator we assume a set of conditions on some fourth order moments/cumulants of finite dimensional distributions of the process X. An s-th order cumulant of the random vector $(Y_1, \ldots, Y_r)$, corresponding to the index set $\{j_1, \ldots, j_s\}$ for $1 \leq j_1, \ldots, j_s \leq r$, is given by

$$cum(Y_{j_1}, \ldots, Y_{j_s}) = \sum_{\boldsymbol{\nu}} (-1)^{\vartheta-1}(\vartheta-1)! \left(E \prod_{j \in \ \nu_1} Y_j \right) \times \cdots \times \left(E \prod_{j \in \ \nu_\vartheta} Y_j \right), \quad (2.4)$$

where the summation is over all partitions $\boldsymbol{\nu} = (\nu_{j_1}, \ldots, \nu_{j_\vartheta})$ of size $\vartheta = 1, \ldots, s$, of the index set $\{j_1, j_2, \ldots, j_s\}$. The relation between moments and cumulants up to any finite order is well known (Brillinger, 2001).

Assumption 2.5. *The fourth moment $E|X(t)|^4$ exists for every t, and the fourth order cumulant function $cum\,[X(t+t_1), X(t+t_2), X(t+t_3), X(t)]$ 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 $[0, \infty)$.

Remark 2.3. The cumulant $Q(t_1, t_2, t_3)$ can be written as

$$Q(t_1, t_2, t_3) = P(t_1, t_2, t_3) - P_G(t_1, t_2, t_3),$$

where

$$\begin{aligned} P(t_1, t_2, t_3) &= E\{X(t)X(t+t_1)X(t+t_2)X(t+t_3)\}, \\ P_G(t_1, t_2, t_3) &= C(t_1)C(t_2-t_1) + C(t_2)C(t_3-t_1) + C(t_3)C(t_1-t_2). \end{aligned}$$

Assumption 2.5 is satisfied by a Gaussian process, as the function $P(\cdot)$ reduces to the function $P_G(\cdot)$.

Theorem 2.2. *Under Assumptions 2.1–2.5, the variance of the estimator $\widehat{\phi}_n(\cdot)$ given by (2.1) converges as follows:*

$$\lim_{n\to\infty} nb_n Var[\widehat{\phi}_n(f)] = \left(1 + 1_{\{0\}}(f)\right) [\phi(f)]^2 \int_{-\infty}^{\infty} K^2(x)dx.$$

The convergence is uniform over any closed and finite interval that does not include the frequency 0. In particular, the variance converges to 0.

It follows from Theorems 2.1 and 2.2 that, under Assumptions 2.1–2.5, the estimator $\widehat{\phi}_n(f)$ is consistent, and is uniformly consistent over any closed and finite frequency interval that does not include the point 0.

2.3 Rate of convergence

The rate of convergence of the variance of $\widehat{\phi}_n(f)$ follows from Theorem 2.2. We assume a few further conditions in order to arrive at a rate of convergence for its bias. These include additional conditions on the shapes of the covariance function and the kernel function.

Assumption 2.1A. *The function $h_q(\cdot)$, defined over the real line as $h_q(t) = \sup_{s\geq|t|} |s|^q|C(s)|$ is integrable, for some positive number q greater than 1.*

Assumption 2.1B. *The spectral density is such that, for some $p > 1$, $\lim\limits_{f\to\infty} |f|^p\phi(f) = A$ for some positive number A.*

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

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

for each positive number r such that the limit exists. The characteristic exponent of the kernel is defined as the largest number r, such that the limit exists and is non-zero

(Parzen, 1957). In other words, the characteristic exponent is the number r such that $1-K(1/y)$ is $O(y^{-r})$.

Assumption 2.2A. *The characteristic exponent of the kernel $K(\cdot)$ is a number, for which Assumption 2.1A holds.*

Remark 2.4. Assumption 2.1A implies Assumption 2.1 (see Remark 2.1), and also that $\phi(\cdot)$ 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 2.1A holds for a particular value of q, then it would also hold for smaller values.

Remark 2.5. The number p indicates the rate of decay of the spectral density. The following are two interesting situations, where Assumption 2.1B holds.

1. The spectral density $\phi(\cdot)$ is a rational function, i.e., $\phi(f) = \frac{P(f)}{Q(f)}$, 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 $C(\cdot)$ has the following smoothness property: $C(\cdot)$ is p times differentiable and the p^{th} derivative of $C(\cdot)$ is in L^1.

Remark 2.6. The number p 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.

Theorem 2.3. *Under Assumption 2.2–2.4, 1A, 1B and 2A, the bias of the estimator $\widehat{\phi}_n(f)$ given by (2.1) is*

$$\begin{aligned}
E[\widehat{\phi}_n(f) - \phi(f)] = & \left[-k_q \int_{-\infty}^{\infty} |t|^q C(t) e^{-i2\pi f t} dt\right] (\rho_n b_n)^q + o\left((\rho_n b_n)^q\right) \\
& + \left[-\int_{-\infty}^{\infty} |t| C(t) e^{-i2\pi f t} dt\right] \left(\frac{\rho_n}{n}\right) + o\left(\frac{\rho_n}{n}\right) \\
& + \left[A \sum_{|l|>0} \frac{1}{|l|^p}\right] \frac{1}{(\rho_n)^p} + o\left(\frac{1}{(\rho_n)^p}\right),
\end{aligned}$$

i.e.,

$$E[\widehat{\phi}_n(f) - \phi(f)] = O\left((\rho_n b_n)^q\right) + O\left(\frac{\rho_n}{n}\right) + O\left(\frac{1}{\rho_n^p}\right),$$

uniformly in f over any closed and finite interval.

Remark 2.7. Assumption 2.2A can be relaxed to the extent that the characteristic exponent of the kernel $K(\cdot)$ is required to be greater than or equal to the number q, for which Assumption 2.1A is assumed to hold. If it is strictly greater than q, then the term $O\left((\rho_n b_n)^q\right)$ in the above theorem would have to be replaced by $o\left((\rho_n b_n)^q\right)$. This follows from equation (2.33) in the Section 2.4 and the fact that $k_q = 0$ in this case. On the other hand, if a kernel with characteristic exponent less than q is used, then one does not fully utilize the strength of the assumption on the smoothness of the spectral density, implied by Assumption 2.1A, and hence gets a slower rate of convergence.

2.3.1 Choice of sampling rate and kernel bandwidth

From Theorem 2.3 and Theorem 2.2, it is observed that the bias and the variance of the estimator $\widehat{\phi}_n(f)$ converge to zero at different rates. We set out to choose the sampling rate ρ_n and the bandwidth b_n in order to ensure that the MSE of $\widehat{\phi}_n(f)$ converges to zero as fast as possible. It would turn out that this happens when the squared bias and the variance go to zero at the same rate.

Theorem 2.4. *Under Assumptions 2.2–2.5, 2.1A, 2.1B and 2.2A, the optimal rate of convergence of the MSE of the estimator $\widehat{\phi}_n(f)$ is given by*

$$MSE[\widehat{\phi}_n(f)] = O\left(n^{-\frac{2pq}{p+q+2pq}}\right), \tag{2.5}$$

which corresponds to the optimal choices

$$\rho_n = P\, n^{\frac{q}{p+q+2pq}} + o\left(n^{\frac{q}{p+q+2pq}}\right), \tag{2.6}$$

$$\text{and } b_n = Q\, n^{-\frac{p+q}{p+q+2pq}} + o\left(n^{\frac{p+q}{p+q+2pq}}\right), \tag{2.7}$$

where P and Q are positive constants.

The above optimal rates of ρ_n and b_n lead to the following corollaries to Theorems 2.2 and 2.3, respectively.

Corollary 2.1. *Under the Assumptions 2.1, 2.2, 2.5 and the choices of ρ_n and b_n given by (2.6–2.7), we have*

$$\lim_{n\to\infty} n^{\frac{2pq}{p+q+2pq}} Var[\widehat{\phi}(f)] = \frac{1}{Q}\left(1 + 1_{\{0\}}(f)\right)[\phi(f)]^2 \int_{-\infty}^{\infty} K^2(x)dx. \tag{2.8}$$

Corollary 2.2. *Under the Assumptions 2.1A, 2.1B, 2.2, 2.2A and the choices of ρ_n and b_n given by (2.6–2.7), we have*

$$\lim_{n\to\infty} n^{\frac{pq}{p+q+2pq}} E[\widehat{\phi}_n(f) - \phi(f)] = -(PQ)^q k_q \int_{-\infty}^{\infty} |t|^q C(t) e^{-i2\pi f t} dt + \frac{1}{P^p} A \sum_{|l|>0} \frac{1}{|l|^p}, \tag{2.9}$$

where the constant A is as in Assumption 2.1B.

2.3.2 Comparison with Poisson sampling estimator

Among the various schemes for sampling a continuous time stochastic process at irregular intervals, Poisson sampling proposed by Silverman (Shapiro and Silverman, 1960) is the simplest and most popular. Here, we compare the asymptotic behaviour of the estimator $\widehat{\phi}_n(\cdot)$ with the corresponding estimator based on Poisson sampling.

Let $\{t_j\}_{j=0}^n$ be the sampling points from a Poisson process with average rate β. Masry (1978b) proved that, under Assumptions 2.1, 2.2, 2.3 and 2.5, the estimator $\widehat{\psi}_n(\cdot)$ defined in (1.14) as

$$\widehat{\psi}_n(f) = \frac{2}{\beta n} \sum_{l=1}^{n-1} \sum_{j=1}^{n-l} X(t_j) X(t_{j+l}) K(b_n(t_{j+l} - t_j)) \cos(2\pi f(t_{j+l} - t_j)),$$

is consistent for $\phi(f)$ for any choice of β.

Under the above Assumptions, the asymptotic variance of $\widehat{\psi}_n(f)$ satisfies

$$\lim_{n\to\infty} (nb_n) Var[\widehat{\psi}_n(f)] = \beta \left[\phi(f) + \frac{C(0)}{\beta}\right]^2 \left(1 + 1_{\{0\}}(f)\right) \int_{-\infty}^{\infty} K^2(t)dt. \tag{2.10}$$

For specifying the rate of convergence of the bias, Masry (1978b) assumed the following additional conditions.

Assumption 2.1C. *The function $|t|^q C(t)$ is integrable for some positive integer q.*

Assumption 2.2B. *The covariance averaging kernel $K(\cdot)$ is q times differentiable with* *bounded derivatives, where q is an integer for which Assumption 2.1C holds.*

Note that Assumption 2.1C is implied by Assumption 2.1A with the same or higher value of q as is used here. Masry (1978b) showed that, under Assumptions 2.1, 2.2, 2.3, 2.1C and 2.2B, the bias of the estimator $\widehat{\psi}_n(f)$ is given as

$$\begin{aligned} Bias[\widehat{\psi}_n(f)] =& E[\widehat{\psi}_n(f)] - \phi(f) \\ =& \sum_{l=1}^{q-1} \frac{(i)^l K^{(l)}(0) b_n^l}{l!} \phi^{(l)}(f) + O(b_n^q) + O\left(\frac{1}{n}\right). \end{aligned} \tag{2.11}$$

It follows from (2.11) that the bias of $\widehat{\psi}_n(f)$ is $O\left(\max\left\{b_n^m, n^{-1}\right\}\right)$, where

$$m = \begin{cases} q & \text{if } K^{(l)}(0) = 0 \text{ for } 1 \le l < q, \\ l_0 & \text{if } K^{(l_0)}(0) \neq 0 \text{ and } K^{(l)}(0) = 0 \text{ for } 1 \le l < l_0 < q. \end{cases}$$

The fastest possible rate of convergence is $O\left(\max\left\{b_n^q, n^{-1}\right\}\right)$, and this is achieved when one uses a kernel, which further satisfies Assumption 2.2A with the same or higher value of q as is used here. In such a case, we have

$$Bias[\widehat{\psi}_n(f)] = O(b_n^q) + O\left(\frac{1}{n}\right). \tag{2.12}$$

If the Assumption 2.2A holds with a higher value of q, then the term $O(b_n^q)$ has to be replaced by $o(b_n^q)$

Let us now assume that the kernel is chosen appropriately to ensure (2.12). Note that the bias and the variance of $\widehat{\psi}_n(f)$ converge to zero at different rates. One can choose the rate of convergence of the bandwidth b_n such that the MSE converges as fast as possible. It turns out that the fastest convergence of the MSE happens when $b_n = O(n^{-\frac{1}{2q+1}})$, in which case the squared bias and variance of $\widehat{\psi}_n(f)$ are both $O(n^{-\frac{2q}{2q+1}})$.

In summary, under Assumptions 2.1, 2.1C, 2.2, 2.2A, 2.2B, 2.5 and

$$b_n = R\, n^{-\frac{1}{2q+1}} + o(n^{-\frac{1}{2q+1}}), \tag{2.13}$$

the MSE of $\widehat{\psi}_n(f)$ is

$$MSE[\widehat{\psi}_n(f)] = O\left(n^{-\frac{2q}{2q+1}}\right).$$

The rate of convergence of MSE of $\widehat{\phi}_n(f)$ is given in Theorem 2.4 under Assumptions 2.1A, 2.1B, 2.2, 2.2A, 2.5 and (2.6–2.7). Both the results hold when ρ_n and b_n for $\widehat{\phi}_n(\cdot)$ are chosen as in (2.6–2.7), b_n for $\widehat{\psi}_n(\cdot)$ is chosen as in (2.13) and the following conditions hold simultaneously: Assumption 2.1A for some q greater than 1 (which implies Assumption 2.1C for $[q]$ and Assumption 2.1), Assumption 2.2, Assumption 2.2A (for the same q as in Assumption 2.1A), Assumption 2.2B for $[q]$, and Assumption 2.5. Under this common set of conditions, we have

$$MSE[\widehat{\psi}_n(f)] = O\left(n^{-\frac{2[q]}{2[q]+1}}\right),$$
$$MSE[\widehat{\phi}_n(f)] = O\left(n^{-\frac{2q}{2q+1+q/p}}\right).$$

When q is an integer, i.e., $[q] = q$, the rate of convergence of the MSE of $\widehat{\phi}_n(f)$ is slower than that of $\widehat{\psi}_n(f)$. The two rates are comparable if p is much larger than q. When q is not an integer, the MSE of $\widehat{\phi}_n(f)$ converges faster when $p > q[q]/(q - [q])$, and in particular when p is very large. As we have indicated in Remark 2.6, for every fixed q, one can make p suitably large through low pass filtering.

If the rates of convergence are comparable, the constants associated with these rates become important. We will compare the constants of the asymptotic bias of $\widehat{\phi}_n(f)$ and $\widehat{\psi}_n(f)$ as well as the constants of their asymptotic variance separately, assuming that q is an integer.

Under Assumptions 2.1, 2.2, 2.5 and (2.13), we have

$$\lim_{n\to\infty} n^{\frac{2q}{2q+1}} Var[\widehat{\psi}_n(f)] = \frac{1}{R}\beta\left[\phi(f) + \frac{C(0)}{\beta}\right]^2 \left(1 + 1_{\{0\}}(f)\right)\int_{-\infty}^{\infty} K^2(x)dx. \tag{2.14}$$

On the other hand, we have from Corollary 2.1 that under the Assumptions 2.1, 2.2, 2.5 and (2.6–2.7),

$$\lim_{n\to\infty} n^{\frac{2q}{2q+1+q/p}} Var[\widehat{\phi}_n(f)] = \frac{1}{Q}\left(1 + 1_{\{0\}}(f)\right)[\phi(f)]^2\int_{-\infty}^{\infty} K^2(x)dx.$$

The ratio of the constants for the asymptotic variances of $\widehat{\psi}_n(f)$ and $\widehat{\phi}_n(f)$ is

$$\frac{Q}{R}\beta\left[1 + \frac{C(0)}{\beta\phi(f)}\right]^2. \tag{2.15}$$

This ratio depends on the Poisson sampling rate β and the true value of the power spectral density $\phi(f)$. This ratio can be much larger than 1, particularly for larger values of f. In fact, even if β is chosen to minimize this ratio for a given value of $\phi(f)$ (though this is not practically possible), the minimum value happens to be $2QC(0)/[R\phi(f)]$, which can be arbitrarily large for large values of f. Thus, the variance of $\widehat{\psi}_n(f)$ can generally be expected to be larger than that of $\widehat{\phi}_n(f)$.

We now turn to the comparison of the expressions for bias. Under Assumptions 2.1C, 2.2, 2.2A and 2.2B along with (2.13), we have

$$\lim_{n\to\infty} n^{\frac{q}{2q+1}}[E[\widehat{\psi}_n(f)] - \phi(f)] = -R^q k_q \int_{-\infty}^{\infty} |t|^q C(t) e^{-i2\pi f t} dt. \tag{2.16}$$

On the other hand, we have from Corollary 2.2 that under the Assumptions 2.1A, 2.1B, 2.2, 2.2A and (2.6–2.7),

$$\lim_{n\to\infty} n^{\frac{pq}{p+q+2pq}} E[\widehat{\phi}_n(f) - \phi(f)] = -(PQ)^q k_q \int_{-\infty}^{\infty} |t|^q C(t) e^{-i2\pi f t} dt + \frac{1}{P^p} A \sum_{|l|>0} \frac{1}{|l|^p}.$$

The first term of the expression on the right hand side is proportional to the expression on the right hand side of (2.16). These terms are small for large values of f, while the second term of the expression on the right hand side of the above inequality does not depend on f. Even if the value of the second term is small, it would make a difference for large values of f. Consequently, $\widehat{\psi}_n(f)$ can generally be expected to have a smaller bias than $\widehat{\phi}_n(f)$.

In summary, even though both $\widehat{\phi}_n(f)$ and $\widehat{\psi}_n(f)$ are consistent estimators under the stated conditions, there is a trade-off between $\widehat{\phi}_n(f)$ and $\widehat{\psi}_n(f)$ in terms of bias and variance. There is no clear order between the constants of the MSE's of the two estimators.

In order to examine the validity of the asymptotic results and the above comparisons for small samples, we turn to Monte Carlo simulations, reported in Section 2.5.

2.4 Proofs

We denote by $K_1(\cdot)$ a function that bounds the covariance averaging kernel $K(\cdot)$ as in Assumption 2.2. Further, we denote $K_1(0)$ by M.

Proof of Theorem 2.1. We shall show that the bias of the estimator $\widehat{\phi}_n(f)$ given by (2.1) converges to 0 uniformly over $[f_l, f_u]$ for any f_l, f_u such that $f_l < f_u$. In order to compute the bias, we evaluate $E[\widehat{c}_{vn}]$:

$$E[\widehat{c}_{vn}] = E\left[\frac{1}{n}\sum_{j=1}^{n-|v|} X\left(\frac{j}{\rho_n}\right) X\left(\frac{j+|v|}{\rho_n}\right)\right] = \left(1-\frac{|v|}{n}\right) C\left(\frac{v}{\rho_n}\right). \tag{2.17}$$

Therefore, we have

$$E[\widehat{\phi}_n(f)] = \frac{1}{\rho_n}\sum_{|v|<n}\left(1-\frac{|v|}{n}\right) C\left(\frac{v}{\rho_n}\right) K(b_n v) e^{\frac{-i2\pi f v}{\rho_n}} 1_{[-\rho_n/2,\rho_n/2]}(f).$$

Consider the simple function $S_n(\cdot)$, defined over $[f_l, f_u] \times (-\infty, \infty)$, by

$$S_n(f,t) = \sum_{|v|<n}\left(1-\frac{|v|}{n}\right) C\left(\frac{v}{\rho_n}\right) K\left(b_n v\right) e^{\frac{-i2\pi f v}{\rho_n}} 1_{[-\rho_n/2,\rho_n/2]}(f) 1_{\left(\frac{v-1}{\rho_n},\frac{v}{\rho_n}\right]}(t).$$

Observe that $\int_{-\infty}^{\infty} S_n(f,t)dt = E[\widehat{\phi}_n(f)]$. Define the function $S(\cdot)$, over $[f_l, f_u] \times (-\infty, \infty)$, by

$$S(f,t) = C(t)e^{-i2\pi f t}.$$

Observe that $\int_{-\infty}^{\infty} S(f,t)dt = \phi(f)$ which is continuous.

For any $t \in (-\infty, \infty)$, let $v_n(t)$ be the smallest integer greater than or equal to $\rho_n t$. Note that the interval $(\frac{v_{n-1}(t)}{\rho_n}, \frac{v_n(t)}{\rho_n}]$ contains the point t and $\lim_{n\to\infty} \frac{v_n(t)}{\rho_n} = t$. For sufficiently large n, we have from Assumptions 2.3 and 2.4,

$$S_n(f,t) = \left(1-\frac{|v_n(t)|}{\rho_n}\cdot\frac{\rho_n}{n}\right) C\left(\frac{v_n(t)}{\rho_n}\right) K\left(b_n\rho_n\frac{v_n(t)}{\rho_n}\right) e^{\frac{-i2\pi f v_n(t)}{\rho_n}} 1_{[-\rho_n/2,\rho_n/2]}(f).$$

Proving the uniform convergence of $Bias[\widehat{\phi}_n(f)]$ over finite interval $[f_l,\ f_u]$ amounts to proving

$$\lim_{n\to\infty}\int_{-\infty}^{\infty} S_n(f,t)dt = \int_{-\infty}^{\infty} S(f,t)dt,$$

uniformly over $[f_l,\ f_u]$. By virtue of the continuity of the limiting function, this in turn is equivalent to proving that $\int_{-\infty}^{\infty} S_n(f,t)dt$ converges continuously over this interval (Resnick, 1987), i.e., for any sequence of frequencies $f_n \to f$,

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} S_n(f_n,t)dt = \int_{-\infty}^{\infty} S(f,t)dt,$$

where $f_n, f \in [f_l,\ f_u]$.

By continuity of the function $S_n(f,t)$ with respect to t and f, we have from Assumptions 2.3 and 2.4, for any fixed t,

$$\lim_{n\to\infty} |S_n(f_n,t) - S(f,t)| = 0.$$

Note that from Assumptions 2.1 and 2.2, we have the dominance

$$|S_n(f_n,t)| \leq M \sum_{|v|<n} \left| C\left(\frac{v}{\rho_n}\right)\right| 1_{\left(\frac{v-1}{\rho_n},\frac{v}{\rho_n}\right]}(t) \leq M h_0(t).$$

where $h_0(\cdot)$ is the function described in Assumption 2.1. Thus, by applying the dominated convergence theorem (DCT), we have

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} S_n(f_n,t)dt = \int_{-\infty}^{\infty} S(f,t)dt.$$

Hence, $E[\widehat{\phi}_n(f)] \to \phi(f)$ uniformly on $[f_l,\ f_u]$. □

Proof of Theorem 2.2. The estimator $\widehat{\phi}_n(f)$, given by (2.1), can be written as

$$\widehat{\phi}_n(f) = \frac{1}{\rho_n}\widehat{c}_{0n} + \frac{2}{\rho_n}\sum_{j=1}^{n-1} \widehat{c}_{jn} K(b_n j) \cos\left(\frac{2\pi f j}{\rho_n}\right).$$

Therefore,

$$Var[\widehat{\phi}_n(f)] = T_1 + 2T_2(f) + T_3(f), \tag{2.18}$$

where

$$T_1 = \frac{1}{\rho_n^2} Var[\widehat{c}_{0n}],$$

$$
\begin{aligned}
T_2(f) &= Cov\left[\frac{1}{\rho_n}\widehat{c}_{0n}, \frac{2}{\rho_n}\sum_{j=1}^{n-1}\widehat{c}_{jn}K(b_nj)\cos\left(\frac{2\pi fj}{\rho_n}\right)\right], \\
T_3(f) &= Var\left[\frac{2}{\rho_n}\sum_{j=1}^{n-1}\widehat{c}_{jn}K(b_nj)\cos\left(\frac{2\pi fj}{\rho_n}\right)\right].
\end{aligned}
$$

Before we consider the convergence of the above three terms, we simplify the computation of $Cov(\widehat{c}_{j_1n}, \widehat{c}_{j_2n})$ for non negative j_1 and j_2.

Note from Assumption 2.5 that

$$
\begin{aligned}
E[\widehat{c}_{j_1n}\widehat{c}_{j_2n}] &= E\left[\frac{1}{n^2}\sum_{k_1=1}^{n-j_1}\sum_{k_2=1}^{n-j_2} X\left(\frac{k_1}{\rho_n}\right)X\left(\frac{k_1+j_1}{\rho_n}\right)X\left(\frac{k_2}{\rho_n}\right)X\left(\frac{k_2+j_2}{\rho_n}\right)\right] \\
&= \frac{1}{n^2}\sum_{k_1=1}^{n-j_1}\sum_{k_2=1}^{n-j_2}\left[C\left(\frac{j_1}{\rho_n}\right)C\left(\frac{j_2}{\rho_n}\right) + C\left(\frac{k_1-k_2+j_1}{\rho_n}\right)C\left(\frac{k_1-k_2-j_2}{\rho_n}\right)\right. \\
&\qquad \left. +C\left(\frac{k_1-k_2}{\rho_n}\right)C\left(\frac{k_1-k_2+j_1-j_2}{\rho_n}\right) + Q\left(\frac{j_1}{\rho_n}, \frac{k_2-k_1}{\rho_n}, \frac{k_2-k_1+j_2}{\rho_n}\right)\right] \\
&= \left(1-\frac{j_1}{n}\right)C\left(\frac{j_1}{\rho_n}\right)\left(1-\frac{j_2}{n}\right)C\left(\frac{j_2}{\rho_n}\right) \\
&\qquad +\frac{1}{n}\sum_{|k|<n} U_n(k,j_1,j_2)\, C\left(\frac{k+j_1}{\rho_n}\right)C\left(\frac{k-j_2}{\rho_n}\right) \\
&\qquad +\frac{1}{n}\sum_{|k|<n} U_n(k,j_1,j_2)\, C\left(\frac{k}{\rho_n}\right)C\left(\frac{k+j_1-j_2}{\rho_n}\right) \\
&\qquad +\frac{1}{n}\sum_{|k|<n} U_n(k,j_1,j_2)\, Q\left(\frac{j_1}{\rho_n}, \frac{-k}{\rho_n}, \frac{-k+j_2}{\rho_n}\right),
\end{aligned}
\tag{2.19}
$$

where $U_n(k, j_1, j_2)$ is a function with values between 0 and 1 defined as follows

$$
U_n(k,j_1,j_2) = \begin{cases}
0, & k \le -n+j_2, \\
1-\frac{j_2-k}{n} & -n+j_2 < k < \min(0, j_2-j_1), \\
1-\frac{\max(j_1,j_2)}{n} & \min(0,j_2-j_1) \le k \le \max(0,j_2-j_1), \\
1-\frac{k+j_1}{n} & \max(0,j_2-j_1) < k < n-j_1, \\
0 & k \ge n-j_1.
\end{cases}
\tag{2.20}
$$

Therefore, by using (2.17) and (2.19), we have

$$\begin{aligned} Cov(\widehat{c}_{j_1n},\widehat{c}_{j_2n}) &= E[\widehat{c}_{j_1n}\widehat{c}_{j_2n}] - E[\widehat{c}_{j_1n}]E[\widehat{c}_{j_2n}] \\ &= \frac{1}{n}\sum_{|k|<n} U_n(k,j_1,j_2)C\left(\frac{k+j_1}{\rho_n}\right)C\left(\frac{k-j_2}{\rho_n}\right) \\ &\quad +\frac{1}{n}\sum_{|k|<n} U_n(k,j_1,j_2)C\left(\frac{k}{\rho_n}\right)C\left(\frac{k+j_1-j_2}{\rho_n}\right) \\ &\quad +\frac{1}{n}\sum_{|k|<n} U_n(k,j_1,j_2)Q\left(\frac{j_1}{\rho_n},\frac{-k}{\rho_n},\frac{-k+j_2}{\rho_n}\right). \end{aligned} \tag{2.21}$$

We now use this simplified form of $Cov(\widehat{c}_{j_1n},\widehat{c}_{j_2n})$ to establish the convergence of the three terms, T_1, $T_2(f)$ and $T_3(f)$.

By using (2.21) and Assumption 2.5, T_1 can be written as

$$\begin{aligned} T_1 =& \frac{1}{\rho_n^2}\left[\frac{2}{n}\sum_{|u|<n} U_n(u,0,0)C\left(\frac{u}{\rho_n}\right)^2 + \frac{1}{n}\sum_{|u|<n} U_n(u,0,0)Q\left(0,\frac{-u}{\rho_n},\frac{-u}{\rho_n}\right)\right] \\ \leq& \frac{2C(0)}{n\rho_n}\sum_{|u|<n}\left|C\left(\frac{u}{\rho_n}\right)\right|\frac{1}{\rho_n} + \frac{g_1(0)g_2(0)}{n\rho_n}\sum_{|u|<n} g_3\left(\frac{u}{\rho_n}\right)\frac{1}{\rho_n}. \end{aligned}$$

As in Theorem 2.1, we can view $\sum_{|u|<n}\left|C\left(\frac{u}{\rho_n}\right)\right|\frac{1}{\rho_n}$ as the integral of the function $s_n(\cdot)$ defined by

$$s_n(t) = \sum_{|u|<n}\left|C\left(\frac{u}{\rho_n}\right)\right| 1_{(\frac{u-1}{\rho_n},\frac{u}{\rho_n}]}(t).$$

Since $s_n(t) \to |C(t)|$, and $s_n(t) \leq h_0(t)$ holds from Assumption 2.1, we get $\lim_{n\to\infty}\sum_{|u|<n}\left|C\left(\frac{u}{\rho_n}\right)\right|\frac{1}{\rho_n} = \int_{-\infty}^{\infty}|C(t)|dt$ by applying the DCT as in Theorem 2.1, under Assumption 2.4. A similar argument, together with Assumptions 2.4 and 2.5, ensures that $\lim_{n\to\infty}\sum_{|u|<n} g_3\left(\frac{u}{\rho_n}\right)\frac{1}{\rho_n} = \int_{-\infty}^{\infty}|g_3(u)|du$. Both the limiting integrals are finite. So $nb_nT_1 \to 0$ as $n\to\infty$.

By using (2.21) and Assumption 2.5, the term $T_2(f)$ is given as

$$\begin{aligned} |T_2(f)| \leq& \frac{2}{\rho_n^2}\sum_{v=1}^{n-1}\left|\frac{1}{n}\sum_{|u|<n} U_n(u,0,v)\ C\left(\frac{u}{\rho_n}\right)C\left(\frac{u+v}{\rho_n}\right)\right||K(b_nv)| \\ &+\frac{2}{\rho_n^2}\sum_{v=1}^{n-1}\left|\frac{1}{n}\sum_{|u|<n} U_n(u,0,v)\ C\left(\frac{u}{\rho_n}\right)C\left(\frac{u-v}{\rho_n}\right)\right||K(b_nv)| \end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{\rho_n^2}\sum_{v=1}^{n-1}\left|\frac{1}{n}\sum_{|u|<n}U_n(u,0,v)\;Q\Big(\frac{v}{\rho_n},\frac{-u}{\rho_n},\frac{-u}{\rho_n}\Big)\right||K(b_nv)|\\
\leq&\frac{4|C(0)|}{nb_n\rho_n}\left(\sum_{|u|<n}\left|C\Big(\frac{u}{\rho_n}\Big)\right|\frac{1}{\rho_n}\right)\left(\sum_{v=1}^{n-1}|K(b_nv)|b_n\right)\\
&+\frac{2|g_1(0)g_2(0)|}{nb_n\rho_n}\left(\sum_{|u|<n}g_3\Big(\frac{u}{\rho_n}\Big)\frac{1}{\rho_n}\right)\left(\sum_{v=1}^{n-1}|K(b_nv)|b_n\right).
\end{aligned}
$$

The last expression does not depend on f. An argument as in the case of T_1 will show that $\lim_{n\to\infty}\sum_{v=1}^{n-1}|K(b_nv)|b_n = \int_0^\infty K(x)dx$, under Assumptions 2.2 and 2.3. The convergence of the other sums have already been discussed in connection with the term T_1. Hence, $nb_nT_2(f)\to 0$ as $n\to\infty$ uniformly for all f.

Now, we will consider $T_3(f)$. By using (2.21), this term can be written as

$$
\begin{aligned}
T_3(f)=&\frac{4}{\rho_n^2}\sum_{j_1=1}^{n-1}\sum_{j_2=1}^{n-1}Cov[\widehat{c}_{j_1n},\widehat{c}_{j_2n}]K(b_nj_1)K(b_nj_2)\cos\left(\frac{2\pi fj_1}{\rho_n}\right)\cos\left(\frac{2\pi fj_2}{\rho_n}\right)\\
=&T_{31}(f)+T_{32}(f)+T_{33}(f),
\end{aligned}
$$

where

$$
\begin{aligned}
T_{31}(f)=&\frac{4}{\rho_n^2}\sum_{j_1=1}^{n-1}\sum_{j_2=1}^{n-1}K(b_nj_1)K(b_nj_2)\cos\left(\frac{2\pi fj_1}{\rho_n}\right)\cos\left(\frac{2\pi fj_2}{\rho_n}\right)\\
&\times\left\{\frac{1}{n}\sum_{k=-(n-1)}^{(n-1)}U_n(k,j_1,j_2)\;C\Big(\frac{k+j_1}{\rho_n}\Big)\,C\Big(\frac{k-j_2}{\rho_n}\Big)\right\},\\
T_{32}(f)=&\frac{4}{\rho_n^2}\sum_{j_1=1}^{n-1}\sum_{j_2=1}^{n-1}K(b_nj_1)K(b_nj_2)\cos\left(\frac{2\pi fj_1}{\rho_n}\right)\cos\left(\frac{2\pi fj_2}{\rho_n}\right)\\
&\times\left\{\frac{1}{n}\sum_{k=-(n-1)}^{(n-1)}U_n(k,j_1,j_2)\;C\Big(\frac{k}{\rho_n}\Big)\,C\Big(\frac{k+j_1-j_2}{\rho_n}\Big)\right\},\\
T_{33}(f)=&\frac{4}{\rho_n^2}\sum_{j_1=1}^{n-1}\sum_{j_2=1}^{n-1}K(b_nj_1)K(b_nj_2)\cos\left(\frac{2\pi fj_1}{\rho_n}\right)\cos\left(\frac{2\pi fj_2}{\rho_n}\right)\\
&\times\left\{\frac{1}{n}\sum_{k=-(n-1)}^{(n-1)}U_n(k,j_1,j_2)\;Q\left(\frac{j_1}{\rho_n},\frac{-k}{\rho_n},\frac{-k+j_2}{\rho_n}\right)\right\}.
\end{aligned}
$$

From Assumptions 2.2 and 2.5,

$$|nb_nT_{33}(f)| \leq 4M^2\rho_n b_n \frac{1}{\rho_n^3}\sum_{j_1=1}^{n-1}\sum_{j_2=1}^{n-1}\sum_{k=-(n-1)}^{(n-1)} g_1\Big(\frac{j_1}{\rho_n}\Big)\, g_2\Big(\frac{k}{\rho_n}\Big)\, g_3\Big(\frac{k+j_2}{\rho_n}\Big).$$

By using a similar argument as in the case of T_1, we have

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{\rho_n^3}\sum_{v_1=1}^{n-1}\sum_{v_2=1}^{n-1}\sum_{u=-(n-1)}^{(n-1)} g_1\left(\frac{v_1}{\rho_n}\right) g_2\left(\frac{u}{\rho_n}\right) g_3\left(\frac{u+v_2}{\rho_n}\right)\\ =\int_0^\infty g_1(v_1)dv_1\int_0^\infty\left[\int_{-\infty}^\infty g_2(u)g_3(u+v_2)du\right]dv_2.\end{aligned}$$

Hence, $nb_nT_{33}(f)\to 0$ as $n\to\infty$ uniformly for all f.

Consider the term $T_{31}(f)$, let $u{=}j_1+j_2,\ v{=}k-j_2,\ w{=}j_2$.

$$\begin{aligned}T_{31}(f) =& \frac{4}{\rho_n^2}\sum_{w=1}^{n-1}\sum_{u=w+1}^{n-1+w} K(b_nw)K(b_n(u{-}w))\cos\left(\frac{2\pi fw}{\rho_n}\right)\cos\left(\frac{2\pi f(u{-}w)}{\rho_n}\right)\\ &\times\left\{\frac{1}{n}\sum_{v=-(n-1)-w}^{(n-1)-w} U_n(v+w,u-w,w)\; C\Big(\frac{u+v}{\rho_n}\Big)\, C\Big(\frac{v}{\rho_n}\Big)\right\}.\end{aligned}$$

From Assumption 2.2, observe that

$$nb_n|T_{31}(f)| \leq 4M\frac{b_n}{\rho_n^2}\sum_{w=1}^{n-1}\sum_{u=w+1}^{n-1+w}\sum_{v=-(n-1)-w}^{(n-1)-w}\left|K(b_nw)\; C\Big(\frac{u+v}{\rho_n}\Big)\, C\Big(\frac{v}{\rho_n}\Big)\right|.$$

Consider the simple function S_n, defined over $(0,\infty)\times(0,\infty)\times(-\infty,\infty)$ by

$$\begin{aligned}S_n(x,t,t') =& 4M\frac{1}{\pi^2}\sum_{w=1}^{n-1}\sum_{u=w+1}^{n-1+w}\sum_{v=-(n-1)-w}^{(n-1)-w}\left|K(b_nw)\; C\Big(\frac{u+v}{\rho_n}\Big)\, C\Big(\frac{v}{\rho_n}\Big)\right|\\ &\times 1_{((w-1)b_n,wb_n]}(x)1_{(\frac{u-1}{\rho_n},\frac{u}{\rho_n}]}(t)1_{(\frac{v-1}{\rho_n},\frac{v}{\rho_n}]}(t'),\end{aligned}$$

so that

$$nb_n|T_{31}(f)| \leq \int_0^\infty\int_0^\infty\int_{-\infty}^\infty S_n(x,t,t')dxdtdt'.$$

Since $\lim_{n\to\infty}\rho_nb_n = 0$, we have, for any fixed $(x,t,t')\in(0,\infty)\times(0,\infty)\times(-\infty,\infty)$ and for large enough n, the inequality $\rho_nb_n < x/t$, i.e., $t\rho_n < x/b_n$. Therefore, for large n, the unique integer u for which $1_{(\frac{u-1}{\rho_n},\frac{u}{\rho_n}]}(t)$ is non-zero is smaller than the unique integer w for which $1_{((w-1)b_n, wb_n]}(x)$ is non-zero. However, the ranges of summations in the definition of $S_n(x,t,t^{'})$ do not permit the order $u \leq w$. Therefore, $\lim_{n\to\infty} S_n(x,t,t^{'}) = 0$. From Assumption 2.1 and 2.2, we have the dominance

$$|S_n(x,t,t^{'})| \leq 4M|K_1(x)h_0(t+t^{'})h_0(t^{'})| \in L^1.$$

By applying the DCT, we have

$$\lim_{n\to\infty} \int_0^\infty \int_0^\infty \int_{-\infty}^\infty S_n(x,t,t^{'})dxdtdt^{'} = 0.$$

Since the function $S_n(x,t,t^{'})$ does not depend on f, we have $nb_nI_{31}(f) \to 0$ uniformly for all f.

In view of the convergence of the terms T_1, $T_2(f)$, $T_{33}(f)$ and $T_{31}(f)$, we have

$$\lim_{n\to\infty} nb_n \left[Var\left(\widehat{\phi}_n(f)\right) - T_{32}(f)\right] = 0 \tag{2.22}$$

uniformly for all f, and so we need to prove the convergence of $nb_nI_{32}(f)$ only.

Now, consider $T_{32}(f)$ and let $u = j_1 - j_2,\ w = j_2,\ v = k$.

$$\begin{aligned} T_{32}(f) \quad &= \quad \frac{4}{\rho_n^2}\sum_{w=1}^{n-1}\sum_{u=-w+1}^{n-1-w} K(b_nw)K(b_n(w+u))\cos\left(\frac{2\pi fw}{\rho_n}\right)\cos\left(\frac{2\pi f(w+u)}{\rho_n}\right) \\ &\quad \times\left\{\frac{1}{n}\sum_{v=-(n-1)}^{(n-1)} U_n(v,w+u,w)\ C\left(\frac{v}{\rho_n}\right)C\left(\frac{v+u}{\rho_n}\right)\right\}. \end{aligned} \tag{2.23}$$

For $f = 0$, it follows from (2.22) and Lemma 2.1 below that

$$\lim_{n\to\infty} nb_nVar[\widehat{\phi}_n(0)] = 2[\phi(0)]^2\int_{-\infty}^\infty K^2(x)dx. \tag{2.24}$$

For $f \neq 0$, we will further decompose $T_{32}(f)$ as follows. By applying the formula $2\cos(a)\cos(b) = \cos(a-b)+\cos(a+b)$ and $\cos(a+b) = \cos(a)\cos(b) - \sin(a)\sin(b)$, we have

$$T_{32}(f) = T_{321}(f) + T_{322}(f) - T_{323}(f),$$

where

$$\begin{aligned}
T_{321}(f) &= \frac{2}{n\rho_n^2}\sum_{w=1}^{n-1}\sum_{u=-w+1}^{n-1-w} K(b_n w)K(b_n(w+u))\cos\left(\frac{2\pi f u}{\rho_n}\right) \\
&\quad \times\left\{\sum_{v=-(n-1)}^{(n-1)} U_n(v,w+u,w)\ C\left(\frac{v}{\rho_n}\right)C\left(\frac{v+u}{\rho_n}\right)\right\}, \quad (2.25)\\
T_{322}(f) &= \frac{2}{n\rho_n^2}\sum_{w=1}^{n-1}\sum_{u=-w+1}^{n-1-w} K(b_n w)K(b_n(w+u))\cos\left(\frac{2\pi f u}{\rho_n}\right)\cos\left(\frac{4\pi f w}{\rho_n}\right) \\
&\quad \times\left\{\sum_{v=-(n-1)}^{(n-1)} U_n(v,w+u,w)\ C\left(\frac{v}{\rho_n}\right)C\left(\frac{v+u}{\rho_n}\right)\right\}, \quad (2.26)\\
T_{323}(f) &= \frac{2}{n\rho_n^2}\sum_{w=1}^{n-1}\sum_{u=-w+1}^{n-1-w} K(b_n w)K(b_n(w+u))\sin\left(\frac{2\pi f u}{\rho_n}\right)\sin\left(\frac{4\pi f w}{\rho_n}\right) \\
&\quad \times\left\{\sum_{v=-(n-1)}^{(n-1)} U_n(v,w+u,w)\ C\left(\frac{v}{\rho_n}\right)C\left(\frac{v+u}{\rho_n}\right)\right\}. \quad (2.27)
\end{aligned}$$

It follows from equation (2.22), Lemma 2.2 and Lemma 2.3 below that

$$\begin{aligned}
\lim_{n\to\infty} nb_n Var[\widehat{\phi}_n(f)] &= \lim_{n\to\infty} nb_n I_{321}(f) \\
&=[\phi(f)]^2\int_{-\infty}^{\infty} K^2(x)dx,
\end{aligned} \quad (2.28)$$

and the convergence is uniform over any closed interval that does not include the frequency 0. This completes the proof. □

Lemma 2.1.

$$\lim_{n\to\infty} nb_n I_{32}(0) = 4\int_0^{\infty} K^2(x)\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} C(t+t')C(t')\ dt'\right] dt\ dx.$$

Proof of Lemma 2.1. Consider the simple function $S_n(\cdot)$, defined over $(0,\infty)\times(-\infty,\infty)\times(-\infty,\infty)$ by

$$\begin{aligned}
S_n(x,t,t') =& 4\sum_{w=1}^{n-1}\sum_{u=-w+1}^{n-1-w}\sum_{v=-(n-1)}^{(n-1)} K(b_n w)K(b_n(w+u))U_n(v,w+u,w) \\
&\times C\left(\frac{v}{\rho_n}\right)C\left(\frac{v+u}{\rho_n}\right) 1_{((w-1)b_n,wb_n]}(x)1_{(\frac{u-1}{\rho_n},\frac{u}{\rho_n}]}(t)1_{(\frac{v-1}{\rho_n},\frac{v}{\rho_n}]}(t').
\end{aligned}$$

Observe from (2.23) that

$$nb_nI_{32}(0) = \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty S_n(x,t,t^{'})dxdtdt^{'}.$$

Define $w_n(x)$, $u_n(t)$ and $v_n(t^{'})$ as the smallest integers greater than or equal to x/b_n, $\rho_n t$ and $\rho_n t'$, respectively. Thus, $(x,t,t^{'}) \in (b_nw_{n-1}(x), b_nw_n(x)] \times \left(\frac{u_{n-1}(t)}{\rho_n}, \frac{u_n(t)}{\rho_n}\right] \times \left(\frac{v_{n-1}(t')}{\rho_n}, \frac{v_n(t')}{\rho_n}\right]$ and $b_nw_n(x) \to x$, $\frac{u_n(t)}{\rho_n} \to t$, $\frac{v_n(t^{'})}{\rho_n} \to t^{'}$ 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 $(x,t,t^{'}) \in (0,\infty)\times(-\infty,\infty)\times(-\infty,\infty)$ and large enough n, the inequalities $-\frac{x}{b_n\rho_n} < t < \frac{nb_n - x}{b_n\rho_n}$, i.e., $-w_n(x) < u_n(t) < n - w_n(x)$. Thus, for sufficiently large n, we have

$$\begin{aligned} S_n(x,t,t^{'}) = {} & 4K(b_nw_n(x))K\left(b_nw_n(x)+b_n\rho_n\frac{u_n(t)}{\rho_n}\right)U_n(v_n(t^{'}), w_n(x)+u_n(t), w_n(x)) \\ & \times C\left(\frac{v_n(t^{'})}{\rho_n}\right)C\left(\frac{v_n(t^{'})+u_n(t)}{\rho_n}\right). \end{aligned}$$

Also, for large n, we have

$$-(n-xb_n) + \frac{t}{\rho_n} < \frac{t'}{\rho_n} < (n-xb_n),$$

and so $U_n(v_n(t^{'}), w_n(x)+u_n(t), w_n(x))$ is positive and it converges to 1. Therefore, by virtue of Assumptions 2.1 and 2.2, we have

$$\lim_{n\to\infty} S_n(x,t,t^{'}) = \frac{1}{\pi^2}K^2(x)C(t^{'})C(t^{'}+t).$$

Again, from Assumptions 2.1 and 2.2, we have the dominance

$$|S_n(x,t,t^{'})| \le 4M|K_1(x)h_0(t+t^{'})h_0(t^{'})| \in L^1.$$

By applying the DCT, we have

$$\begin{aligned} \lim_{n\to\infty} nb_nI_{32}(0) = {} & \lim_{n\to\infty}\int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty S_n(x,t,t^{'})dxdtdt' \\ = {} & 4\int_0^\infty K^2(x)\int_{-\infty}^\infty \left[\int_{-\infty}^\infty C(t+t^{'})C(t^{'})dt^{'}\right]dtdx. \qquad \square \end{aligned}$$

Lemma 2.2. *The function $T_{321}(\cdot)$ converges as follows:*

$$\lim_{n\to\infty} nb_n I_{321}(f) = 2\int_0^\infty K^2(x)\int_{-\infty}^\infty \left[\cos(2\pi ft)\left[\int_{-\infty}^\infty C(t+t^{'})C(t^{'})dt^{'}\right]dt\right]dx.$$

The convergence is uniform on $[f_l, f_u]$ for arbitrary f_l and f_u such that $f_l < f_u$ and $f_l f_u > 0$.

Proof of Lemma 2.2. Consider the simple function $S_n(\cdot)$, defined over $[f_l, f_u] \times (0,\infty)\times(-\infty,\infty)\times(-\infty,\infty)$ by

$$\begin{aligned} S_n(f,x,t,t^{'}) =& 2\sum_{w=1}^{n-1}\sum_{u=-w+1}^{n-1-w}\sum_{v=-(n-1)}^{(n-1)} K(b_n w)K(b_n(w+u))\cos\left(\frac{2\pi fu}{\rho_n}\right)U_n(v,w+u,w) \\ & \times C\left(\frac{v}{\rho_n}\right)C\left(\frac{v+u}{\rho_n}\right)1_{((w-1)b_n, wb_n]}(x)1_{(\frac{u-1}{\rho_n},\frac{u}{\rho_n}]}(t)1_{(\frac{v-1}{\rho_n},\frac{v}{\rho_n}]}(t^{'}), \end{aligned}$$

so that, from (2.25),

$$nb_n I_{321}(f) = \int_0^\infty\int_{-\infty}^\infty\int_{-\infty}^\infty S_n(f,x,t,t^{'})dxdtdt^{'}.$$

A similar argument as in the proof of Lemma 2.1 will show that for $(x,t,t^{'}) \in (0,\infty)\times(-\infty,\infty)\times(-\infty,\infty)$ and sufficiently large n,

$$\begin{aligned} S_n(f,x,t,t^{'}) =& 2K(b_n w_n(x))\; K\left(b_n w_n(x) + b_n\rho_n\frac{u_n(t)}{\rho_n}\right)\cos\left(\frac{2\pi f u_n(t)}{\rho_n}\right) \\ & \times U_n(v_n(t^{'}), w_n(x)+u_n(t), w_n(x))C\left(\frac{v_n(t^{'})}{\rho_n}\right)C\left(\frac{v_n(t^{'})+u_n(t)}{\rho_n}\right), \end{aligned}$$

where $w_n(x)$, $u_n(t)$ and $v_n(t^{'})$ are the smallest integers greater than or equal to x/b_n, $\rho_n t$ and $\rho_n t'$, respectively, and that the function $S_n(f,x,t,t^{'})$ converges to the function $S(\cdot)$, defined over $(0,\infty)\times(-\infty,\infty)\times(-\infty,\infty)$ by

$$S(f,x,t,t^{'}) = 2K^2(x)\cos(2\pi ft)C(t^{'})C(t^{'}+t).$$

Observe also that $\int_0^\infty\int_{-\infty}^\infty\int_{-\infty}^\infty S(f,x,t,t^{'})dxdtdt^{'}$ is a continuous function in f. As in the proof of Theorem 2.1, we prove the convergence of $\int_0^\infty\int_{-\infty}^\infty\int_{-\infty}^\infty S_n(f,x,t,t^{'})dxdtdt^{'}$

uniformly on $[f_l, f_u]$, by showing that for any sequence $f_n \to f$,

$$\lim_{n\to\infty} \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty S_n(f_n, x, t, t^{'}) dxdtdt^{'} = \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty S(f, x, t, t^{'}) dxdtdt^{'}$$

for $f_n, f \in [f_l, f_u]$. The latter convergence follows, through Assumption 2.1 and 2.2 and the DCT, from the dominance

$$|S_n(f, x, t, t^{'})| \le 2M|K_1(x)h_0(t + t^{'})h_0(t^{'})| \in L^1,$$

and the convergence of the integrand, which holds because of the continuity of the kernel, the cosine and the covariance function.

Hence, $nb_nI_{321}(\cdot)$ converges as stated uniformly on $[f_l, f_u]$. □

Lemma 2.3. *The functions $nb_nI_{322}(\cdot)$ and $nb_nI_{323}(\cdot)$ converge to* 0 *uniformly on* $[f_l, f_u]$ *for arbitrary f_l and f_u such that $f_l < f_u$ and $f_lf_u > 0$.*

Proof of Lemma 2.3. Consider the simple function $S_n(\cdot)$, defined over $[f_l, f_u] \times (0,\infty) \times (-\infty,\infty) \times (-\infty,\infty)$ by

$$\begin{aligned} &S_n(f, x, t, t^{'}) \\ &= 2\sum_{w=1}^{n-1}\sum_{u=-w+1}^{n-1+w}\sum_{v=-(n-1)}^{(n-1)} K(b_nw)K(b_n(w+u))\cos\left(\frac{2\pi fu}{\rho_n}\right)\cos\left(\frac{4\pi fb_nw}{\rho_nb_n}\right) \\ &\quad\times U_n(v, w+u, w)C\left(\frac{v}{\rho_n}\right)C\left(\frac{v+u}{\rho_n}\right)1_{((w-1)b_n, wb_n]}(x)\times 1_{(\frac{u-1}{\rho_n},\frac{u}{\rho_n}]}(t)1_{(\frac{v-1}{\rho_n},\frac{v}{\rho_n}]}(t^{'}), \end{aligned}$$

so that, from (2.26),

$$nb_nI_{322}(f) = \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty S_n(f, x, t, t^{'})dxdtdt^{'}. \qquad (2.29)$$

A similar argument as in the proof of Lemma 2.1 will show that for $(x, t, t^{'}) \in (0,\infty) \times (-\infty,\infty) \times (-\infty,\infty)$ and sufficiently large n,

$$\begin{aligned}
&S_n(f,x,t,t') \\
&\quad = 2K(b_n w_n(x))K\left(b_n w_n(x) + b_n\rho_n\frac{u_n(t)}{\rho_n}\right)\cos\left(\frac{2\pi f u_n(t)}{\rho_n}\right)\cos\left(\frac{4\pi f w_n(x) b_n}{b_n\rho_n}\right) \\
&\quad\quad \times U_n(v_n(t'), u_n(t)+w_n(x), w_n(x))C\left(\frac{v_n(t')}{\rho_n}\right)C\left(\frac{v_n(t')+u_n(t)}{\rho_n}\right),
\end{aligned}$$

where $w_n(x)$, $u_n(t)$ and $v_n(t')$ are the smallest integers greater than or equal to x/b_n, $\rho_n t$ and $\rho_n t'$, respectively.

For obtaining the uniform convergence of $T_{322}(\cdot)$, consider

$$\begin{aligned}
&\sup_{f\in[f_l,f_u]}\left|\int_0^\infty\int_{-\infty}^\infty\int_{-\infty}^\infty S_n(f,x,t,t')dxdtdt'\right| \\
&\quad\quad \leq \sup_{f\in[f_l,f_u]}\int_0^\infty\int_{-\infty}^\infty\int_{-\infty}^\infty |S_n(f,x,t,t') - g_n(f,x,t,t')|dxdtdt' \\
&\quad\quad\quad + \sup_{f\in[f_l,f_u]}\left|\int_0^\infty\int_{-\infty}^\infty\int_{-\infty}^\infty g_n(f,x,t,t')dxdtdt'\right|, \qquad (2.30)
\end{aligned}$$

where the function $g_n(\cdot)$ is defined over $[f_l, f_u]\times(0,\infty)\times(-\infty,\infty)\times(-\infty,\infty)$ by

$$g_n(f,x,t,t') = 2\cos(2\pi f t)\cos\left(\frac{4\pi f x}{b_n\rho_n}\right)K^2(x)C(t')C(t+t').$$

We shall prove the convergence of $nb_nI_{322}(\cdot)$ given in (2.29) by proving the convergence of the two integrals on the right hand side of (2.30).

In order to prove the first convergence, we follow the route taken in Theorem 2.1, i.e., we show that for any sequence $f_n \to f$

$$\lim_{n\to\infty}\int_0^\infty\int_{-\infty}^\infty\int_{-\infty}^\infty |S_n(f_n,x,t,t') - g_n(f_n,x,t,t')|dxdtdt' = 0,$$

for $f, f_n \in [f_l, f_u]$. The above integral can be written as

$$\begin{aligned}
&\int_0^\infty\int_{-\infty}^\infty\int_{-\infty}^\infty |S_n(f_n,x,t,t') - g_n(f_n,x,t,t')|dxdtdt' \\
&\quad\quad \leq \int_0^\infty\int_{-\infty}^\infty\int_{-\infty}^\infty |S_n(f_n,x,t,t') - G_n(f_n,x,t,t')|dxdtdt' \\
&\quad\quad\quad + \int_0^\infty\int_{-\infty}^\infty\int_{-\infty}^\infty |G_n(f_n,x,t,t') - g_n(f_n,x,t,t')|dxdtdt', \qquad (2.31)
\end{aligned}$$

where the function $G_n(\cdot)$ is defined over $[f_l, f_u] \times (0,\infty) \times (-\infty,\infty) \times (-\infty,\infty)$ by

$$G_n(f,x,t,t^{'}) = 2\cos(2\pi ft)\cos\left(\frac{4\pi f w_n(x) b_n}{b_n\rho_n}\right) K^2(x)C(t^{'})C(t+t^{'})).$$

Now, observe that

$$|S_n(f_n,x,t,t^{'}) - G_n(f_n,x,t,t^{'})| \le 2M \left|\cos\left(\frac{4\pi f_n w_n(x) b_n}{b_n\rho_n}\right)\alpha_n(f_n,x,t,t^{'})\right|,$$

where

$$\begin{aligned}\alpha_n(f_n,x,t,t^{'}) =& K(b_n w_n(x))\cos\left(\frac{2\pi f_n u_n(t)}{\rho_n}\right) U_n(v_n(t^{'}), u_n(t)+w_n(x), w_n(x)) \\ & \times C\left(\frac{v_n(t^{'})}{\rho_n}\right) C\left(\frac{v_n(t^{'})+u_n(t)}{\rho_n}\right) - \cos(2\pi f_n t)K(x)C(t^{'})C(t+t^{'}).\end{aligned}$$

Since $\alpha_n(f_n,x,t,t^{'}) \to 0$ as $n \to \infty$, it is easy to see that

$$\lim_{n\to\infty} |S_n(f_n,x,t,t^{'}) - G_n(f_n,x,t,t^{'})| = 0.$$

From Assumption 2.1 and 2.2, we have the dominance

$$|S_n(f_n,x,t,t^{'}) - G_n(f_n,x,t,t^{'})| \le 4M|K_1(x)h_0(t')h_0(t+t')| \in L^1.$$

By applying the DCT, we have

$$\lim_{n\to\infty}\int_0^\infty\int_{-\infty}^\infty\int_{-\infty}^\infty |S_n(f_n,x,t,t^{'}) - G_n(f_n,x,t,t^{'})| dxdtdt' = 0.$$

Turning to the second term on the right hand side of (2.31), observe that for any fixed x, $|x - w_n(x)b_n| \le b_n$. By applying the Mean Value Theorem to the cosine function in the interval $[\frac{f_n x}{\rho_n b_n}, \frac{f_n w_n(x) b_n}{\rho_n b_n}]$, we have

$$\cos\left(\frac{2\pi f_n w_n(x) b_n}{\rho_n b_n}\right) - \cos\left(\frac{2\pi f_n x}{\rho_n b_n}\right) = -\sin(\theta)\left|\frac{2\pi f_n w_n(x) b_n}{\rho_n b_n} - \frac{2\pi f_n x}{\rho_n b_n}\right|,$$

for some $\theta \in \left[\frac{2\pi f_n x}{\rho_n b_n}, \frac{2\pi f_n w_n(x) b_n}{\rho_n b_n}\right]$. Therefore

$$\left|\cos\left(\frac{2\pi f_n w_n(x) b_n}{\rho_n b_n}\right) - \cos\left(\frac{2\pi f_n x}{\rho_n b_n}\right)\right| \le \frac{2\pi f_n}{\rho_n}.$$

Thus,

$$|G_n(f_n,x,t,t^{'}) - g_n(f_n,x,t,t^{'})| \le 2M^2C^2(0)\left|\cos\left(\frac{2\pi f_n w_n(x) b_n}{\rho_n b_n}\right) - \cos\left(\frac{2\pi f_n x}{\rho_n b_n}\right)\right|$$
$$\le 2M^2C^2(0)\frac{2\pi f_n}{\rho_n}.$$

So

$$\lim_{n\to\infty}|G_n(f_n,x,t,t^{'}) - g_n(f_n,x,t,t^{'})| \to 0.$$

From Assumption 2.1 and 2.2, we have the dominance

$$|G_n(f_n,x,t,t^{'}) - g_n(f_n,x,t,t^{'})| \le 4M|K_1(x)h_0(t+t^{'})h_0(t^{'})| \in L^1,$$

which leads us, through another use of the DCT, the convergence of the second integral of (2.31). This establishes that the first term on the right hand side of (2.30) converges to 0. We only have to deal with the second term.

Let

$$s_n(f) = \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty g_n(f,x,t,t^{'})dxdtdt^{'}.$$

In order to establish the uniform convergence of $s_n(\cdot)$ over $[f_l, f_u]$, it is enough to show that $s_n(f_n) \to 0$ for any sequence $f_n \to f$, where $f, f_n \in [f_l, f_u]$. By using the Reimann-Lebesgue lemma, we have $s_n(f_n) \to 0$. Thus, the second term on the right hand side of (2.30) also converges to 0. Hence, $nb_nI_{322}(f)$ converges to 0 uniformly on $[f_l, f_u]$ as $n \to \infty$.

Convergence of $nb_nI_{323}(f)$ to 0 can be established in a similar manner. □

Proof of Theorem 2.3. Assumption 2.1 ensures absolute summability of the covariance sequence $\{C(k/\rho_n),\ k = \ldots,-2,-1,0,1,2,\ldots\}$ of the uniformly sampled process $\left\{X\left(\frac{k}{\rho_n}\right),\ k = \ldots,-2,-1,0,1,2,\ldots\right\}$, for fixed ρ_n. The corresponding spectral density $\phi_{\rho_n}(\cdot)$ is defined as

$$\phi_{\rho_n}(f) = \frac{1}{\rho_n}\sum_{j=-\infty}^{\infty} C\left(\frac{j}{\rho_n}\right)e^{-\frac{i2\pi f j}{\rho_n}},\quad f \in (-\infty,\infty).$$

The function $\phi_{\rho_n}(\cdot)$ is periodic with period ρ_n and is related to $\phi(\cdot)$ as follows:

$$\phi_{\rho_n}(f) = \sum_{l=-\infty}^{\infty} \phi(f + l\rho_n), \qquad f \in (-\rho_n/2, \rho_n/2].$$

In particular, for $f \in [-\rho_n/2, \rho_n/2]$,

$$\phi(f) = \phi_{\rho_n}(f) - \sum_{l=1}^{\infty} \phi(f + l\rho_n) - \sum_{l=-\infty}^{-1} \phi(f + l\rho_n).$$

For sufficiently large n, $\rho_n/2$ lies outside any finite interval $[f_l, f_u]$, and the bias of the estimator $\widehat{\phi}_n(f)$ given by (2.1) on $[f_l, f_u]$ can be decomposed as follows.

$$\begin{aligned}
&E[\widehat{\phi}_n(f)] - \phi(f) \\
&= \frac{1}{\rho_n} \sum_{|v|<n} \left(1 - \frac{|v|}{n}\right) C\left(\frac{v}{\rho_n}\right) K(b_n v) e^{-\frac{i2\pi f v}{\rho_n}} - \phi(f) \\
&= \frac{1}{\rho_n} \sum_{|v|<n} \left(1 - \frac{|v|}{n}\right) C\left(\frac{v}{\rho_n}\right) K(b_n v) e^{-\frac{i2\pi f v}{\rho_n}} - \phi_{\rho_n}(f) + \sum_{l=1}^{\infty} \phi(f + l\rho_n) + \sum_{l=-\infty}^{-1} \phi(f + l\rho_n) \\
&= B_1(f) + B_2(f) + B_3(f) + B_4(f) + B_5(f), \qquad (2.32)
\end{aligned}$$

where

$$\begin{aligned}
B_1(f) &= -\frac{1}{\rho_n} \sum_{|v|<n} (1 - K(b_n v)) C\left(\frac{v}{\rho_n}\right) e^{-\frac{i2\pi f v}{\rho_n}}, \\
B_2(f) &= -\frac{1}{\rho_n} \sum_{|v|<n} \frac{|v|}{n} C\left(\frac{v}{\rho_n}\right) K(b_n v) e^{-\frac{i2\pi f v}{\rho_n}}, \\
B_3(f) &= -\frac{1}{\rho_n} \sum_{|v|\geq n} C\left(\frac{v}{\rho_n}\right) e^{-\frac{i2\pi f v}{\rho_n}}, \\
B_4(f) &= \sum_{l=1}^{\infty} \phi(f + l\rho_n), \\
B_5(f) &= \sum_{l=-\infty}^{-1} \phi(f + l\rho_n).
\end{aligned}$$

We will consider each $B_i(f)$, $i = 1, \ldots, 5$, separately.

$$\left(\frac{1}{\rho_n b_n}\right)^q B_1(f) = -\sum_{|v|<n} \left(\frac{1 - K(b_n v)}{b_n^q |v|^q}\right) \frac{|v|^q}{\rho_n^q} C\left(\frac{v}{\rho_n}\right) e^{-\frac{i2\pi f v}{\rho_n}} \frac{1}{\rho_n}.$$

Consider the simple function S_n defined over $[f_l, f_u] \times (-\infty, \infty)$ as

$$S_n(f,t) = -\sum_{|v|<n} \left(\frac{1-K(b_n v)}{b_n^q |v|^q} \right) \frac{|v|^q}{\rho_n^q} C\left(\frac{v}{\rho_n}\right) e^{-\frac{i2\pi f v}{\rho_n}} 1_{\left(\frac{v-1}{\rho_n}, \frac{v}{\rho_n}\right]}(t).$$

Observe that $\left(\frac{1}{\rho_n b_n}\right)^q B_1(f) = \int_{-\infty}^{\infty} S_n(f,t)dt$.

For any $t \in (-\infty, \infty)$, we define $v_n(t)$ as the smallest integer greater than or equal to $t\rho_n$. It follows that $\frac{v_n(t)}{\rho_n} \to t$ as $n \to \infty$, and for sufficiently large n and any $t \in (-\infty, \infty)$, we can write

$$S_n(f,t) = -\left(\frac{1-K(b_n v_n(t))}{b_n^q |v_n(t)|^q} \right) \frac{|v_n(t)|^q}{\rho_n^q} C\left(\frac{v_n(t)}{\rho_n}\right) e^{-\frac{i2\pi f v_n(t)}{\rho_n}}.$$

From Assumptions 2.4 and 2.2A, we have

$$\lim_{n\to\infty} S_n(t) = -k_q |t|^q C(t) e^{-i2\pi f t}.$$

Also, Assumption 2.1A implies

$$|S_n(t)| \leq M_1 h_q(t) \text{ where } M_1 = \sup_x \frac{1-K(x)}{x^q}.$$

By applying the DCT, we have

$$\lim_{n\to\infty} \left(\frac{1}{\rho_n b_n}\right)^q B_1(f) = \lim_{n\to\infty} \int_{-\infty}^{\infty} S_n(f,t)dt = -k_q \int_{-\infty}^{\infty} |t|^q C(t) e^{-i2\pi f t} dt. \qquad (2.33)$$

Thus, $B_1(f)$ is $O\left((\rho_n b_n)^q\right)$. The fact that this convergence is uniform over the interval $[f_l, f_u]$ can be established by choosing any sequence f_n in this interval that converges to f, and showing that $\int_{-\infty}^{\infty} S_n(f_n,t)dt$ converges to the right hand side of (2.33).

The term $B_2(f)$ can be written as

$$\frac{n}{\rho_n} B_2(f) = \int_{-\infty}^{\infty} S_n(f,t)dt,$$

where $S_n(\cdot)$ is defined over $[f_l, f_u] \times (-\infty, \infty)$ as

$$S_n(f,t) = -\sum_{|v|<n} \frac{|v|}{\rho_n} C\left(\frac{v}{\rho_n}\right) K(b_n v) e^{-\frac{i2\pi f v}{\rho_n}} 1_{\left(\frac{v-1}{\rho_n}, \frac{v}{\rho_n}\right]}(t).$$

As in the case of $B_1(f)$, it can be shown that

$$\lim_{n\to\infty} S_n(f,t) = -|t|C(t)e^{-i2\pi ft}.$$

From Assumption 2.1A, it follows that $|S_n(f,t)| \le h_q(t)$. Again, by applying the DCT, we have

$$\lim_{n\to\infty} \frac{n}{\rho_n} B_2(f) = \lim_{n\to\infty} \int_{-\infty}^{\infty} S_n(f,t)dt = -\int_{-\infty}^{\infty} |t|C(t)e^{-i2\pi ft}dt.$$

Thus, $B_2(f)$ is $O(\frac{\rho_n}{n})$. The uniform convergence can be argued similarly as in the case of $B_1(f)$.

The term $B_3(f)$ satisfies

$$\begin{aligned}
|B_3(f)| \le & \frac{1}{\rho_n} \sum_{|v|\ge n} \left| C\left(\frac{v}{\rho_n}\right) \right| = \left(\frac{\rho_n}{n}\right)^q \frac{1}{2\pi} \sum_{|v|\ge n} \left(\frac{n}{\rho_n}\right)^q \left| C\left(\frac{v}{\rho_n}\right) \frac{1}{\rho_n} \right| \\
\le & \left(\frac{\rho_n}{n}\right)^q \frac{1}{\pi} \left[\sum_{v\ge 1} \left(\frac{v}{\rho_n}\right)^q \left| C\left(\frac{v}{\rho_n}\right) \frac{1}{\rho_n} \right| \right].
\end{aligned}$$

Observe that for each fixed n, we have from Assumption 2.1A,

$$\begin{aligned}
\lim_{m\to\infty} \sum_{v=1}^{m} \left(\frac{v}{\rho_n}\right)^q \left| C\left(\frac{v}{\rho_n}\right) \right| \frac{1}{\rho_n} &\le \lim_{n\to\infty} \sum_{v=1}^{m} h_q\left(\frac{v}{\rho_n}\right) \frac{1}{\rho_n} \\
&\le \lim_{n\to\infty} \int_0^{\frac{m}{\rho_n}} h_q(t)dt \le \int_0^{\infty} h_q(t)dt.
\end{aligned}$$

So

$$\limsup_{n\to\infty} \sum_{v=1}^{\infty} \left(\frac{v}{\rho_n}\right)^q \left| C\left(\frac{v}{\rho_n}\right) \right| \frac{1}{\rho_n} \le \int_0^{\infty} h_q(t)dt.$$

Hence, $|B_3(f)|$ is bounded by an $O((\frac{\rho_n}{n})^q)$ term, which converges to zero faster than $B_2(f)$.

As for the term $B_4(f)$, we have from Assumption 2.1B and the DCT

$$\begin{aligned}
\lim_{n\to\infty} (\rho_n)^p B_4(f) = & \lim_{n\to\infty} \sum_{l=1}^{\infty} \frac{(\rho_n)^p}{|f + l\rho_n|^p} |f + l\rho_n|^p \phi(f + l\rho_n) \\
= & \frac{1}{(2\pi)^p} \sum_{l=1}^{\infty} \lim_{n\to\infty} \frac{1}{|\frac{f}{\rho_n} + l|^p} \left(\lim_{n\to\infty} |f + l\rho_n|^p \phi(f + l\rho_n) \right) = A \sum_{l=1}^{\infty} \frac{1}{|l|^p}.
\end{aligned}$$

Hence $B_4(f) = O\left(\frac{1}{(\rho_n)^p}\right)$.

Similarly it can be proved that $B_5(f) = O\left(\frac{1}{(\rho_n)^p}\right)$.

The theorem is proved by combining the five terms. □

Proof of Theorem 2.4. It follows from Theorems 2.2 and 2.3 that the MSE of the estimator $\widehat{\phi}_n(\cdot)$ can be written as

$$\begin{aligned} MSE[\widehat{\phi}_n(f)] &= [E\{\widehat{\phi}_n(f) - \phi(f)\}]^2 + Var[\widehat{\phi}_n(f)] \\ &= O\left((\rho_n b_n)^{2q}\right) + O\left(\frac{\rho_n^2}{n^2}\right) + O\left(\frac{1}{\rho_n^{2p}}\right) \\ &\quad +O\left(\frac{1}{nb_n}\right). \end{aligned} \tag{2.34}$$

Let us first fix n and ρ_n and minimize the MSE with respect to b_n. The squared bias is an increasing functions of b_n, while the variance is a decreasing function of b_n. Therefore, the maximum possible value is minimized (i.e., the fastest rate of convergence is achieved) when $(\rho_n b_n)^{2q} \propto (nb_n)^{-1}$, i.e., when

$$b_n \propto \left(n\rho_n^{2q}\right)^{-\frac{1}{2q+1}}. \tag{2.35}$$

By substituting this value in the expression for the MSE, and making use of the fact that $\frac{2q}{2q+1} < 2$ and $\rho_n/n < 1$, we have

$$MSE[\widehat{\phi}_n(f)] = O\left(\left(\frac{\rho_n}{n}\right)^{\frac{2q}{2q+1}}\right) + O\left(\frac{1}{\rho_n^{2p}}\right).$$

The first term on the right hand side is an increasing function of ρ_n, while the second term is a decreasing function of ρ_n. Therefore, the maximum of the two terms is minimized when $\left(\frac{\rho_n}{n}\right)^{\frac{2q}{2q+1}} \propto \rho_n^{-2p}$, i.e., when ρ_n is chosen as in (2.6). The optimal rate for b_n, as given in (2.7), is obtained by substituting the expression for ρ_n in (2.35). Further substitution of these two optimal rates in (2.34) gives (2.5). □

2.5 Simulation study

In this section, we shall present the reslts of a simulation study of performance of the spectral density estimators based on uniform and Poisson sampled data. We consider the continuous time autoregressive (AR(4)) process having the spectral density

$$\phi(f) = \sigma^2 \cdot \frac{1}{((2\pi f)^2 + \alpha_1^2)((2\pi f)^2 + \alpha_2^2)((2\pi f)^2 + \alpha_3^2)((2\pi f)^2 + \alpha_4^2)}, \tag{2.36}$$

where $\alpha_1 = 0.65$, $\alpha_2 = 0.75$, $\alpha_3 = 0.85$, $\alpha_4 = 0.95$ and $\sigma = 1/2$. A procedure for generating the samples is outlined in Section A.1 of the Appendix. For estimation, we assume that the underlying power spectral density satisfies Assumption 2.1A with $q = 2$. Accordingly, we use the Hanning Kernel

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

which has characteristic exponent 2.

2.5.1 Performance of the estimator

Here, we consider the performance of $\widehat{\phi}_n(f)$ over the frequency range $[0, 0.5]$. We used the optimal choice of sampling rate developed in Section 2.3.1 to generate uniformly spaced samples of the process for sample sizes $n = 100$, 1000 and 10000. We assume Assumption 2.1A with $q = 2$ and Assumption 2.1B with $p = 8$ (both of which actually hold for the underlying power spectral density). For the above choices, the optimal powers of n for the sampling rate and the bandwidth are $\rho_n \propto n^{1/21}$ and $b_n \propto n^{-5/21}$. We choose $\rho_n = n^{1/21}$ and $b_n = \frac{1}{4}n^{-5/21}$.

Figure 2.1 shows the average of the estimated power spectral density computed from 500 simulation runs, the empirically observed bias and variance, together with the true power spectral density and the theoretical (asymptotic) bias and variance, respectively, for the three samples sizes.

From these figures, it can be observed that as the sample size goes from 100 to 10000, the empirical values of bias and variance get closer to the asymptotic results. Moreover, the theoretical (asymptotic) computations are quite comparable to the empirical values, even for sample size 100.

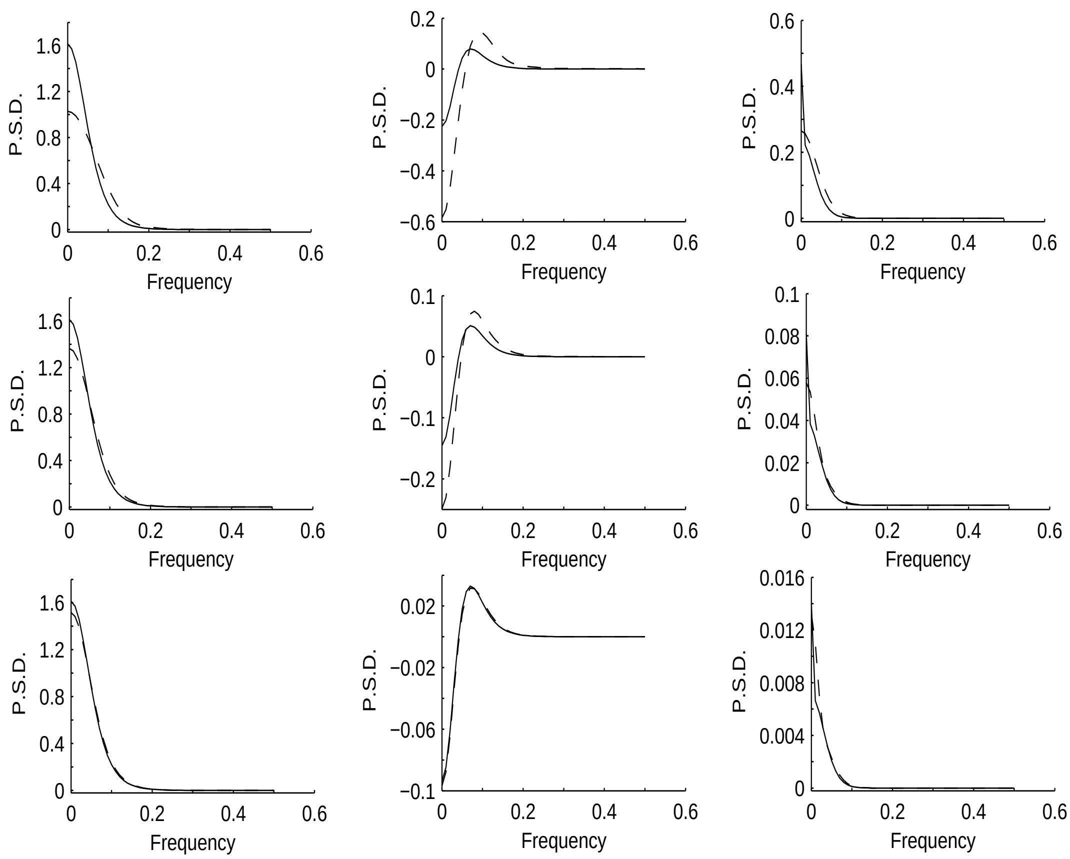

Figure 2.1: The average estimated power spectral density $\widehat{\phi}_n(\cdot)$ (left column), the bias (middle column) and the variance (right column) for sample sizes 100 (top row), 1000 (middle row) and 10000 (bottom row). The solid and the dotted lines correspond to theoretical (asymptotic) and empirical values, respectively.

2.5.2 Comparison with Poisson sampling estimator

We generate Poisson sampled data with the average sampling rate $\beta = 1$ for sample sizes $n = 100$, 1000 and 10000, and compute the estimator $\widehat{\psi}_n(f)$ on $[0, 0.5]$. Here, the optimal power of n for the bandwidth is $b_n \propto n^{-1/5}$. We use $b_n = \frac{1}{4}n^{-1/5}$.

Figure 2.2 shows the empirical bias, variance and MSE of the estimators $\widehat{\phi}_n(f)$ and $\widehat{\psi}_n(f)$ computed from 500 simulation runs, as a function of the frequency, for sample sizes 100, 1000 and 10000.

From these figures, it can be observed that the bias of $\widehat{\psi}_n(\cdot)$ is generally less than that of $\widehat{\phi}_n(\cdot)$ while the variance of $\widehat{\psi}_n(\cdot)$ is larger than that of $\widehat{\phi}_n(\cdot)$. The differences diminish with larger sample size. These patterns are in accordance with the large sample

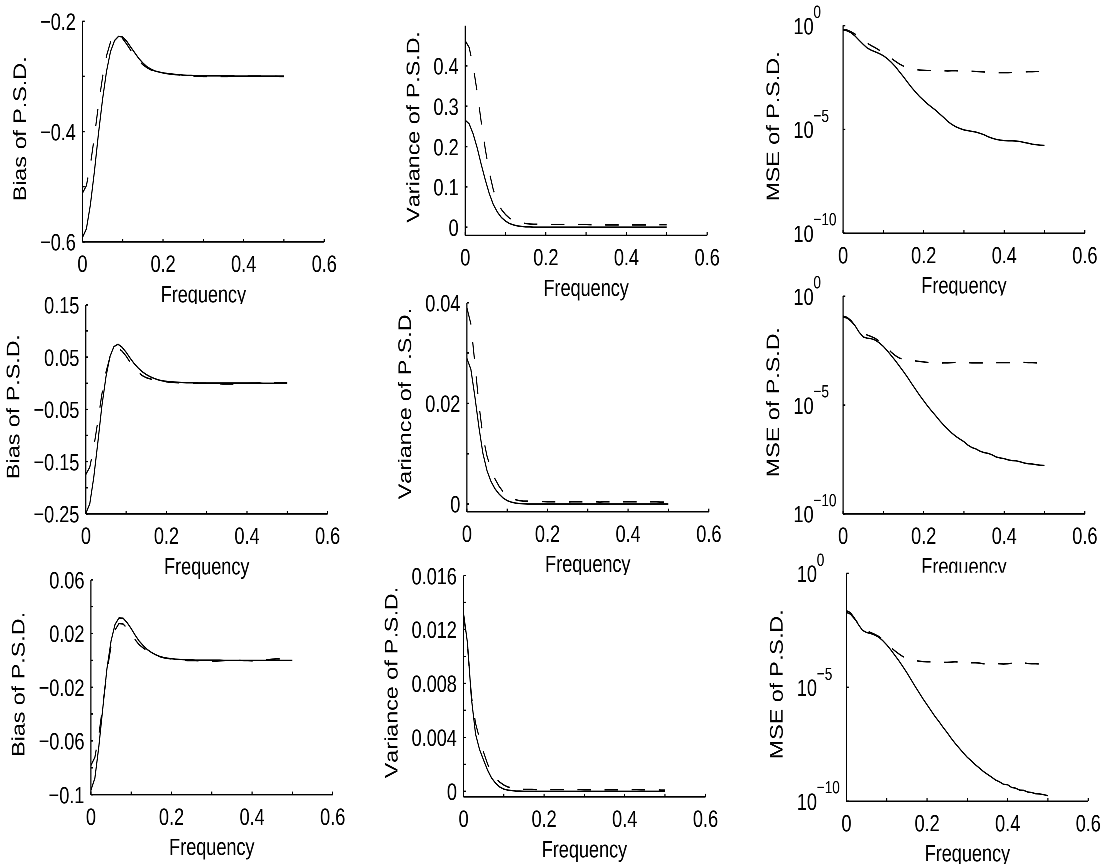

Figure 2.2: The estimated bias of the power spectral density $\widehat{\phi}_n(\cdot)$ and $\widehat{\psi}_n(\cdot)$ (left column), the variance (middle column) and the MSE (in log scale, right column) for sample sizes 100 (top row), 1000 (middle row) and 10000 (bottom row). The solid and the dotted lines correspond to uniform and Poisson samplings, respectively.

comparisons made in Section 2.3.2. The MSE of $\widehat{\psi}_n(\cdot)$ is larger than that of the $\widehat{\phi}_n(\cdot)$ for larger frequencies. The MSE is plotted in log-scale in order to highlight the fact that this quantity, in the case of $\widehat{\psi}_n(\cdot)$, levels off to a constant value for larger frequencies, while in the case of $\widehat{\phi}_n(\cdot)$, it continues to decline. This difference in behaviour is in accordance with the variance expressions given in (2.8) and (2.14).

2.6 Summary and discussion

In this chapter, we have shown that the smoothed periodogram based on uniformly spaced samples of a continuous time stationary stochastic process is consistent, under certain conditions, provided that the sampling rate increases appropriately as the sample size goes to infinity. We have also shown that, under the conditions used in the proofs, the estimators based on uniformly and non-uniformly spaced samples have about the same rate of convergence. Thus, our results remove a widely perceived theoretical deficiency of a popular spectral estimator based on uniform sampling.

It has been a common experience, both theoretically and empirically (Roberts and Gaster, 1980), that the smoothed periodogram estimator (1.12) of a non-bandlimited power spectral density has less variance and more bias compared to the corresponding estimator $\widehat{\psi}_n(\cdot)$ based on Poisson sampling. What the results of Section 2.3 show is that, even though the new asymptotic results presented in this chapter establish consistency of the smoothed periodogram $\widehat{\phi}_n(\cdot)$ and the rates of convergence of the estimators $\widehat{\phi}_n(\cdot)$ and $\widehat{\psi}_n(\cdot)$ are comparable, the constants for the first order approximations of the bias and variance of the two estimators exhibit the same type of trade off, i.e., the constant for the bias term is larger in the case of $\widehat{\phi}_n(\cdot)$, and the constant for the variance term is larger in the case of $\widehat{\psi}_n(\cdot)$.

The new asymptotic calculations provide a common-sense theoretical justification for using the smoothed periodogram, even if the underlying power spectral density is not bandlimited. The approach consists of appropriate filtering of the continuous time process followed by sampling at a suitably uniform rate. Remark 2.6 and Theorem 2.4 give guidelines for choosing a suitable filter and an appropriate sampling rate, respectively, which may be useful for practitioners.

The simulation results reported in Section 2.5 illustrate how one can choose an appropriate sampling rate for estimating the power spectral density, and obtain results in line with the theoretical results. Even though the underlying spectral density in this example is not band-limited, the estimator $\widehat{\phi}_n(\cdot)$ (based on uniformly spaced samples) is found to have smaller MSE than $\widehat{\psi}_n(\cdot)$ (based on Poisson samples) for larger frequencies. The reverse order holds for smaller frequencies. This shows that there is no clear dominance of one kind of sampling over another. This finding for finite samples complements our asymptotic results.

Our results do not take anything away from the vast literature on spectrum estimation through non-uniformly sampled data. These methods may be quite appropriate when one does not have control over the sampling mechanism, when non-uniform sampling is logistically feasible and methodologically not limited, or when uniform sampling has to be avoided for a specific reason (other than its perceived inconsistency). Further, a non-uniform sampling scheme can be used where an estimator based on it is expected, either through theoretical analysis or through simulation studies, to have smaller MSE than the corresponding estimator based on uniform sampling.

We have shown in Section 2.3.2 how our theoretical results can be used to compare uniform and Poisson sampling schemes. The results compiled there may be used to make further comparisons under different constraints. For example, if there is a limit to the maximum average sampling rate and/or the maximum sample size, one may make an optimal choice of the bandwidth for fixed values of these two parameters, and then determine the corresponding MSE. In the case of $\widehat{\phi}_n(\cdot)$, the optimal choice of the bandwidth (for given sample size and sampling rate) is given by (2.35), while the choice in the case of $\widehat{\psi}_n(\cdot)$ can be derived similarly from (2.10) and (2.11) (Masry, 1978b). The best rates and constants achievable under the two sampling schemes, under appropriate constraints, may then be used to make a suitable choice of the sampling scheme.

The asymptotic arguments of this chapter do not apply when there is a hard limit on the minimum separation between two successive samples (rather than a restriction on the *average* sampling rate). This is because the sampling rate cannot be allowed to go to infinity. Thus, the scope of spectrum estimation through uniform sampling becomes rather restricted under this constraint. On the other hand, under this constraint, one cannot use Poisson sampling at all. The scope of spectrum estimation through non-uniform sampling in such situations is investigated in Chapters 4 and 5.

The existence of a consistent spectrum estimator makes it possible to conceive of interval estimators shrinking asymptotically to the true (possibly non-bandlimited) spectral density. This problem is followed up in the next chapter.

Chapter 3

Uniformly sampled non-bandlimited processes: Point and interval estimation of cross-spectra

3.1 Introduction

Many methods for constructing estimators as well as confidence intervals of bandlimited spectra have been proposed (Swanepoel and van Wyk, 1986; Brillinger, 2001; Politis et al., 1992). For uniformly sampled non-bandlimited processes, the problem of aliasing leads to estimators that are biased even if the sample size goes to infinity. Such an estimator is not an appropriate pivot for constructing confidence intervals.

Masry's work on consistent estimation of non-bandlimited spectra through Poisson sampled data (Masry, 1978a) has been used to construct asymptotic confidence intervals (Lii and Masry, 1994) on the basis of such data. These intervals shrink to the true (possibly non-bandlimited) power spectral density as the sample size goes to infinity. The fact of consistency of a spectrum estimator computed from uniformly spaced samples, proved in Chapter 2, opens up the possibility of constructing similar confidence intervals through uniform sampling, which is easier to implement than Poisson sampling. This is what we do in this chapter. We also extend the scope of the analysis so as to include cross-spectra of multivariate processes.

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

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

where

$$C_{a_1a_2}(\tau) = E\left[X_{a_1}(t+\tau)X_{a_2}(t)\right] \text{ for } a_1, a_2 \in \{1, 2, \ldots, r\}.$$

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

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

where

$$\phi_{a_1a_2}(f) = \int_{-\infty}^{\infty} C_{a_1a_2}(t)e^{-i2\pi ft}dt, \quad \text{for} -\infty < f < \infty, \quad a_1, a_2 \in \{1, 2, \ldots, r\}.$$

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

$$\widehat{\phi}_{a_1a_2}(f) = \frac{1}{n\rho_n}\sum_{j_1=1}^{n}\sum_{j_2=1}^{n} K(b_n(j_1-j_2))X_{a_1}\left(\frac{j_1}{\rho_n}\right)X_{a_2}\left(\frac{j_2}{\rho_n}\right)e^{-\frac{i2\pi f(j_1-j_2)}{\rho_n}}1_{[-\rho_n/2,\rho_n/2]}(f), \tag{3.1}$$

where n is the sample size, $K(\cdot)$ is a covariance averaging kernel, b_n is the kernel bandwidth and ρ_n is the sampling rate. Note that the estimator (2.1), considered in Chapter 2, is a special case of the above estimator for $a_1 = a_2$.

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 3.2, we establish the consistency of the spectrum estimator (3.1) for non-bandlimited processes. This is done through generalization of some results of Chapter 2 to the case of multivariate time series. These results also pave the way for computation of the asymptotic distribution of the estimator. A formal description of the confidence intervals is given in Section 3.3, which also contains some discussion on optimal rates of shrinkage of these intervals. Section 3.4 contains the proofs of the theorems presented in Sections 3.2 and 3.3. In Section 3.5, we investigate the question as to how large the sample size should be, in order that the asymptotic confidence intervals are applicable. We look for answers through a Monte Carlo simulation study and report the findings. We make some concluding remarks in Section 3.6.

3.2 Large sample results for point estimates

3.2.1 Consistency

In order to establish the consistency of the estimator $\widehat{\phi}_{a_1a_2}(\cdot)$ given in (3.1), we make a few assumptions on the process $\boldsymbol{X}$, the kernel $K(\cdot)$ and the sequences b_n and ρ_n. These are either the same as the corresponding assumptions made in Section 2.2, or similar.

Assumption 3.1. *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 3.2. *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 3.3. *The kernel bandwidth is such that $nb_n \to \infty$ as $n \to \infty$.*

Assumption 3.4. *The sampling rate is such that $\rho_n \to \infty$ and $\rho_n b_n \to 0$ as $n \to \infty$.*

Theorem 3.1. *Under Assumptions 3.1–3.4, 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 some fourth order moments/cumulants of the underlying process $\boldsymbol{X}$.

Assumption 3.5. *The fourth moment* $E\left[\left(X_{a_j}(t)\right)^4\right]$ *of the process* $\boldsymbol{X}$ *is finite for all* $a_j \in \{1, \ldots, r\}$*, while the fourth order cumulant function*

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

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

$$|Q_{a_1a_2a_3a_4}(t_1, t_2, t_3)| \leq \prod_{j=1}^{3} g_{a_j}(t_j),$$

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

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

$$\begin{pmatrix} Re\left(\widehat{\phi}_{a_1a_2}(f)\right) \\ Im\left(\widehat{\phi}_{a_1a_2}(f)\right) \end{pmatrix}, \tag{3.2}$$

where

$$\begin{aligned} &Re\left(\widehat{\phi}_{a_1a_2}(f)\right) \\ &\quad= \frac{1}{n\rho_n} \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} K(b_n(j_2 - j_1)) X_{a_1}\left(\frac{j_1}{\rho_n}\right) X_{a_2}\left(\frac{j_2}{\rho_n}\right) \\ &\qquad \times \cos\left(\frac{2\pi f(j_2 - j_1)}{\rho_n}\right) 1_{[-\rho_n/2, \rho_n/2]}(f), \end{aligned}$$

$$\begin{aligned} &Im\left(\widehat{\phi}_{a_1a_2}(f)\right) \\ &\quad= \frac{1}{n\rho_n} \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} K(b_n(j_2 - j_1)) X_{a_1}\left(\frac{j_1}{\rho_n}\right) X_{a_2}\left(\frac{j_2}{\rho_n}\right) \\ &\qquad \times \sin\left(\frac{2\pi f(j_2 - j_1)}{\rho_n}\right) 1_{[-\rho_n/2, \rho_n/2]}(f). \end{aligned}$$

Theorem 3.2. *Under Assumptions 3.1–3.5, 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:*

$$\lim_{n\to\infty} nb_n Cov\left[\begin{pmatrix} Re\left(\widehat{\phi}_{a_1a_2}(f_1)\right) \\ Im\left(\widehat{\phi}_{a_1a_2}(f_1)\right) \end{pmatrix}, \begin{pmatrix} Re\left(\widehat{\phi}_{a_3a_4}(f_2)\right) \\ Im\left(\widehat{\phi}_{a_3a_4}(f_2)\right) \end{pmatrix}\right] = \begin{bmatrix} \sigma_{11}(f_1,f_2) & \sigma_{12}(f_1,f_2) \\ \sigma_{21}(f_1,f_2) & \sigma_{22}(f_1,f_2) \end{bmatrix},$$

where

$$\begin{aligned}
\sigma_{11}(f_1,f_2) &= B\cdot Re\{\phi^*_{a_1a_3}(f_2)\phi_{a_2a_4}(f_2)+\phi^*_{a_1a_4}(f_2)\phi_{a_2a_3}(f_2)\}\,[1_{E_2}(f_1,f_2)+1_{E_3}(f_1,f_2) \\
&\quad +2\times 1_{E_4}(f_1,f_2)], \\
\sigma_{12}(f_1,f_2) &= B\cdot Im\{\phi_{a_1a_3}(f_2)\phi^*_{a_2a_4}(f_2)+\phi^*_{a_1a_4}(f_2)\phi_{a_2a_3}(f_2)\}\,[1_{E_2}(f_1,f_2)+1_{E_3}(f_1,f_2)], \\
\sigma_{21}(f_1,f_2) &= B\cdot Im\{\phi^*_{a_1a_3}(f_2)\phi_{a_2a_4}(f_2)+\phi^*_{a_1a_4}(f_2)\phi_{a_2a_3}(f_2)\}\,[1_{E_2}(f_1,f_2)-1_{E_3}(f_1,f_2)], \\
\sigma_{22}(f_1,f_2) &= B\cdot Re\{\phi_{a_1a_3}(f_2)\phi^*_{a_2a_4}(f_2)-\phi_{a_1a_4}(f_2)\phi^*_{a_2a_3}(f_2)\}\,[1_{E_2}(f_1,f_2)-1_{E_3}(f_1,f_2)], \\
B &= \frac{1}{2}\int_{-\infty}^{\infty} K^2(x)dx, \\
E_1 &= \{(f_1,f_2): f_1-f_2\neq 0,\ f_1+f_2\neq 0,\ -\infty< f_1,f_2<\infty\}, \\
E_2 &= \{(f_1,f_2): f_1-f_2=0,\ -\infty< f_1,f_2<\infty\}\setminus\{(0,0)\}, \\
E_3 &= \{(f_1,f_2): f_1+f_2=0,\ -\infty< f_1,f_2<\infty\}\setminus\{(0,0)\}, \\
E_4 &= \{(0,0)\}.
\end{aligned}$$

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 3.1 and 3.2, which are generalizations of Theorems 2.1 and 2.2 of Chapter 2 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 3.1. 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 (Brockwell and Davis, 1991). In the case of the pair $(\widehat{\phi}_{a_1a_2}(f_1), \widehat{\phi}_{a_3a_4}(f_2))$, the limiting covariance according to this notion can be easily be computed from Theorem 3.2.

3.2.2 Asymptotic normality

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

Assumption 3.5A. *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* k*th order joint cumulant denoted by*

$$Q_{a_1a_2\ldots a_k}(t_1, t_2, \ldots, t_{k-1}) = cum\left(X_{a_1}(t_1+t), X_{a_2}(t_2+t), \ldots, X_{a_{k-1}}(t_{k-1}+t), X_{a_k}(t)\right),$$

satisfies

$$|Q_{a_1a_2\ldots a_k}(t_1, t_2, \ldots, \ldots, t_{k-1})| \leq \prod_{j=1}^{k-1} g_{a_j}(t_j),$$

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

Note that Assumption 3.5A is stronger than Assumption 3.5.

Theorem 3.3. *Under Assumptions 3.1–3.4 and 3.5A, a vector of real and imaginary parts of the estimated spectra or cross-spectra converges weakly as follows.*

$$\sqrt{nb_n}\left[\begin{pmatrix} Re\{\widehat{\phi}_{a_1a_2}(f_1)\} \\ Im\{\widehat{\phi}_{a_1a_2}(f_1)\} \\ \vdots \\ Re\{\widehat{\phi}_{a_{2J-1}a_{2J}}(f_J)\} \\ Im\{\widehat{\phi}_{a_{2J-1}a_{2J}}(f_J)\} \end{pmatrix} - E\begin{pmatrix} Re\{\widehat{\phi}_{a_1a_2}(f_1)\} \\ Im\{\widehat{\phi}_{a_1a_2}(f_1)\} \\ \vdots \\ Re\{\widehat{\phi}_{a_{2J-1}a_{2J}}(f_J)\} \\ Im\{\widehat{\phi}_{a_{2J-1}a_{2J}}(f_J)\} \end{pmatrix}\right] \xrightarrow{D} N_{2J}(0, \Sigma), \qquad (3.3)$$

where $a_1, a_2, \ldots, a_{2J} \in \{1, 2, \ldots, r\}$*, and the elements of* Σ *are defined in accordance with Theorem 3.2.*

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 vector of true spectra and cross-spectra remains to be established. Note that

$$\begin{aligned}\sqrt{nb_n}&\left(\widehat{\phi}_{a_1a_2}(f)-\phi_{a_1a_2}(f)\right)\\&=\sqrt{nb_n}\left(\widehat{\phi}_{a_1a_2}(f)-E[\widehat{\phi}_{a_1a_2}(f)]\right)+\sqrt{nb_n}\left(E[\widehat{\phi}_{a_1a_2}(f)]-\phi_{a_1a_2}(f)\right).\end{aligned}\tag{3.4}$$

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 $E[\widehat{\phi}_{a_1a_2}(f)]-\phi_{a_1a_2}(f)$. These are either the same as the corresponding assumptions made in Section 2.3, or are similar to them.

Assumption 3.1A. *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 3.1B. *The power spectral density is such that, for all $a_1,a_2\in\{1,2,\ldots,r\}$ and for some $p>1$, $\lim_{f\to\infty}|f^p\phi_{a_1a_2}(f)|=A_{a_1a_2}$ for some non-negative number $A_{a_1a_2}$.*

Assumption 3.2A. *The characteristic exponent of the kernel $K(\cdot)$ is a number, for which Assumption 3.1A holds.*

Note that Assumption 3.1A is stronger than Assumption 3.1. As mentioned in Remarks 2.4 and 2.5, the numbers q and p indicate the degree of smoothness and the rate of decay, respectively, of the elements of the power spectral density matrix. (In the present case, 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 3.1B holds.

1. The real and imaginary parts of the components of the power spectral density matrix are rational functions of the form $\frac{P(f)}{Q(f)}$, 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 3.4. *Under Assumptions 3.2–3.4, 3.1A, 3.1B and 3.2A, the bias of the estimator $\widehat{\phi}_{a_1a_2}(f)$ given by (3.1), for $a_1, a_2 \in \{1, 2, \ldots, r\}$, is*

$$\begin{aligned} E[\widehat{\phi}_{a_1a_2}(f) - \phi_{a_1a_2}(f)] &= \left[-k_q \int_{-\infty}^{\infty} |t|^q C_{a_1a_2}(t) e^{-i2\pi f t} dt\right] (\rho_n b_n)^q + o\left((\rho_n b_n)^q\right) \\ &+ \left[-\int_{-\infty}^{\infty} |t| C_{a_1a_2}(t) e^{-i2\pi f t} dt\right] \left(\frac{\rho_n}{n}\right) + o\left(\frac{\rho_n}{n}\right) \\ &+ \left[A_{a_1a_2} \sum_{|l|>0} \frac{1}{|l|^p}\right] \frac{1}{(\rho_n)^p} + o\left(\frac{1}{(\rho_n)^p}\right). \end{aligned}$$

The convergence is uniform over any closed and finite interval.

Theorem 3.4 shows that the second term on the right hand side of (3.4) would go to zero if the sampling rate ρ_n satisfies additional conditions.

Assumption 3.4A. *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 3.3 holds, Assumption 3.4A is stronger than Assumption 3.4. With this assumption, the expected values of the estimators in Theorem 3.3 can be replaced by their true values.

Theorem 3.5. *Under Assumptions 3.1–3.3, 3.1A, 3.1B, 3.2A, 3.4A and 3.5A, we have the following weak convergence.*

$$\sqrt{nb_n}\left[\begin{pmatrix} Re\{\widehat{\phi}_{a_1a_2}(f_1)\} \\ Im\{\widehat{\phi}_{a_1a_2}(f_1)\} \\ \vdots \\ Re\{\widehat{\phi}_{a_{2J-1}a_{2J}}(f_J)\} \\ Im\{\widehat{\phi}_{a_{2J-1}a_{2J}}(f_J)\} \end{pmatrix} - \begin{pmatrix} Re\{\phi_{a_1a_2}(f_1)\} \\ Im\{\phi_{a_1a_2}(f_1)\} \\ \vdots \\ Re\{\phi_{a_{2J-1}a_{2J}}(f_J)\} \\ Im\{\phi_{a_{2J-1}a_{2J}}(f_J)\} \end{pmatrix}\right] \xrightarrow{D} N_{2J}(0, \Sigma),$$

where $a_1, a_2, \ldots, a_{2J} \in \{1, 2, \ldots, r\}$, and the elements of Σ are defined in accordance with Theorem 3.2.

3.3 Confidence intervals of spectra and cross-spectra

3.3.1 Construction of confidence intervals

By using the result of Theorem 3.5, 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

$$\left[\widehat{\phi}_{a_1a_1}(f) - \frac{z_{\alpha/2}\hat{\sigma}_{a_1a_1}}{nb_n}, \widehat{\phi}_{a_1a_1}(f) + \frac{z_{\alpha/2}\hat{\sigma}_{a_1a_1}}{nb_n}\right], \tag{3.5}$$

where $\hat{\sigma}^2_{a_1a_1} = [1 + 1_{\{0\}}(f)] \left(\int_{-\infty}^{\infty} K^2(x)dx\right) \widehat{\phi}^2_{a_1a_1}(f)$, which is a plug-in estimator obtained from a simplified expression of the variance given in Theorem 3.5, and $z_{\alpha/2}$ is the $(\alpha/2)^{\text{th}}$ quantile of the standard normal distribution. Theorem 3.5 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 (3.1) implies that $\hat{\sigma}_{a_1a_1}$ converges in probability to $\sigma_{a_1a_1}$. Therefore, it follows from Assumption 3.3 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 $\phi_{a_1a_2}(f)$ as a vector, a confidence region with level $1-\alpha$, for any $a_1, a_2 \in \{1, 2, \ldots, r\}$, is constructed as

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

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

$$\widehat{\Sigma} = \begin{bmatrix} \widehat{\sigma}_1^2(f) & \widehat{\sigma}_{12}(f) \\ \widehat{\sigma}_{12}(f) & \widehat{\sigma}_2^2(f) \end{bmatrix},$$

$$\widehat{\sigma}_1^2(f) = \{1 + 1_{\{0\}}(f)\}B[\widehat{\phi}_{a_1a_1}(f)\widehat{\phi}_{a_2a_2}(f) + \{Re(\widehat{\phi}_{a_1a_2}(f))\}^2 - \{Im(\widehat{\phi}_{a_1a_2}(f))\}^2], \tag{3.6}$$

$$\widehat{\sigma}_2^2(f) = \{1 - 1_{\{0\}}(f)\}B[\widehat{\phi}_{a_1a_1}(f)\widehat{\phi}_{a_2a_2}(f)\{Re(\widehat{\phi}_{a_1a_2}(f))\}^2 + \{Im(\widehat{\phi}_{a_1a_2}(f))\}^2], \tag{3.7}$$

$$\widehat{\sigma}_{12}(f) = \{1_{\{0\}}(f) - 1\}B\{2Re(\widehat{\phi}_{a_1a_2}(f))Im(\widehat{\phi}_{a_1a_2}(f))\}, \tag{3.8}$$

$$B = \frac{1}{2}\int_{-\infty}^{\infty} K^2(x)dx,$$

$\widehat{\Sigma}$ being a plug-in estimator of Σ defined in Theorem 3.5 for $J = 1$.

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

$$\left[Re[\widehat{\phi}_{a_1a_2}(f)] - \frac{z_{\alpha/2}\widehat{\sigma}_1(f)}{nb_n}, Re[\widehat{\phi}_{a_1a_2}(f)] + \frac{z_{\alpha/2}\widehat{\sigma}_1(f)}{nb_n}\right], \tag{3.9}$$

and

$$\left[Im[\widehat{\phi}_{a_1a_2}(f)] - \frac{z_{\alpha/2}\widehat{\sigma}_2(f)}{nb_n}, Im[\widehat{\phi}_{a_1a_2}(f)] + \frac{z_{\alpha/2}\widehat{\sigma}_2(f)}{nb_n}\right], \tag{3.10}$$

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 (Brillinger, 2001).

3.3.2 Optimal rate of shrinkage

It can be seen from the expressions of confidence intervals and confidence regions based on Theorem 3.5 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 3.5.

Theorem 3.6. *Under Assumptions 3.3 and 3.4A, the reciprocal of the scale factor ($\sqrt{nb_n}$) used in Theorem 3.5 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 Section 2.3 that under the assumptions of Theorems 2.4, the optimal rate of convergence for mean square consistency of the estimator (2.1) is $O\left(n^{-\frac{2pq}{p+q+2pq}}\right)$. Similar arguments would show that under similar conditions, the estimator (3.1) for any general cross-spectral density (i.e., a_1 not necessarily equal to a_2) has the same property, i.e., the optimal rate of convergence of the MSE is

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

which corresponds to the choices

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

Theorem 3.6 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 has already been observed in Remark 2.6 that for every fixed value of q, the number p, which indicates the 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. For fixed q, the best rate of weak convergence given in Theorem 3.6, 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 (Parzen, 1957).

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 large q (strong assumption) is $o\left(n^{-\frac{1}{2}}\right)$, assuming that p can be allowed to be large. For $q = 1$ (weakest possible assumption) and correspondingly large p, the rate approaches $o\left(n^{-\frac{1}{3}}\right)$. For $q = 1$ and p nearly equal to one, the rate approaches $o\left(n^{-\frac{1}{4}}\right)$.

3.4 Proofs

We denote by $K_1(\cdot)$ a function that bounds the covariance averaging kernel $K(\cdot)$ as in Assumption 3.2. Further, we denote $K_1(0)$ by M. We represent the real line by $\mathbb{R}$.

Proof of Theorem 3.1. We shall show that the bias of the estimator $\widehat{\phi}_{a_1a_2}(f)$ given by (3.1) converges to 0 uniformly over $[f_l, f_u]$ for any f_l, f_u such that $f_l < f_u$. Note that

$$E[\widehat{\phi}_{a_1a_2}(f)] = \frac{1}{\rho_n} \sum_{u=-(n-1)}^{n-1} \left(1 - \frac{|u|}{\rho_n}\right) K(b_n u) C_{a_1a_2}\left(\frac{u}{\rho_n}\right) e^{-\frac{i2\pi f u}{\rho_n}} 1_{[-\rho_n/2,\rho_n/2]}(f).$$

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

$$S_n(f,x) = \sum_{u=-(n-1)}^{n-1} \left(1 - \frac{|u|}{\rho_n}\right) K(b_n u) C_{a_1a_2}\left(\frac{u}{\rho_n}\right) e^{-\frac{i2\pi f u}{\rho_n}} 1_{[-\rho_n/2,\rho_n/2]}(f) 1_{\left(\frac{u-1}{\rho_n}, \frac{u}{\rho_n}\right]}(x).$$

Observe that $\int_{-\infty}^{\infty} S_n(f,x)dx = E[\widehat{\phi}_{a_1a_2}(f)]$. Define the function $S(\cdot)$, over $[f_l, f_u] \times \mathbb{R}$, by $S(f,x) = C_{a_1a_2}(x)e^{-i2\pi f x}$.

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 3.3 and 3.4,

$$S_n(f,x) = \left(1 - \frac{|u_n(x)|}{\rho_n}\frac{\rho_n}{n}\right) K\left(b_n \rho_n \frac{u_n(x)}{\rho_n}\right) C_{a_1a_2}\left(\frac{u_n(x)}{\rho_n}\right) e^{-\frac{i2\pi f u_n(x)}{\rho_n}} 1_{[-\rho_n/2,\rho_n/2]}(f).$$

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

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} S_n(f,x)dx = \int_{-\infty}^{\infty} S(f,x)dx, \tag{3.11}$$

uniformly over $[f_l,\ f_u]$.

Observe that $\int_{-\infty}^{\infty} S(f,t)dt = \phi_{a_1a_2}(f)$, which is continuous. By virtue of the continuity of the limiting function, (3.11) is equivalent to proving that $\int_{-\infty}^{\infty} S_n(f,x)dx$ converges continuously over this interval (Resnick, 1987), i.e., for any sequence $f_n \to f$,

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} S_n(f_n,x)dx = \int_{-\infty}^{\infty} S(f,x)dx, \tag{3.12}$$

where $f_n, f \in [f_l,\ f_u]$.

By continuity of the function $S_n(f,x)$ with respect to x and f, we have from Assumptions 3.3 and 3.4, for any fixed x,

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

Note that from Assumptions 3.1 and 3.2, we have the dominance

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

where $g_{a_1a_2}(\cdot)$ is the function described in Assumption 3.1. Thus, by applying the DCT, we have (3.12).

Hence, $E[\widehat{\phi}_{a_1a_2}(f)] \to \phi(f)$ uniformly on $[f_l,\, f_u]$. □

Proof of Theorem 3.2. We begin by calculating the covariance between the estimators $Re(\widehat{\phi}_{a_1a_2}(\cdot))$ and $Re(\widehat{\phi}_{a_3a_4}(\cdot))$.

$$\begin{aligned}
&Cov\left[Re(\widehat{\phi}_{a_1a_2}(f_1)), Re(\widehat{\phi}_{a_3a_4}(f_2))\right]\\
=&\frac{1}{(n\rho_n)^2}\sum_{j_1=1}^{n}\sum_{j_2=1}^{n}\sum_{j_3=1}^{n}\sum_{j_4=1}^{n} K(b_n(j_2-j_1))K(b_n(j_4-j_3))\\
&\times Cov\left[X_{a_1}\left(\frac{j_1}{\rho_n}\right)X_{a_2}\left(\frac{j_2}{\rho_n}\right), X_{a_3}\left(\frac{j_3}{\rho_n}\right)X_{a_4}\left(\frac{j_4}{\rho_n}\right)\right]\\
&\times \cos\left(\frac{2\pi f_1(j_2-j_1)}{\rho_n}\right)\cos\left(\frac{2\pi f_2(j_4-j_3)}{\rho_n}\right)\\
=&\frac{1}{(n\rho_n)^2}\sum_{j_1=1}^{n}\sum_{j_2=1}^{n}\sum_{j_3=1}^{n}\sum_{j_4=1}^{n} K(b_n(j_2-j_1))K(b_n(j_4-j_3))\left[C_{a_1a_3}\left(\frac{j_1-j_3}{\rho_n}\right)C_{a_2a_4}\left(\frac{j_2-j_4}{\rho_n}\right)\right.\\
&\left.+C_{a_1a_4}\left(\frac{j_1-j_4}{\rho_n}\right)C_{a_2a_3}\left(\frac{j_2-j_3}{\rho_n}\right)+Q_{a_1a_2a_3a_4}\left(\frac{j_1-j_4}{\rho_n}, \frac{j_2-j_4}{\rho_n}, \frac{j_3-j_4}{\rho_n}\right)\right]\\
&\times \cos\left(\frac{2\pi f_1(j_1-j_2)}{\rho_n}\right)\cos\left(\frac{2\pi f_2(j_3-j_4)}{\rho_n}\right)\\
=&T_1(f_1,f_2)+T_2(f_1,f_2)+T_3(f_1,f_2),
\end{aligned}$$

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

Now, consider the function $T_1(f_1, f_2)$. By using the transformations $u_1 = j_1 - j_2$, $u_2 = j_1 - j_3$ and $u_3 = j_2 - j_4$, we have

$$\begin{aligned}
T_1(f_1,f_2) =&\frac{1}{(n\rho_n)^2}\sum_{j_1=1}^{n}\sum_{u_1=j_1-1}^{j_1-n}\sum_{u_2=j_1-1}^{n-j_1}\sum_{u_3=j_1-1-u_1}^{j_1-n-u_1} K(b_nu_1)K(b_n(u_1-u_2+u_3))\\
&\times C_{a_1a_3}\left(\frac{u_2}{\rho_n}\right)C_{a_2a_4}\left(\frac{u_3}{\rho_n}\right)\cos\left(\frac{2\pi f_1u_1}{\rho_n}\right)\cos\left(\frac{2\pi f_2(u_1-u_2+u_3)}{\rho_n}\right).
\end{aligned}$$

The range of the four summations on the right hand side is described by the set of inequalities $1 \le j_1 \le n$ and $j_1 - n \le u_1, u_2, u_1 + u_3 \le j_1 - 1$, which is equivalent to the inequalities $-(n-1) \le u_1, u_2, u_1 + u_3 \le (n-1)$ and $\max\{u_1, u_2, u_1 + u_3\} + 1 \le j_1 \le \min\{u_1, u_2, u_1 + u_3\}$. Therefore, the expression for $T_1(f_1, f_2)$ simplifies to

$$\frac{1}{n\rho_n^2} \sum_{u_1=-(n-1)}^{(n-1)} \sum_{u_2=-(n-1)}^{(n-1)} \sum_{u_3=-(n-1)-u_1}^{(n-1)-u_1} U_n(u_1, u_2, u_3) K(b_n u_1) K(b_n(u_1 - u_2 + u_3))$$
$$\times C_{a_1 a_3}\left(\frac{u_2}{\rho_n}\right) C_{a_2 a_4}\left(\frac{u_3}{\rho_n}\right) \cos\left(\frac{2\pi f_1 u_1}{\rho_n}\right) \cos\left(\frac{2\pi f_2 (u_1 - u_2 + u_3)}{\rho_n}\right),$$

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

$$\begin{aligned}
&T_1(f_1, f_2)\\
=&\frac{1}{4n\rho_n^2} \sum_{u_1=-(n-1)}^{(n-1)} \sum_{u_2=-(n-1)}^{(n-1)} \sum_{u_3=-(n-1)-u_1}^{(n-1)-u_1} U_n(u_1, u_2, u_3) K(b_n u_1) K(b_n(u_1 - u_2 + u_3))\\
&\times C_{a_1 a_3}\left(\frac{u_2}{\rho_n}\right) C_{a_2 a_4}\left(\frac{u_3}{\rho_n}\right) \Bigg\{ e^{-i\frac{2\pi(f_1 - f_2)u_1}{\rho_n}} e^{-i\frac{2\pi f_2 u_2}{\rho_n}} e^{i\frac{2\pi f_2 u_3}{\rho_n}} + e^{i\frac{2\pi(f_1 - f_2)u_1}{\rho_n}} e^{i\frac{2\pi f_2 u_2}{\rho_n}} e^{-i\frac{2\pi f_2 u_3}{\rho_n}}\\
&+ e^{i\frac{2\pi(f_1 + f_2)u_1}{\rho_n}} e^{-i\frac{2\pi f_2 u_2}{\rho_n}} e^{i\frac{2\pi f_2 u_3}{\rho_n}} + e^{-i\frac{2\pi(f_1 + f_2)u_1}{\rho_n}} e^{i\frac{2\pi f_2 u_2}{\rho_n}} e^{-i\frac{2\pi f_2 u_3}{\rho_n}} \Bigg\}\\
=&T_{11}(f_1, f_2) + T_{12}(f_1, f_2) + T_{13}(f_1, f_2) + T_{14}(f_1, f_2), \qquad (3.13)
\end{aligned}$$

where the four terms correspond to the four summands appearing within braces in the last factor on the right hand side of (3.13).

By using the results of Lemmas 3.1 and 3.2 given below, we have the convergence

$$\lim_{n\to\infty} n b_n T_{11}(f_1, f_2) = \frac{1}{2} B \phi_{a_1 a_3}(f_2) \phi^*_{a_2 a_4}(f_2) 1_{E_2 \cup E_4}(f_1, f_2),$$

and similar arguments show that

$$\begin{aligned}
\lim_{n\to\infty} n b_n T_{12}(f_1, f_2) &= \frac{1}{2} B \phi^*_{a_1 a_3}(f_2) \phi_{a_2 a_4}(f_2) 1_{E_2 \cup E_4}(f_1, f_2),\\
\lim_{n\to\infty} n b_n T_{13}(f_1, f_2) &= \frac{1}{2} B \phi_{a_1 a_3}(f_2) \phi^*_{a_2 a_4}(f_2) 1_{E_3 \cup E_4}(f_1, f_2),\\
\lim_{n\to\infty} n b_n T_{14}(f_1, f_2) &= \frac{1}{2} B \phi^*_{a_1 a_3}(f_2) \phi_{a_2 a_4}(f_2) 1_{E_3 \cup E_4}(f_1, f_2).
\end{aligned}$$

For the function $T_2(f_1, f_2)$, one can similarly use the transformations $u_1 = j_1 - j_2$, $u_2 = j_1 - j_4$ and $u_3 = j_2 - j_3$, interchange the order of summation and expand the cosine functions in terms of complex exponentials to obtain

$$
\begin{aligned}
&T_2(f_1, f_2)\\
=&\frac{1}{n\rho_n^2}\sum_{u_1=-(n-1)}^{(n-1)}\sum_{u_2=-(n-1)}^{(n-1)}\sum_{u_3=-(n-1)-u_1}^{(n-1)-u_1} U_n(u_1,u_2,u_3)K(b_nu_1)K(b_n(-u_1+u_2-u_3))\\
&\times C_{a_1a_4}\left(\frac{u_2}{\rho_n}\right)C_{a_2a_3}\left(\frac{u_3}{\rho_n}\right)\cos\left(\frac{2\pi f_1u_1}{\rho_n}\right)\cos\left(\frac{2\pi f_2(-u_1+u_2-u_3)}{\rho_n}\right)\\
=&\frac{1}{4n\rho_n^2}\sum_{u_1=-(n-1)}^{(n-1)}\sum_{u_2=-(n-1)}^{(n-1)}\sum_{u_3=-(n-1)-u_1}^{(n-1)-u_1} U_n(u_1,u_2,u_3)K(b_nu_1)K(b_n(-u_1+u_2-u_3))\\
&\times C_{a_1a_4}\left(\frac{u_2}{\rho_n}\right)C_{a_2a_3}\left(\frac{u_3}{\rho_n}\right)\Big\{e^{-i\frac{2\pi(f_1-f_2)u_1}{\rho_n}}e^{-i\frac{2\pi f_2u_2}{\rho_n}}e^{i\frac{2\pi f_2u_3}{\rho_n}}+e^{i\frac{2\pi(f_1-f_2)u_1}{\rho_n}}e^{i\frac{2\pi f_2u_2}{\rho_n}}e^{-i\frac{2\pi f_2u_3}{\rho_n}}\\
&+e^{i\frac{2\pi(f_1+f_2)u_1}{\rho_n}}e^{-i\frac{2\pi f_2u_2}{\rho_n}}e^{i\frac{2\pi f_2u_3}{\rho_n}}+e^{-i\frac{2\pi(f_1+f_2)u_1}{\rho_n}}e^{i\frac{2\pi f_2u_2}{\rho_n}}e^{-i\frac{2\pi f_2u_3}{\rho_n}}\Big\}\\
&=T_{21}(f_1,f_2)+T_{22}(f_1,f_2)+T_{23}(f_1,f_2)+T_{24}(f_1,f_2).
\end{aligned}
$$

By using similar arguments as in the case of $nb_nT_{11}(f_1, f_2)$, it can be shown that

$$
\begin{aligned}
\lim_{n\to\infty} nb_nT_{21}(f_1,f_2) &= \frac{1}{2}B\phi_{a_1a_4}(f_2)\phi^*_{a_2a_3}(f_2)1_{E_2\cup E_4}(f_1,f_2),\\
\lim_{n\to\infty} nb_nT_{22}(f_1,f_2) &= \frac{1}{2}B\phi^*_{a_1a_4}(f_2)\phi_{a_2a_3}(f_2)1_{E_2\cup E_4}(f_1,f_2),\\
\lim_{n\to\infty} nb_nT_{23}(f_1,f_2) &= \frac{1}{2}B\phi_{a_1a_4}(f_2)\phi^*_{a_2a_3}(f_2)1_{E_3\cup E_4}(f_1,f_2),\\
\lim_{n\to\infty} nb_nT_{24}(f_1,f_2) &= \frac{1}{2}B\phi^*_{a_1a_4}(f_2)\phi_{a_2a_3}(f_2)1_{E_3\cup E_4}(f_1,f_2).
\end{aligned}
$$

Finally, for the term $T_3(f_1, f_2)$, we use the transformations $u_1 = j_1 - j_4$, $u_2 = j_2 - j_4$ and $u_3 = j_3 - j_4$ and interchange the order of summations to have

$$
\begin{aligned}
T_3(f_1,f_2) =&\frac{1}{(n\rho_n)^2}\sum_{u_1=-(n-1)}^{(n-1)}\sum_{u_2=-(n-1)}^{n-1}\sum_{u_3=-(n-1)}^{(n-1)}\{n-\min_i(u_i)+\max_i(u_i)\}K(b_n(u_1-u_2))\\
&\times K(b_nu_3)Q_{a_1a_2a_3a_4}\left(\frac{u_1}{\rho_n},\frac{u_2}{\rho_n},\frac{u_3}{\rho_n}\right)\cos\left(\frac{2\pi f_1(u_1-u_2)}{\rho_n}\right)\cos\left(\frac{2\pi f_2u_3}{\rho_n}\right).
\end{aligned}
$$

From Assumptions 3.2 and 3.5, we have

$$nb_n|T_3(f_1,f_2)| \\ \leq \ {}_nb_nM^2 \sum_{u_1=-(n-1)}^{n-1} \sum_{u_2=-(n-1)}^{n-1} \sum_{u_2=-(n-1)}^{n-1} g_{a_1}\left(\frac{u_1}{\rho_n}\right) g_{a_2}\left(\frac{u_2}{\rho_n}\right) g_{a_3}\left(\frac{u_3}{\rho_n}\right) \frac{1}{\rho_n^3}. \quad (3.14)$$

Now, consider the function $S_n(\cdot)$ defined over $\mathbb{R}$ as $S_n(x) = \sum_{u_1=-(n-1)}^{n-1} g_{a_1}\left(\frac{u_1}{\rho_n}\right) 1_{\left(\frac{u_1-1}{\rho_n}, \frac{u_1}{\rho_n}\right]}(x)$. 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 the 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 $nb_nT_3(f_1,f_2)$ given by (3.14) is $O(\rho_n b_n)$. Assumption 3.4 ensures that $nb_nT_3(f_1,f_2)$ converges to zero uniformly.

By combining all these terms, we have the convergence of $nb_nCov\left[Re(\widehat{\phi}_{a_1a_2}(f_1)), Re(\widehat{\phi}_{a_3a_4}(f_2))\right]$ as given in the theorem. Convergence of the other three covariances follow from a similar argument. □

Lemma 3.1. *For $f_1 - f_2 = 0$, the function $T_{11}(f_1,f_2)$ converges as follows.*

$$\lim_{n\to\infty} nb_nT_{11}(f_1,f_2) = \frac{1}{2}B\phi_{a_1a_3}(f_2)\phi^*_{a_2a_4}(f_2).$$

The convergence is uniform on any compact subset of the set

$$E = \{(f_1,f_2) : f_1 - f_2 = 0, \ -\infty < f_1, f_2 < \infty\}.$$

Proof of Lemma 3.1. Consider a compact subset E' of the set E. Consider the simple function $S_n(\cdot)$, defined over $E' \times \mathbb{R}^3$ by

$$S_n(f_1,f_2,x_1,x_2,x_3) \\ = \sum_{u_1=-(n-1)}^{(n-1)} \sum_{u_2=-(n-1)}^{(n-1)} \sum_{u_3=-(n-1)-u_1}^{(n-1)-u_1} U_n(u_1,u_2,u_3)K(b_nu_1)K(b_n(u_1-u_2+u_3))C_{a_1a_3}\left(\frac{u_2}{\rho_n}\right) \\ \times e^{-i\frac{2\pi f_2 u_2}{\rho_n}} C_{a_2a_4}\left(\frac{u_3}{\rho_n}\right) e^{i\frac{2\pi f_2 u_3}{\rho_n}} 1_{((u_1-1)b_n, u_1b_n]}(x_1) 1_{\left(\frac{u_2-1}{\rho_n}, \frac{u_2}{\rho_n}\right]}(x_2) 1_{\left(\frac{u_3-1}{\rho_n}, \frac{u_3}{\rho_n}\right]}(x_3),$$

so that

$$nb_nT_{11}(f_1,f_2)=\frac{1}{4}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}S_n(f_1,f_2,x_1,x_2,x_3)dx_1dx_2dx_3.$$

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, $\rho_n x_2$ and $\rho_n x_3$, 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 $b_nu_{1n}(x_1)\to x_1$, $\frac{u_{2n}(x_2)}{\rho_n}\to x_2$, $\frac{u_{3n}(x_3)}{\rho_n}\to x_3$ 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

$$\begin{aligned}&S_n(f_1,f_2,x_1,x_2,x_3)\\&=U_n(u_{1n}(x_1),u_{2n}(x_2),u_{3n}(x_3))K(b_nu_{1n}(x_1))K(b_n(u_{1n}(x_1)-u_{2n}(x_2)+u_{3n}(x_3)))\\&\quad\times C_{a_1a_3}\left(\frac{u_{2n}(x_2)}{\rho_n}\right)e^{-i\frac{2\pi f_2u_{2n}(x_2)}{\rho_n}}C_{a_2a_4}\left(\frac{u_{3n}(x_3)}{\rho_n}\right)e^{i\frac{2\pi f_2u_{3n}(x_3)}{\rho_n}}.\end{aligned}\tag{3.15}$$

Observe that, under Assumptions 3.1,3.3 and 3.4, the function $S_n(f_1,f_2,x_1,x_2,x_3)$ converges to the function $S(\cdot)$, defined over $E'\times\mathbb{R}^3$ by

$$S(f_1,f_2,x_1,x_2,x_3)=K^2(x_1)C_{a_1a_3}(x_2)\,e^{-i2\pi f_2x_2}C_{a_2a_4}(x_3)\,e^{i2\pi f_2x_3}.$$

Note that $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}S(f_1,f_2,x_1,x_2,x_3)dx_1dx_2dx_3$ is a continuous function in (f_1,f_2). As in the proof of Theorem 3.1, we prove the convergence of the left hand side of (3.15) uniformly on E', by showing that for any sequence $(f_{1n},f_{2n})\to(f_1,f_2)$,

$$\begin{aligned}\lim_{n\to\infty}&\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}S_n(f_{1n},f_{2n},x_1,x_2,x_3)dx_1dx_2dx_3\\&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}S(f_1,f_2,x_1,x_2,x_3)dx_1dx_2dx_3.\end{aligned}$$

for $(f_{1n},f_{2n}),(f_1,f_2)\in E'$. The latter convergence follows, through Assumption 3.1 and 3.2 and the DCT, from the dominance

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

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 3.2. *For $f_1 - f_2 \neq 0$, the function $nb_nT_{11}(f_1, f_2)$ converges to zero. The convergence is uniform on any compact subset of the set E_1 given by*

$$E = \{(f_1, f_2) : f_1 - f_2 \neq 0, \ -\infty < f_1, f_2 < \infty\}.$$

Proof of Lemma 3.2. 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

$$\begin{aligned}
&S_n(f_1, f_2, x_1, x_2, x_3) \\
&= \sum_{u_1=-(n-1)}^{(n-1)} \sum_{u_2=-(n-1)}^{(n-1)} \sum_{u_3=-(n-1)-u_1}^{(n-1)-u_1} U_n(u_1, u_2, u_3) K(b_n u_1) K(b_n(u_1 - u_2 + u_3)) \\
&\quad \times e^{-i\frac{u_1 2\pi(f_1-f_2)}{\rho_n}} C_{a_1a_3}\left(\frac{u_2}{\rho_n}\right) e^{-i\frac{2\pi f_2 u_2}{\rho_n}} C_{a_2a_4}\left(\frac{u_3}{\rho_n}\right) e^{i\frac{2\pi f_2 u_3}{\rho_n}} \\
&\quad \times 1_{((u_1-1)b_n, u_1 b_n)]}(x_1) 1_{\left(\frac{(u_2-1)}{\rho_n}, \frac{u_2}{\rho_n}\right]}(x_2) 1_{\left(\frac{(u_3-1)}{\rho_n}, \frac{u_3}{\rho_n}\right]}(x_3),
\end{aligned}$$

so that

$$nb_nT_{11}(f_1, f_2) = \frac{1}{4} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S_n(f_1, f_2, x_1, x_2, x_3) dx_1 dx_2 dx_3.$$

An argument similar to that used in the proof of Lemma 3.1 shows that for $(x_1, x_2, x_3) \in \mathbb{R}^3$ and sufficiently large n,

$$\begin{aligned}
&S_n(f_1, f_2, x_1, x_2, x_3) \\
&= U_n(u_{1n}(x_1), u_{2n}(x_2), u_{3n}(x_3)) K(b_n u_{1n}(x_1)) K(b_n(u_{1n}(x_1) - u_{2n}(x_2) + u_{3n}(x_3))) \\
&\quad \times e^{-i\frac{2\pi(f_1-f_2)u_{1n}(x_1)}{\rho_n}} C_{a_1a_3}\left(\frac{u_{2n}(x_2)}{\rho_n}\right) e^{-i\frac{2\pi f_2 u_{2n}(x_2)}{\rho_n}} C_{a_2a_4}\left(\frac{u_{3n}(x_3)}{\rho_n}\right) e^{i\frac{2\pi f_2 u_{3n}(x_3)}{\rho_n}},
\end{aligned}$$

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, $\rho_n x_2$ and $\rho_n x_3$, respectively.

For obtaining the uniform convergence of $nb_nT_{11}(f_1,f_2)$, consider

$$\begin{aligned}
&\sup_{(f_1,f_2)\in E'}\left|\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S_n(f_1,f_2,x_1,x_2,x_3)dx_1dx_2dx_3\right| \\
&\quad\le \sup_{(f_1,f_2)\in E'}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left| S_n(f_1,f_2,x_1,x_2,x_3)-g_n(f_1,f_2,x_1,x_2,x_3)\right| dx_1dx_2dx_3 \\
&\qquad+\sup_{(f_1,f_2)\in E'}\left|\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_n(f_1,f_2,x_1,x_2,x_3)dx_1dx_2dx_3\right|, \qquad (3.16)
\end{aligned}$$

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

$$g_n(f_1,f_2,x_1,x_2,x_3)=K^2(x_1)e^{-i\frac{x_1 2\pi(f_1-f_2)}{b_n\rho_n}}C_{a_1a_3}(x_2)\,e^{-i2\pi f_2x_2}C_{a_2a_4}(x_3)\,e^{i2\pi f_2x_3}.$$

We will show the uniform convergence of the right hand side of (3.16) by considering the two terms separately. For the first term, we follow the route taken in the proof of Theorem 3.1, i.e., show that for any sequence $(f_{1n},f_{2n})\to(f_1,f_2)$,

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

for $(f_{1n},f_{2n}),(f_1,f_2)\in E'$. For this purpose, we write the above integral as

$$\begin{aligned}
&\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left| S_n(f_{1n},f_{2n},x_1,x_2,x_3)-g_n(f_{1n},f_{2n},x_1,x_2,x_3)\right| dx_1dx_2dx_3 \\
&\quad\le\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left| S_n(f_{1n},f_{2n},x_1,x_2,x_3)-G_n(f_{1n},f_{2n},x_1,x_2,x_3)\right| dx_1dx_2dx_3 \\
&\qquad+\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left| G_n(f_{1n},f_{2n},x_1,x_2,x_3)-g_n(f_{1n},f_{2n},x_1,x_2,x_3)\right| dx_1dx_2dx_3,
\end{aligned} \qquad (3.17)$$

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

$$G_n(f_1,f_2,x_1,x_2,x_3)=K^2(x_1)e^{-i\frac{u_{1n}(x_1)b_n2\pi(f_1-f_2)}{b_n\rho_n}}C_{a_1a_3}(x_2)\,e^{-i2\pi f_2x_2}C_{a_2a_4}(x_3)\,e^{i2\pi f_2x_3}.$$

Now, observe that

$$\begin{aligned}
&|S_n(f_{1n},f_{2n},x_1,x_2,x_3)-G_n(f_{1n},f_{2n},x_1,x_2,x_3)| \\
&\qquad\le M\left|e^{-i\frac{u_{1n}(x_1)b_n2\pi(f_{1n}-f_{2n})}{b_n\rho_n}}\alpha_n(f_{1n},f_{2n},x_1,x_2,x_3)\right|,
\end{aligned}$$

where

$$\begin{aligned}&\alpha_n(f_{1n}, f_{2n}, x_1, x_2, x_3)\\ =&U_n(u_{1n}(x_1), u_{2n}(x_2), u_{3n}(x_3))K(b_n u_{1n}(x_1))K(b_n(u_{1n}(x_1) - u_{2n}(x_2) + u_{3n}(x_3)))\\ &\times C_{a_1a_3}\left(\frac{u_{2n}(x_2)}{\rho_n}\right) e^{-i\frac{2\pi f_{2n}u_{2n}(x_2)}{\rho_n}} C_{a_2a_4}\left(\frac{u_{3n}(x_3)}{\rho_n}\right) e^{i\frac{2\pi f_{2n}u_{3n}(x_3)}{\rho_n}}\\ &- K^2(x_1)C_{a_1a_3}(x_2)\, e^{-i2\pi f_{2n}x_2} C_{a_2a_4}(x_3)\, e^{i2\pi f_{2n}x_3}.\end{aligned}$$

Since $\alpha_n(f_n, x, t, t^{'}) \to 0$ as $n \to \infty$, it follows that

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

From Assumption 3.1 and 3.2, we have the dominance

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

By applying the DCT, we have

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

For the second term on the right hand side of (3.17), observe that for any fixed x_1,

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

Thus,

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

and so

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

From Assumption 3.1 and 3.2, we have the dominance

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

which leads us, through another use of the DCT, to the convergence of the second integral of (3.17). This establishes that the first term on the right hand side of (3.16) converges to 0. We only have to deal with the second term. Let

$$s_n(f_1, f_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_n(f_1, f_2, x_1, x_2, x_3)dx_1dx_2dx_3.$$

In order to establish the uniform convergence of $s_n(\cdot)$ over E', it is enough to show that $s_n(f_{1n}, f_{2n}) \to 0$ for any sequence $(f_{1n}, f_{2n}) \to (f_1, f_2)$, where $(f_{1n}, f_{2n}), (f_1, f_2) \in E'$. By using the Reimann-Lebesgue lemma, we have $s_n(f_1, f_2) \to 0$. Thus, the second term on the right hand side of (3.16) also converges to 0. Hence, $nb_nT_{11}(f_1, f_2)$ converges to 0 uniformly on E' as $n \to \infty$. □

In order to prove Theorem 3.3, 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 (2.4) may be complex-valued.

Lemma 3.3. *Under the Assumptions 3.1–3.4 and 3.5A, for $L > 2$, the Lth order joint cumulant of the vector $\left(\widehat{\phi}_{a_1a_2}(f_1), \ \ldots, \widehat{\phi}_{a_{2L-1}a_{2L}}(f_L)\right)$ for $a_1, a_2, \ldots, a_{2L} \in \{1, 2, \ldots, r\}$ is bounded from above as follows.*

$$\left|cum\left(\widehat{\phi}_{a_1a_2}(f_1), \ldots, \widehat{\phi}_{a_{2L-1}a_{2L}}(f_L)\right)\right| \leq Q \cdot (nb_n)^{-(L-1)}, \tag{3.18}$$

where the constant Q does not depend on $f_1, \ldots, f_L$.

Proof of Lemma 3.3. $cum(\widehat{\phi}_{a_1a_2}(f_1), \widehat{\phi}_{a_3a_4}(f_2), \ldots, \widehat{\phi}_{a_{2L-1}a_{2L}}(f_L))$ can be written as

$$\begin{aligned}
&cum(\widehat{\phi}_{a_1a_2}(f_1), \widehat{\phi}_{a_3a_4}(f_2), \ldots, \widehat{\phi}_{a_{2L-1}a_{2L}}(f_L)) \\
=&\frac{1}{(n\rho_n)^L}\sum_{s_1=1}^{n}\sum_{s_2=1}^{n}\cdots\sum_{s_{2L-1}=1}^{n}\sum_{s_{2L}=1}^{n} K(b_n(s_1-s_2))\ldots K(b_n(s_{2L-1}-s_{2L}))e^{-\frac{i2\pi f_1(s_1-s_2)}{\rho_n}} \\
&\times\cdots\times e^{-\frac{i2\pi f_L(s_{2L-1}-s_{2L})}{\rho_n}} cum\left(X_{a_1}\left(\frac{s_1}{\rho_n}\right)X_{a_2}\left(\frac{s_2}{\rho_n}\right), \ldots, X_{a_{2L-1}}\left(\frac{s_{2L-1}}{\rho_n}\right)X_{a_{2L}}\left(\frac{s_{2L}}{\rho_n}\right)\right).
\end{aligned} \tag{3.19}$$

It follows that

$$|cum(\widehat{\phi}_{a_1a_2}(f_1), \widehat{\phi}_{a_3a_4}(f_2), \ldots, \widehat{\phi}_{a_{2L-1}a_{2L}}(f_L))|$$

$$\leq \frac{1}{(n\rho_n)^L} \sum_{s_1=1}^{n} \sum_{s_2=1}^{n} \cdots \sum_{s_{2L-1}=1}^{n} \sum_{s_{2L}=1}^{n} |K(b_n(s_1 - s_2)) \ldots K(b_n(s_{2L-1} - s_{2L}))|$$

$$\times \left| cum\left(X_{a_1}\left(\frac{s_1}{\rho_n}\right) X_{a_2}\left(\frac{s_2}{\rho_n}\right), \ldots, X_{a_{2L-1}}\left(\frac{s_{2L-1}}{\rho_n}\right) X_{a_{2L}}\left(\frac{s_{2L}}{\rho_n}\right)\right)\right|.$$

Now,

$$cum\left(X_{a_1}\left(\frac{s_1}{\rho_n}\right) X_{a_2}\left(\frac{s_2}{\rho_n}\right), \ldots, X_{a_{2L-1}}\left(\frac{s_{2L-1}}{\rho_n}\right) X_{a_{2L}}\left(\frac{s_{2L}}{\rho_n}\right)\right)$$

$$= \sum_{\boldsymbol{\nu}} Q_{a_{j_{11}}a_{j_{12}}\ldots a_{j_{1k_1}}}\left(\frac{s_{j_{11}} - s'_1}{\rho_n}, \ldots, \frac{s_{j_{1,k_1-1}} - s'_1}{\rho_n}\right) \times \cdots$$

$$\times Q_{a_{j_{P1}}a_{j_{P2}}\ldots a_{j_{Pk_P}}}\left(\frac{s_{j_{P1}} - s'_P}{\rho_n}, \ldots, \frac{s_{j_{P,k_P-1}} - s'_P}{\rho_n}\right),$$

where the summation is over all *indecomposable* (Brillinger, 2001; Leonov and Shiryayev, 1959) partitions $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_P)$, such that $\nu_\vartheta = (j_{\vartheta 1}, \ldots, j_{\vartheta k_\vartheta})$, $\vartheta = 1, \ldots, P$, of the table

1	2
3	4
⋮	⋮
2L-1	2L

and $s'_\vartheta = s_{j_{\vartheta k_\vartheta}}$, $\vartheta = 1, \ldots, P$. Since the partition $\boldsymbol{\nu}$ is indecomposable, we have

$$s_{j_{\vartheta l}} - s'_\vartheta \neq s_{2m} - s_{2m-1}$$

for $l = 1, \ldots, k_\vartheta;\ \vartheta = 1, \ldots, P;\ m = 1, \ldots, L.$

Define

$$u_{j_{\vartheta l}} = s_{j_{\vartheta l}} - s'_\vartheta;\ l = 1, \ldots, k_\vartheta; \vartheta = 1, \ldots, P.$$

Note that $u_{j_{\vartheta k_\vartheta}} = 0$ for $\vartheta = 1, \ldots, P$. Then the joint cumulant of ($\widehat{\phi}_{a_1a_2}(f_1)$, $\widehat{\phi}_{a_3a_4}(f_2)$, $\ldots$, $\widehat{\phi}_{a_{2L-1}a_{2L}}(f_L)$) given by (3.19) is absolutely bounded by

$$\frac{1}{(n\rho_n)^L}\sum_{\boldsymbol{\nu}}\sum_{s'_1=1}^{n}\sum_{u_{j_{11}}=-(s'_1-1)}^{n-s'_1}\cdots\sum_{u_{j_{1,k_1-1}}=-(s'_1-1)}^{n-s'_1}\cdots\sum_{s'_P=1}^{n}\sum_{u_{j_{P1}}=-(s'_P-1)}^{n-s'_P}\cdots\sum_{u_{j_{P,k_P-1}}=-(s'_P-1)}^{n-s'_P}$$

$$\left| K[b_n(u_1+s'_{\vartheta_1}-u_2-s'_{\vartheta_2})]\times\cdots\times K[b_n(u_{2L-1}+s'_{\vartheta_{2L-1}}-u_{2L}-s'_{\vartheta_{2L}})]\right|$$

$$\times\left|Q_{a_{j_{11}}a_{j_{12}}\ldots a_{j_{1k_1}}}\left(\frac{u_{j_{11}}}{\rho_n},\ldots,\frac{u_{j_{1,k_1-1}}}{\rho_n}\right)\times\cdots\times Q_{a_{j_{P1}}a_{j_{P2}}\ldots a_{j_{Pk_P}}}\left(\frac{u_{j_{P1}}}{\rho_n},\ldots,\frac{u_{j_{P,k_P-1}}}{\rho_n}\right)\right|, \tag{3.20}$$

where ϑ_m is that member of the set $\{1,2,\ldots,P\}$ which satisfies $s_m\in\nu_{\vartheta_m}$ for $m=1,\ldots,L$.

We will now show that the set $A=\{s'_{\vartheta_1}-s'_{\vartheta_2},\ldots,s'_{\vartheta_{2L-1}}-s'_{\vartheta_{2L}}\}$ has $P-1$ linearly independent elements. Note that the set A consists of differences of pairs of elements of the set $\{s'_1,s'_2,\ldots,s'_P\}$. 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

$$A_1=\left\{s'_{\vartheta_{2k_1-1}}-s'_{\vartheta_{2k_1}},s'_{\vartheta_{2k_2-1}}-s'_{\vartheta_{2k_2}},\ldots,s'_{\vartheta_{2k_{P-j}-1}}-s'_{\vartheta_{2k_{P-j}}}\right\},$$

where $k_1,\ldots,k_{P-j}\in\{1,2,\ldots,L\}$. Let, if possible, $j>1$, and consider a difference $s'_{l_1}-s'_{l_2}$ for $l_1,l_2\in\{1,2,\ldots,P\}$ 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* (Leonov and Shiryayev, 1959). Therefore, there exists an index set $\{\alpha_1,\alpha_2,\ldots,\alpha_r\}$ with $r\geq 2$, which is a proper subset of $\{1,2,\ldots,P\}$, such that $\alpha_1=l_1$, $\alpha_r=l_2$ and the pairs $(\nu_{\alpha_1},\nu_{\alpha_2}),(\nu_{\alpha_2},\nu_{\alpha_3}),\ldots,(\nu_{\alpha_{r-1}},\nu_{\alpha_r})$ are *hook* (Leonov and Shiryayev, 1959). Consequently, there exist indices $j_1,\ldots,j_{r-1}\in\{1,\ldots,L\}$ such that for $m=1,\ldots,r-1$, one of the points s_{2j_m-1} and s_{2j_m} belongs to ν_{α_m} and the other belongs to $\nu_{\alpha_{m+1}}$. It follows that for $m=1,\ldots,r-1$, $(s'_{\vartheta_{2j_m-1}}-s'_{\vartheta_{2j_m}})$ is in A, and hence, they can be written as linear combinations of the members of A_1. Note that for $m=1,\ldots,r-1$, $(s'_{\alpha_{m-1}}-s'_{\alpha_m})$ is equal to either $(s'_{\vartheta_{2j_m-1}}-s'_{\vartheta_{2j_m}})$ or $-(s'_{\vartheta_{2j_m-1}}-s'_{\vartheta_{2j_m}})$. Thus,

$$s'_{l_1}-s'_{l_2}=s'_{\alpha_1}-s'_{\alpha_r}=(s'_{\alpha_1}-s'_{\alpha_2})+(s'_{\alpha_2}-s'_{\alpha_3})+\cdots+(s'_{\alpha_{r-1}}-s'_{\alpha_r})$$

can be written as a linear combination of the members of A_1. This fact contradicts the assumption that $s'_{l_1} - s'_{l_2}$ 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

$$\begin{aligned} v_1 &= u_{2k_1-1} + s'_{\vartheta_{2k_1-1}} - u_{2k_1} - s'_{\vartheta_{2k_1}}, \\ &\vdots \\ v_{P-1} &= u_{2k_{P-1}-1} + s'_{\vartheta_{2k_{P-1}-1}} - u_{2k_{P-1}} - s'_{\vartheta_{2k_{P-1}}}. \end{aligned}$$

By using the above transformation, and by replacing the P sums over indices $s'_1, \ldots, s'_P$ by $P-1$ sums overs the indices $v_1, \ldots, v_{P-1}$, we find that the joint cumulant given in (3.20) is bounded from above by

$$\frac{1}{n^{L-1}\rho_n^L} \sum_{\boldsymbol{\nu}} M^{L-P+1} \sum_{u_{j_{11}}=-(n-1)}^{n-1} \cdots \sum_{u_{j_{1,k_1-1}}=-(n-1)}^{n-1} \cdots \sum_{u_{j_{P1}}=-(n-1)}^{n-1} \cdots \sum_{u_{j_{P,k_P-1}}=-(n-1)}^{n-1} \sum_{v_1=-3n}^{3n} \cdots$$
$$\sum_{v_{P-1}=-3n}^{3n} |K(b_n v_1)| \times \cdots \times |K(b_n v_{P-1})| \left| Q_{a_{j_{11}} a_{j_{12}} \ldots a_{j_{1k_1}}} \left(\frac{u_{j_{11}}}{\rho_n}, \ldots, \frac{u_{j_{1,k_1-1}}}{\rho_n} \right) \right| \times \cdots$$
$$\times \left| Q_{a_{j_{P1}} a_{j_{P2}} \ldots a_{j_{Pk_P}}} \left(\frac{u_{j_{P1}}}{\rho_n}, \ldots, \frac{u_{j_{P,k_P-1}}}{\rho_n} \right) \right|. \tag{3.21}$$

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 $v_1, \ldots, v_{P-1}$. Now, one can rewrite the expression in (3.21) as follows:

$$\sum_{\boldsymbol{\nu}} M^{L-P+1} \frac{(\rho_n b_n)^{L-P}}{(n b_n)^{L-1}} \left[\sum_{v_1=-3n}^{3n} K(b_n v_1) b_n \right] \times \cdots \times \left[\sum_{v_{P-1}=-3n}^{3n} K(b_n v_{P-1}) b_n \right]$$
$$\times \left\{ \frac{1}{\rho_n^{k_1-1}} \sum_{u_{j_{11}}=-(n-1)}^{n-1} \cdots \sum_{u_{j_{1,k_1-1}}=-(n-1)}^{n-1} \left| Q_{a_{j_{11}} a_{j_{12}} \ldots a_{j_{1k_1}}} \left(\frac{u_{j_{11}}}{\rho_n}, \ldots, \frac{u_{j_{1,k_1-1}}}{\rho_n} \right) \right| \right\} \times \cdots$$
$$\times \left\{ \frac{1}{\rho_n^{k_P-1}} \sum_{u_{j_{P1}}=-(n-1)}^{n-1} \cdots \sum_{u_{j_{P,k_P-1}}=-(n-1)}^{n-1} \left| Q_{a_{j_{P1}} a_{j_{P2}} \ldots a_{j_{Pk_P}}} \left(\frac{u_{j_{P1}}}{\rho_n}, \ldots, \frac{u_{j_{P,k_P-1}}}{\rho_n} \right) \right| \right\}. \tag{3.22}$$

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

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

Note that $\int_{-\infty}^{\infty} S_n(x)dx = \sum_{v_1=-3n}^{3n} K(b_n v_1)b_n$, and from Assumption 3.2 we have the dominance $S_n(x) \leq K_1(x)$. By applying the DCT, we have

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

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

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

Note that

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

From Assumption 3.5A, we have that the function $T_n(\cdot)$ is bounded by an integrable function. Thus, by applying the DCT, we have

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

Likewise, we have the convergence for the remaining $P-1$ sets of summations. By using these above convergence results, the upper bound of (3.20) given in (3.22) 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 3.3. Note that the first moment of the random vector on the left hand side of (3.3) is zero and the second moment converges in accordance with Theorem 3.2. Further,

$$cum(A_1(Y_1-B_1), A_2(Y_2-B_2), \ldots, A_J(Y_J-B_J)) = A_1A_2\cdots A_J \times cum(Y_1, Y_2, \ldots, Y_J),$$

for any set of constants $A_1, \ldots, A_J, B_1, \ldots, B_J$. From the above fact and Lemma 3.3, for all $k > 2$, the absolute value of the kth order joint cumulant of the random vector on the left hand side of (3.3) is bounded from above by an $O((nb_n)^{k/2-k+1})$ term. According to Assumption 3.3, this upper bound tends to 0 as n tends to infinity. This completes the proof. □

Proof of Theorem 3.4. The result can be proved along the lines of the proof of Theorem 2.3 given in Section 2.4. □

Proof of Theorem 3.5. The weak convergence of the first term on the right hand side of (3.4) follows from Theorem 3.3. On the other hand, the second term can be written, in view of Theorem 3.4, as

$$\sqrt{nb_n}\left(E[\widehat{\phi}_{a_1a_2}(f_1)] - \phi_{a_1a_2}(f_1)\right) = \sqrt{nb_n}\left(O\left((\rho_n b_n)^q\right) + O\left(\frac{\rho_n}{n}\right) + O\left(\frac{1}{\rho_n^p}\right)\right). \tag{3.23}$$

Under Assumption 3.3,

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

Therefore, under Assumptions 3.3 and 3.4A, the right hand side of (3.23) goes to zero as $n \to \infty$. This completes the proof. □

Proof of Theorem 3.6. Note that under Assumption 3.4A, we have

$$\lim_{n\to\infty} \sqrt{nb_n}\frac{1}{\rho_n^p} = 0 \tag{3.24}$$

and

$$\begin{aligned}\lim_{n\to\infty} \sqrt{nb_n}\rho_n^q b_n^q = 0 \;\Leftrightarrow\; & \lim_{n\to\infty} (nb_n)^{\frac{1}{2q}} b_n\rho_n = 0 \\ \Leftrightarrow \lim_{n\to\infty} (nb_n)^{1+\frac{1}{2q}} \frac{\rho_n}{n} = 0 \;\Leftrightarrow\; & \lim_{n\to\infty} \sqrt{nb_n} \left(\frac{\rho_n}{n}\right)^{\frac{q}{1+2q}} = 0.\end{aligned} \tag{3.25}$$

From (3.24) and (3.25), we have

$$\frac{1}{\sqrt{nb_n}} = o\left(\left(\frac{\rho_n}{n}\right)^{\frac{q}{1+2q}}\right), \tag{3.26}$$

$$\text{and } \frac{1}{\sqrt{nb_n}} = o\left(\left(\frac{1}{\rho_n}\right)^{p}\right). \tag{3.27}$$

The right hand sides of (3.26) and (3.27) are increasing and decreasing functions, respectively, of ρ_n. Assumption 3.3, together with (3.24), indicate that ρ_n goes to infinity as n goes to infinity. The rate given by (3.26) will be unduly slow if ρ_n goes to infinity too slowly, while the rate given by (3.27) 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, (3.26) implies that

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

This completes the proof. □

3.5 Simulation study

With a view to investigating the applicability of the asymptotic results given in Section 3.2 and 3.3 to finite sample size, we consider the bivariate continuous time linear process

$$\begin{pmatrix} X_1(t) \\ \\ X_2(t) \end{pmatrix} = \begin{pmatrix} \int_{-\infty}^{t} h_1(t-u)Y_1(u)du + \int_{-\infty}^{t} h_2(t-u)Y_2(u)du \\ \\ \int_{-\infty}^{t} h_3(t-u)Y_1(u)du + \int_{-\infty}^{t} h_4(t-u)Y_3(u)du \end{pmatrix}, \tag{3.28}$$

where $Y_j(u)$, $j = 1, 2, 3$ are independent continuous time white noise with variance σ^2 and $h_j(u) = \beta_j e^{-\alpha_j u}$ for $j \in \{1, 2, 3, 4\}$. The elements of the spectral density matrix

$$\begin{pmatrix} \phi_{11}(f) & \phi_{12}(f) \\ \phi_{12}^*(f) & \phi_{22}(f) \end{pmatrix}$$

are defined as follows (Hoel et al., 1972):

$$\begin{aligned} \phi_{11}(f) &= \frac{1}{\sigma^2} \cdot \frac{\beta_1^2}{\alpha_1^2 + (2\pi f)^2} + \frac{1}{\sigma^2} \cdot \frac{\beta_2^2}{\alpha_2^2 + (2\pi f)^2}, \\ \phi_{22}(f) &= \frac{1}{\sigma^2} \cdot \frac{\beta_3^2}{\alpha_3^2 + (2\pi f)^2} + \frac{1}{\sigma^2} \cdot \frac{\beta_4^2}{\alpha_4^2 + (2\pi f)^2}, \\ Re(\phi_{12}(f)) &= \frac{1}{\sigma^2} \cdot \frac{\beta_1\beta_3(\alpha_1\alpha_3 + (2\pi f)^2)}{(\alpha_1^2 + (2\pi f)^2)(\alpha_3^2 + (2\pi f)^2)}, \\ Im(\phi_{12}(f)) &= \frac{1}{\sigma^2} \cdot \frac{\beta_1\beta_3(\alpha_3 - \alpha_1)2\pi f}{(\alpha_1^2 + (2\pi f)^2)(\alpha_3^2 + (2\pi f)^2)}. \end{aligned}$$

The procedure for drawing samples of this bivariate process is given in Section A.1 of the Appendix. We simulate this bivariate process with the choices $\beta_1 = 1$, $\beta_2 = 1$, $\beta_3 = 2$, $\beta_4 = \frac{2}{5}$, $\alpha_1 = \beta_1 \cdot \sqrt{\frac{3}{2}}$, $\alpha_2 = \beta_2 \cdot \sqrt{3}$, $\alpha_3 = \beta_3 \cdot \sqrt{3}$, $\alpha_4 = \beta_4 \cdot \sqrt{3}$ and $\sigma = 0.5$. Note that for this process, Assumption 3.1A holds with $q \geq 1$ and Assumption 3.1B holds with $p \leq 2$. For the purpose of estimation, we make these assumptions with $p = 2$ and $q = 2$. In accordance with this choice of q, we use the second order kernel function

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

We also use the rates $\rho_n = 5 \cdot n^{\frac{1}{6}}$ and $b_n = \frac{1}{4} n^{-\frac{1}{3}-\delta}$ where $\delta = \frac{1}{6}$.

We estimate the bivariate spectrum matrix for frequencies in the range $[0, 2]$ at intervals of .01 (i.e., 201 uniformly spaced grid points). We subsequently compute the normalized statistics

$$T_1(f) = \sqrt{nb_n}\left(\frac{\widehat{\phi}_{11}(f)-\phi_{11}(f)}{\sqrt{2\{1+1_{\{0\}}(f)\}B\widehat{\phi}^2_{11}(f)}}\right),$$

$$T_2(f) = \sqrt{nb_n}\left(\frac{\widehat{\phi}_{22}(f)-\phi_{22}(f)}{\sqrt{2\{1+1_{\{0\}}(f)\}B\widehat{\phi}^2_{22}(f)}}\right),$$

$$T_3(f) = \sqrt{nb_n}\left(\frac{Re(\widehat{\phi}_{12}(f))-Re(\phi_{12}(f))}{\sqrt{\{1+1_{\{0\}}(f)\}B[\widehat{\phi}_{11}(f)\widehat{\phi}_{22}(f)+\{Re(\widehat{\phi}_{12}(f))\}^2-\{Im(\widehat{\phi}_{12}(f))\}^2]}}\right),$$

$$T_4(f) = \sqrt{nb_n}\left(\frac{Im(\widehat{\phi}_{12}(f))-Im(\phi_{12}(f))}{\sqrt{B[\widehat{\phi}_{11}(f)\widehat{\phi}_{22}(f)-\{Re(\widehat{\phi}_{12}(f))\}^2+\{Im(\widehat{\phi}_{12}(f))\}^2]}}\right)\cdot\left[1-1_{\{0\}}(f)\right],$$

in accordance with Theorem 3.2. (According to Theorem 3.5, the asymptotic distribution of each of these four statistics is standard normal.) This procedure is repeated for 500 simulation runs. By regrading the values of the above statistics for the different simulation runs as four data sets of size 500 each, we calculate the Kolmogorov-Smirnov test statistic (Shorak and Wellner, 1986) for these data sets, and the corresponding p-value. This procedure is repeated for the 301 frequency values mentioned above. The percentage of p-values (across 201 frequency values) exceeding the number 0.05 are reported in Table 1, for sample sizes n = 100, 1000, 10000 and 100000. The table shows that for each statistic, the percentage approaches the ideal value of 95 very slowly as n increases.

sample	observed percentage			
size (n)	ϕ_{11}	ϕ_{22}	$Re(\phi_{12})$	$Im(\phi_{12})$
100	0.0 %	0.0 %	0.0 %	0.5 %
1000	10.9 %	10.9 %	16.4 %	27.9 %
10000	80.1 %	71.1 %	80.1 %	87.1 %
100000	83.1 %	87.7 %	92.0 %	89.1 %

Table 3.1. Observed percentage of frequencies (in the range 0 to 2) for which p-values of the Kolmogorov-Smirnov statistics for testing normality of ϕ_{11}, ϕ_{22}, $Re(\phi_{12})$ and $Im(\phi_{12})$ are greater than 0.05 (ideal percentage is 95%).

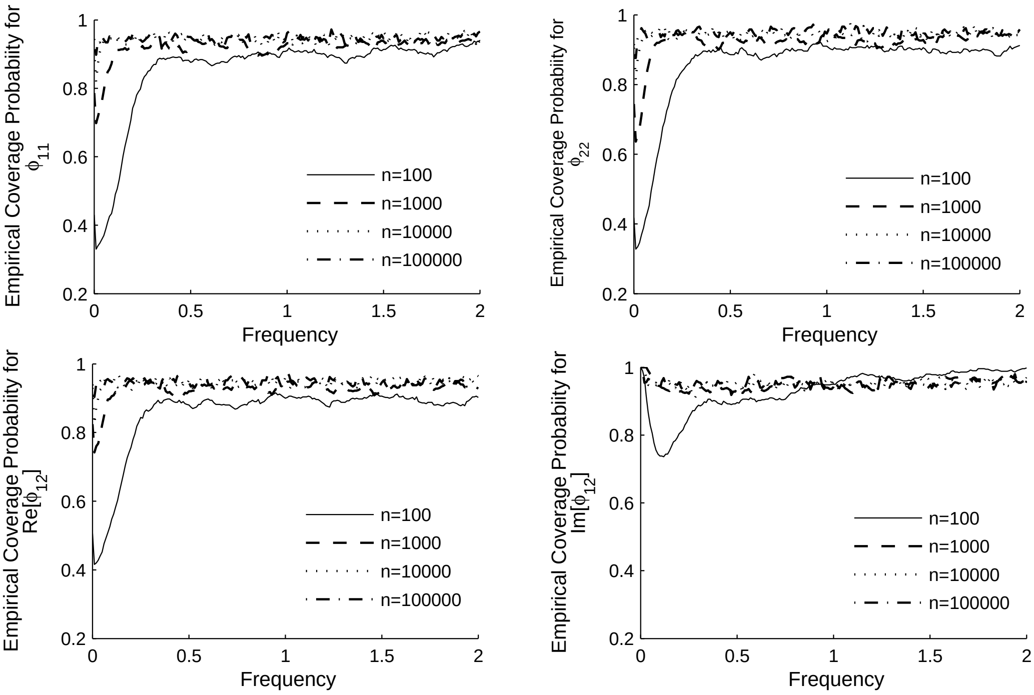

Figure 3.1: Empirical coverage probability (based on 500 simulation runs) of pointwise confidence intervals of ϕ_{11}, ϕ_{22}, $Re(\phi_{12})$ and $Im(\phi_{12})$ for sample sizes 100, 1000, 10000 and 100000.

We now turn to computation of confidence limits of the power spectral density. For each frequency value, we compute the 95% asymptotic confidence intervals of ϕ_{11}, ϕ_{22}, $Re(\phi_{12})$ and $Im(\phi_{12})$ from the statistics $T_1(f)$, $T_2(f)$, $T_3(f)$ and $T_4(f)$, assuming that the latter have the standard normal distribution (see (3.5), (3.9) and (3.10) for the explicit forms of the confidence intervals). Subsequently, we compute the fraction of times (out of 500 simulation runs) the confidence intervals contain the true value of the function. These percentages are plotted against the frequency, for sample sizes $n = 100$, 1000, 10000 and 100000, in Figure 3.1. It is seen that the observed fraction approaches the ideal coverage probability (0.95) for larger sample sizes. Since there is a discontinuity of the asymptotic variance function at the point $f = 0$, while the estimated spectrum is constrained to be continuous, some anomalous behaviour in the neighbourhood of the the point $f = 0$ is expected. This results in substantially lower values of the empirical coverage probability in this region. However, this region of anomaly is observed to shrink as the sample size increases. It is interesting to note that the empirical coverage probability is reasonably close to the ideal coverage probability for most frequency values when the sample size as small as 1000, even though Table 3.1 indicates that the asymptotic distribution is not applicable at this sample size.

3.6 Summary and discussion

In this chapter, we have constructed asymptotic confidence intervals of spectral and cross spectral densities on the basis of uniformly sampled data, using arguments that allow the sampling rate to go to infinity at a suitable rate as the sample size goes to infinity. The simulation results of section 3.5 indicates that these confidence intervals have adequate coverage probability at moderate sample size.

The confidence intervals presented here are the first ones based on uniform sampling, that are applicable to non-bandlimited processes. 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.

Chapter 4

Stochastically sampled processes: Identifiability under inter-sample spacing constraint

4.1 Introduction

In Chapters 2 and 3, we used a special asymptotic argument in connection with point and interval estimation of a possibly non-bandlimited power spectral density, where we allowed the uniform sampling rate to go to infinity at a suitable rate as the sample size goes to infinity. However these asymptotic arguments do not hold when there is a practical constraint on the spacing in between successive samples. This constraint may occur due to technological or economic limitations.

In the presence of a constraint on minimum separation between successive samples, the smoothed periodogram cannot be a consistent estimator of the underlying power spectral density, if the latter is possibly non-bandlimited. However, this fact is not due to a defect of the estimator, this is rather due to a limitation of the data. Let d be the minimum allowable separation between successive samples. Then, the fastest feasible uniform sampling rate is $1/d$. The Nyquist theorem implies that the spectral density of a continuous time, mean square continuous, stationary stochastic process can be identified uniquely from the uniformly sampled process when bandwidth of the underlying process is known to be less than $1/2d$. It can be seen from (1.7) that one can construct a class of continuous time processes, which would lead to uniformly sampled versions having a common wrapped-around power spectral density (Shapiro and Silverman, 1960). Thus, the problem of aliasing is essentially that of non-identifiability of the underlying process from the sampled process.

Given the limitation of uniform sampling in the present situation, one would wonder whether it could be overcome through stochastic sampling. The issue of aliasing/identifiabiity has also been explored in that context, and some formalizations have been made (Shapiro and Silverman, 1960; Beutler, 1970; Masry, 1978a). Masry (1978b) proved that under certain conditions, the estimator (1.14) based on Poisson sampling is consistent for the underlying power spectral density, regardless of whether it is bandlimited, for any average sampling rate. This property implies that one can estimate bandlimited spectra consistently, even if the sampling is done at a sub-Nyquist average rate. Such estimators show that deficiencies in sampling rate can be made up by large sample size, provided one is prepared to sample at irregular intervals. In view of this fact, one might expect that even when there is a constraint on the minimum separation between successive samples, one can judiciously use non-uniform sampling to consistently estimate spectra with much larger bandwidth than what can be achieved through uniform sampling.

It is important to note that a small average sampling rate does not mean that *any* two successive samples are far apart. In the case of Poisson sampling with any average sampling rate, it can be seen that as the sample size goes to infinity, there would be a large number of pairs of consecutive samples which are nearer to each other than any specified threshold. Thus, in order to use Poisson sampling with any average sampling rate, sometimes one has to sample the process very fast. Many other non-uniform and alias-free sampling schemes also have this requirement. All these schemes become infeasible if there is a hard limit on the minimum separation between successive samples.

Books on sampling (Higgins, 1996; Benedetto and Ferreira, 2001) give a clear picture of the limitations of uniform sampling in respect of a constraint on the minimum separation between successive samples. However, suitability of non-uniform sampling schemes in the presence of this constraint has not been studied so far.

In this chapter, we consider the problem of identification of the power spectral density of a stationary stochastic process through non-uniform sampling, under a constraint on the minimum separation between successive samples. In Section 4.2, we describe the underlying set-up and discuss the notions of alias-free sampling provided by Shapiro and Silverman (1960) as well as by Masry (1978a). In Section 4.3, we consider the class of all power spectra, and show that under the above constraint, no stationary point process sampling scheme is alias-free for this class. Subsequently, we study the possibility of alias-free sampling for estimation of spectra that are known to be confined to a certain bandwidth. The proofs of all the results of this section is given in Section 4.4. In Section 4.5, we report the results of a simulation study of the performance of the estimator (1.14) based on Poisson sampling, in the presence of the above constraint. We summarize the findings and provide some concluding remarks in Section 4.6.

4.2 Existing notions of alias-free stochastic sampling

As in Chapter 1, let $X = \{X(t),\ -\infty < t < \infty\}$ be a real, mean square continuous and wide sense stationary stochastic process with mean zero, covariance function $C(\cdot)$ and spectral distribution function $\Phi(\cdot)$. If $\Phi(\cdot)$ has a density, we denote it by $\phi(\cdot)$. Let $\tau = \{t_j,\ j = \ldots, -2, -1, 0, 1, 2, \ldots\}$ be a sequence of real-valued, stochastic sampling times, i.e., a point process on the real line.

Shapiro and Silverman (1960) introduced a notion of alias-free sampling. Beutler (1970) further formalized this definition of alias-free sampling for different classes of power spectra. This notion of alias-free sampling is based on the following assumptions about the sampling process.

Assumption 4.1. *The process τ is independent of X.*

Assumption 4.2. *The sequence of sampling times τ is such that the probability distribution of $(t_{l+m} - t_l)$ does not depend on l.*

Under a sampling scheme that satisfies Assumptions 4.1 and 4.2, the sampled process $X_\tau = \{X(t),\ t \in \tau\}$ is wide sense stationary. Denote the covariance sequence of the sampled process X_τ by $c = \{\ldots, c_{-2}, c_{-1}, c_0, c_1, c_2, \ldots\}$, where

$$c_m = E[X(t_{l+m})X(t_l)],\ \text{for } l,\, m \text{ integers},$$

and the expectation is taken without conditioning on the sampling times. Beutler's definition of alias-free sampling, based on Shapiro and Silverman's earlier idea, is as follows.

Definition 4.1. *The sampling process τ satisfying Assumptions 4.1 and 4.2 is alias-free relative to the class of spectra $\mathbb{S}$ if no two random processes with different spectra belonging to $\mathbb{S}$ yield the same covariance sequence (c) of the sampled process.*

Shapiro and Silverman (1960) had considered the special case where the sampling times constitute a renewal process, and $\mathbb{S}$ is the class of *all* spectra with integrable and square integrable densities. They referred to this scheme as additive random sampling, and showed that it is alias-free, provided the characteristic function of the inter-arrival distribution takes no value more than once on the real line. In particular, Poisson sampling is alias-free for the class of spectra $\mathbb{S}$.

The above definition has the drawback that it does not make use of the information contained in the sampling times. If one wishes to reconstruct $\phi(\cdot)$ using a sampling scheme that is alias-free according to the above definition, then that would be done on the basis of the sequence c only. Beutler (1970) gave a procedure for this reconstruction, and indicated that this procedure may be used to estimate $\phi(\cdot)$ from estimates of c. However, as Masry (1978a) pointed out, consistency of any estimator obtained from this procedure has not been studied.

From all these considerations, this approach appears to be rather restrictive. In practice, one would expect to use the information contained not only in the sampled values, but also in the sampling times, in order to estimate the power spectral density. In order to take into account the sampling times, Masry (1978a) gave an alternative definition of alias-free sampling. This definition requires a property of point processes, which we describe below.

Let $\mathbb{B}$ be the Borel σ-field on the real line. Any point process $\tau = \{t_j,\ j = \ldots, -2, -1, 0, 1, 2, \ldots\}$ on the real line can be represented in terms of a counting process $\{N(B), B \in \mathbb{B}\}$ defined as

$$N(B) = \sum_{j=-\infty}^{\infty} 1_B(t_j). \tag{4.1}$$

A point process τ is said to be orderly if $P\{N(0,x] > 1\} = o(x)$ as $x \downarrow 0$. The point process τ is said to be second order stationary if its second order moments exist and

$$E[N(A+t)] = E[N(A)] \text{ for all } A \in \mathbb{B} \text{ and } -\infty < t < \infty, \tag{4.2}$$

$$Cov[N(A+t), N(B+t)] = Cov[N(A), N(B)] \text{ for all } A, B \in \mathbb{B} \text{ and } -\infty < t < \infty, \tag{4.3}$$

where $A + t$ denotes the set $\{x : x - t \in A\}$. By using (4.2) and additivity property of (4.1), it can be easily shown that

$$E[N(A)] = E[N(0,1]] \times \int_{-\infty}^{\infty} 1_A(t)dt.$$

The number $E[N(0,1]]$ is called the mean intensity of the point process. By using (4.3), it can be shown that

$$Cov[N(A), N(B)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} 1_{A\times B}(t, t+u)\mu_\tau(du)dt,$$

where μ_τ is a σ-finite signed measure also known as the reduced covariance measure of the point process τ (for details see Daley and Vere-Jones (2002)).

Assumption 4.3. *The process τ constitutes a second-order stationary orderly point process on the real line.*

Let β be the mean intensity and μ_τ be the reduced covariance measure of the process τ. Consider the compound process $\{Z(B),\ B \in \mathbb{B}\}$ defined by

$$Z(B) = \sum_{t_j \in B} X(t_j).$$

The process $Z = \{Z(B),\ B \in \mathbb{B}\}$ is second order stationary (i.e., the first and second moments of $Z(B + t)$, for any real number t, does not depend of t). Let μ_z be the covariance measure of the process Z. It can be shown that this measure is given by

$$\mu_z(B) = \int_B C(u)[\beta^2 du + \mu_\tau(du)]. \tag{4.4}$$

Masry's notion of alias-free sampling is as follows.

Definition 4.2. *The sampling process τ satisfying Assumptions 4.1 and 4.3 is alias-free relative to the class of spectra $\mathbb{S}$ if no two random processes with different spectra belonging to $\mathbb{S}$ yield the same covariance measure (μ_z) of the compound process.*

Note that this definition makes use of the information contained in the sampling times, as the covariance measure μ_z involves the mean intensity β as well as the reduced covariance measure μ_τ of the sampling process. It has been shown that, according to Definition 4.2, Poisson sampling is alias-free for the class of all spectra having integrable and square integrable densities (Masry, 1978a).

4.3 Sampling under constraint of minimum inter-sample spacing

As mentioned in Section 4.1, the focus of the present work is on a sampling process τ which satisfies the following constraint.

Assumption 4.4. *The time separation between two successive sample points is at least d (i.e., $t_{m+1} - t_m \geq d$ for any index m) for some fixed $d > 0$.*

In this section, we investigate whether a sampling scheme satisfying this constraint can be alias-free.

4.3.1 General spectra

We present some negative results in the case when $\mathbb{S}$ is the class of *all* spectra – band-limited or otherwise.

Theorem 4.1. *No sampling point process satisfying Assumptions 4.1, 4.2 and 4.4 is alias-free according to Definition 4.1, for the class of all spectra.*

Theorem 4.2. *No sampling point process satisfying Assumptions 4.1, 4.3 and 4.4 is alias-free according to Definition 4.2, for the class of all spectra.*

We prove these theorems in the Section 4.4 by constructing counter-examples, based on the following class of power spectral densities:

$$\mathbb{A} = \left\{ \phi(\cdot) : \int_{-\infty}^{\infty} \phi(f) e^{i2\pi f t} df = 0 \text{ for } |t| > d \right\}. \tag{4.5}$$

The members of this class correspond to covariance functions supported over the interval $[-d, d]$. For example, a member of this class is the power spectral density defined by

$$\phi_a(f) = \frac{2}{a} \cdot \frac{1 - \cos(2\pi f a)}{(2\pi f)^2}, \quad -\infty < f < \infty,$$

for any arbitrary positive $a \in (0, d]$. This density corresponds to the covariance function

$$C_a(t) = \begin{cases} 1 - \frac{|t|}{a} & \text{for } |t| \leq a, \\ 0 & \text{for } |t| > a. \end{cases}$$

Some other members of $\mathbb{A}$ can be constructed by convolving $\phi_a(\cdot)$ with an arbitrary power spectral density. We show in the Section 4.4 that if X_1, X_2 and X_3 are independent mean square continuous stochastic processes such that X_2 and X_3 have different spectra belonging to $\mathbb{A}$ and have the same variance, then the spectra of $X_1 + X_2$ and $X_1 + X_3$ cannot be distinguished from the sequence c or the measure μ_z, leading to aliasing according to Definitions 4.1 and 4.2.

One can easily construct two *integrable and square integrable* power spectral densities that are indistinguishable from c or μ_z. Therefore, the statements of Theorems 4.1 and 4.2 also hold in respect of all spectra having integrable and square integrable densities (rather than all spectra). Thus, the alias-free property of Poisson sampling mentioned in Section 4.2 become inapplicable, once the inter-sample spacings are adjusted in accordance with Assumption 4.4.

These two theorems show that, under the constraint of a minimum inter-sample spacing, any point process sampling scheme would be inadequate for the identification of a completely unrestricted power spectral density – according to the existing notions of alias-free sampling. If the power spectral density of the original continuous time process is not identifiable from the sequence c or the covariance measure μ_z, then one cannot expect to consistently estimate the power spectral density on the basis of estimates of either of these.

It is well known that estimators based on uniformly spaced samples, irrespective of the sampling rate, also suffer from the limitation of non-identifiability. In fact, it is this limitation of uniform sampling that has been historically used as one of the major arguments in favor of non-uniform sampling schemes. The above theorems show that the same difficulty applies to practical non-uniform sampling schemes as well.

4.3.2 Bandlimited spectra

In the case of uniform sampling, it is well known that a bandlimited process would not lead to aliasing provided that the sampling is done at the Nyquist rate or faster. On the other hand, uniform sampling at any fixed rate would be free from the problem of aliasing if the spectrum of the continuous time process is known to be confined to an appropriate band. This fact, together with the limitation of point process sampling in the case of non-bandlimited spectra, gives rise to the question: Can point process sampling under Assumption 4.4 be alias-free for the class of bandlimited spectra? If so, it would be interesting to compare the maximum allowable spectral bandwidths for alias-free sampling, arising from uniform and point process sampling schemes under Assumption 4.4.

It turns out that alias-free sampling under Assumption 4.4 is possible for an important class of stochastic sampling schemes, namely, renewal process sampling. This is a special case of point process sampling, which has received much attention from researchers (Shapiro and Silverman, 1960; Beutler, 1970; Masry, 1978a; Tarczynski and Allay, 2004). Poisson sampling is a further special case of renewal process sampling. However, it is an ideal sampling scheme (see discussion in page 80), in contrast with *implementable* renewal process sampling schemes that would satisfy Assumption 4.4.

Note that the fastest possible rate of uniform sampling under Assumption 4.4 is $1/d$. Uniform sampling at this rate is alias-free for the class of spectra supported on $[-1/2d, 1/2d]$. This frequency interval would be the benchmark for the present study.

First, we present a general result that would be useful in answering the foregoing question, as far as Definition 4.1 of alias-free sampling is concerned.

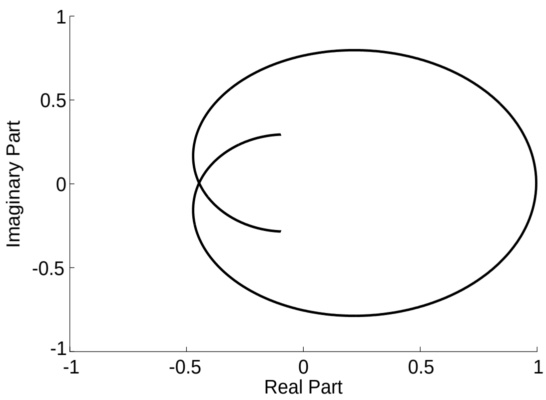

Figure 4.1: Graph of $\eta(f)$ on the complex plane for the left truncated exponential distribution, with mean $2d$ and truncation point d, for $-1/2d < f \leq 1/2d$.

Theorem 4.3. *A renewal process sampling scheme satisfying Assumptions 4.1 and 4.2, and having characteristic function of the inter-sample spacing denoted by η, is alias-free relative to a class of spectra supported on the closed and finite interval I according to Definition 4.1 if and only if the graph of $\eta(f)$ on the complex plane, for $f \in I$, does not divide the complex plane.*[1]

Theorem 4.3 relates the alias-free property of a renewal process sampling scheme to the geometry of the characteristic function of the inter-sample spacing. It may be noted that the distribution of $d+X$, where d is fixed and X has the gamma distribution with any combination of parameters, does not satisfy the necessary and sufficient condition given in Theorem 4.3 for $I = [-1/2d, 1/2d]$. It follows that the corresponding renewal process sampling schemes, including the case of inter-sample spacing having a left-truncated exponential distribution, are not alias-free according to Definition 4.1, relative to a class of spectra limited to the band $[-1/2d, 1/2d]$. The graph of $\eta(f)$ for the left-truncated exponential distribution with mean $2d$ and truncation point d, for $-1/2d \leq f \leq 1/2d$, is shown in Figure 4.1. For such sampling schemes, aliasing can be avoided only if the continuous time process is confined to a bandwidth that is even smaller than $1/2d$, the maximum allowable bandwidth in the case of uniform sampling.

However, there are some other renewal process sampling schemes that satisfy As-

[1] A graph does not divide the complex plane if any two points on the complex plane can be connected by a continuous path, which does not have a point of intersection with the graph (Lavrentieff, 1936).

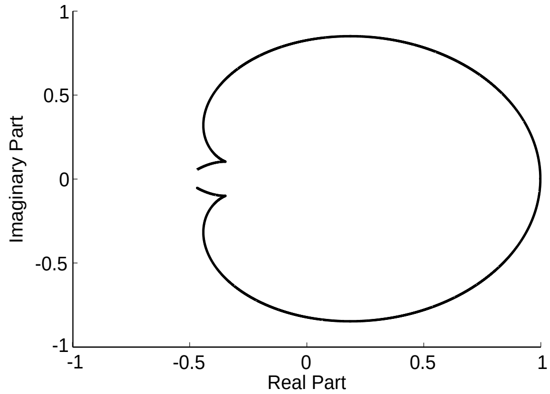

Figure 4.2: Graph of $\eta(f)$ on the complex plane for the example given in the proof of Theorem 4.4, for $-1.1/2d < f \leq 1.1/2d$.

sumption 4.4 and are alias-free for the class of spectra limited to a band larger than $[-1/2d, 1/2d]$, as the next theorem shows.

Theorem 4.4. *There exists a closed and finite interval I, which contains the interval $[-1/2d, 1/2d]$, and a renewal process sampling scheme satisfying Assumptions 4.1, 4.3 and 4.4 which is alias-free relative to the class of spectra supported on I, according to Definition 4.1.*

The proof of Theorem 4.4, given in the Section 4.4, invokes an example, for which I is more than 10% larger than the interval $[-1/2d, 1/2d]$, while the average inter-sample spacing is about 35% more than the minimum allowable spacing (d). The graph of $\eta(f)$ for this inter-sample spacing distribution, for $-1.1/2d \leq f \leq 1.1/2d$, is shown in Figure 4.2.

We now turn to Definition 4.2. Since this notion of alias-free sampling is weaker than that of Definition 4.1, one can expect a stronger result.

Theorem 4.5. *Any renewal process sampling scheme, satisfying Assumptions 4.1, 4.3 and 4.4 and the further assumption that the inter-sample spacing distribution has a density that is positive over an interval, is alias-free according to Definition 4.2, for the class of spectra limited to the band $[-f_0, f_0]$ for every finite $f_0 > 0$.*

Theorem 4.5 shows that, under the constraint of a minimum allowable separation between successive samples, renewal process sampling is alias-free (according to Definition 4.2) for a *wider* range of power spectra than uniform sampling. It is interesting to note that sampling schemes following the assumptions of Theorem 4.5 are alias-free according to Definition 4.2 when the spectral density of the underlying continuous-time process is known to be confined to *any* finite bandwidth (no matter how large), but according to Theorem 4.2, these are not alias-free when the process is non-bandlimited.

It transpires from the foregoing discussion that there are contrasting scopes of alias-free renewal process sampling under the constraint of a minimum allowable inter-sample spacing, according to Definitions 4.1 and 4.2. The limited scope of alias-free sampling in the case of Definition 4.1 stems from the fact that, under that notion, one aims to identify spectra solely from the sequence c, which is rather restrictive.

4.4 Proofs

Proof of Theorem 4.1. Consider independent, zero mean, mean square continuous stationary stochastic processes X_1, X_2 and X_3, having covariance functions $C_j(\cdot)$, $j = 1, 2, 3$, respectively, such that $C_2(0) = C_3(0)$ and X_2 and X_3 have different spectral densities belonging to the class $\mathbb{A}$ defined in (4.5). Consider a sampling point process $\tau = \{t_j,\ j = \dots, -2, -1, 0, 1, 2, \dots\}$ satisfying the Assumptions 4.1, 4.2 and 4.4. Let the processes $X_1{+}X_2$ and $X_1{+}X_3$ have spectral distributions $\Phi_{12}(\cdot)$ and $\Phi_{13}(\cdot)$, respectively, and covariance sequences of sampled processes $c_{12} = \{c_{12,m},\ m = \dots, -2, -1, 0, 1, 2, \dots\}$ and $c_{13} = \{c_{13,m},\ m = \dots, -2, -1, 0, 1, 2, \dots\}$, respectively. We have

$$c_{12,0} = C_1(0) + C_2(0) = C_1(0) + C_3(0) = c_{13,0}.$$

For arbitrary integers l and m, let $F_m(x)$ be the distribution function of $(t_{l+m} - t_l)$. Assumption 4.4 implies that $F_m(x)$ is supported on the interval $[|m|d, \infty)$. It follows that, for $m \neq 0$,

$$\begin{aligned} c_{12,m} =& E[X_1(t_{l+m})X_1(t_l)] + E[X_2(t_{l+m})X_2(t_l)] \\ =& \int_0^\infty C_1(u)dF_m(u) + \int_0^\infty C_2(u)dF_m(u) \end{aligned}$$

$$= \int_{|m|d}^{\infty} C_1(u)dF_m(u) + \int_{|m|d}^{\infty} C_2(u)dF_m(u)$$
$$= \int_{|m|d}^{\infty} C_1(u)dF_m(u),$$

since $C_2(\cdot)$ is supported on $[-d,d]$. Likewise, $c_{13,m}$ is also equal to the last expression. This completes the proof. □

Proof of Theorem 4.2. Consider independent, zero mean, mean square continuous stationary stochastic processes X_1, X_2 and X_3, having covariance functions $C_j(\cdot)$, $j = 1,2,3$, respectively, such that $C_2(0) = C_3(0)$ and X_2 and X_3 have different spectral densities belonging to the class $\mathbb{A}$ defined in (4.5). Consider a sampling point process $\tau = \{t_j,\ j = \ldots,-2,-1,0,1,2,\ldots\}$ satisfying the Assumptions 4.1, 4.3 and 4.4, and having mean intensity β and reduced covariance measure μ_τ. Let the processes X_1+X_2 and X_1+X_3 have spectral distributions $\Phi_{12}(\cdot)$ and $\Phi_{13}(\cdot)$, respectively. As in Section 4.2, consider the compound processes

$$Z_{1j} = \left\{ Z_{1j}(B) = \sum_{t_k \in B} X_1(t_k) + X_j(t_k),\ B \in \mathbb{B} \right\},\ j = 2,3,$$

which have covariance measures $\mu_{z_{12}}$ and $\mu_{z_{13}}$ given by

$$\mu_{z_{1j}}(B) = \int_B \{C_1(u) + C_j(u)\}\,[\beta^2 du + \mu_\tau(du)],\ j = 2,3,$$

respectively.

The reduced covariance measure μ_τ of the point process τ can be expressed as

$$\mu_\tau(B) = \beta\delta_0(B) + \beta \int_B [dS(|u|) - \beta du],\ B \in \mathbb{B},$$

where, for $u > 0$,

$$S(u) = \sum_{m=1}^{\infty} F_m(u),$$

$F_m(u)$ is the conditional probability

$$F_m(u) = \lim_{\epsilon \downarrow 0} P\left[N(t,t+u] \geq m \;\middle|\; N(t-\epsilon,t] \geq 1\right],$$

and $\{N(B),\ B \in \mathbb{B}\}$ is the counting process induced by the process τ (Beutler and Leneman, 1966; Daley and Vere-Jones, 2002). Assumption 4.4 implies that $S(u) = 0$ for $u \in [0, d]$.

It follows from the above representation of μ_τ that, for each Borel set B, the covariance measures $\mu_{z_{12}}$ is given by

$$\mu_{z_{12}}(B) = \int_B C_1(u)[\beta^2 du + \mu_\tau(du)] + \int_{B\cap[-d,d]} C_2(u)[\beta^2 du + \mu_\tau(du)]$$

$$= \int_B C_1(u)[\beta^2 du + \mu_\tau(du)] + \beta C_2(0)\delta_0(B \cap [-d, d]).$$

Since $C_2(0) = C_3(0)$, it is clear that the measures $\mu_{z_{12}}$ and $\mu_{z_{13}}$ agree on all Borel sets. This completes the proof. □

Proof of Theorem 4.3. Here, the sampling process $\tau = \{t_j,\ j = \ldots, -2, -1, 0, 1, 2, \ldots\}$ is such that the inter-sample spacing $t_{m+1} - t_m$ for different values of m are independent and identically distributed, say with distribution function $F(\cdot)$.

Let $\mathbb{S}$ be the class of spectra supported on the closed and finite interval I. Let X be a process as defined in the theorem, and have the power spectral distribution $\Phi(\cdot)$ belonging to $\mathbb{S}$. The covariance sequence c of the sampled process is given by

$$c_m = E[X(t_{l+m})X(t_l)] = E\left[E\left[X(t_{l+m})X(t_l) \;\middle|\; \tau\right]\right]$$

$$= E\left[C(t_{l+m} - t_l)\right] = E\left[\int_I e^{i2\pi f(t_{l+m}-t_l)} d\Phi(f)\right]$$

$$= \int_I E\left(e^{i2\pi f(t_{l+m}-t_l)}\right) d\Phi(f).$$

The interchange of the integral and the expectation is possible by Fubini's theorem, since the power spectral distribution $\Phi(\cdot)$ and the probability distribution of $t_{l+m} - t_l$ are both finite. Since the latter distribution is the m-fold convolution of the inter-sample spacing distribution, we have

$$c_m = \int_I [\eta(f)]^m \, d\Phi(f), \tag{4.6}$$

where $\eta(\cdot)$ is the characteristic function of inter-sample spacing distribution, i.e.,

$$\eta(f) = \int_0^\infty e^{i2\pi fy} dF(y), \quad -\infty < f < \infty.$$

The sampling scheme τ is alias-free relative to the class of spectra $\mathbb{S}$ according to Definition 4.1, if no two different spectra Φ_1 and Φ_2 belonging to $\mathbb{S}$ produce the same covariance sequence c. Since the sequence c satisfies $c_{-m} = c_m$, the foregoing condition is equivalent to the statement:

$$\int_I [\eta(f)]^m (d\Phi_1(f) - d\Phi_2(f)) = 0 \text{ for } m = 0, 1, 2, \ldots \quad \text{implies that } \Phi_1(\cdot) = \Phi_2(\cdot). \tag{4.7}$$

The above integral with respect to the real variable f can be written as a complex integral over the contour

$$\Omega = \{z : \; z = \eta(f), \quad f \in I\}. \tag{4.8}$$

Thus, we can conclude that the sampling scheme τ is alias free relative to the class of spectra $\mathbb{S}$ according to Definition 4.1 if and only if

$$\begin{aligned} &\text{For any signed measure } \nu \text{ defined on the Borel } \sigma\text{-field on } \Omega, \\ &\int_\Omega z^m \nu(dz) = 0 \text{ for } m = 0, 1, 2, \ldots \implies \nu = 0. \end{aligned} \tag{4.9}$$

Note that, since Ω is the image of the continuous function $\eta(f)$ on the closed and finite interval I, the contour Ω is compact. Let $C(\Omega)$ be the Banach space of all complex-valued continuous functions on Ω equipped with the supremum norm. Let $\mathbb{M}$ be the set of all signed measures defined on the Borel σ-field on Ω. For any $\nu \in \mathbb{M}$, define the complex valued bounded linear functional L_ν defined on $C(\Omega)$ as

$$L_\nu(g) = \int_\Omega g(z)\nu(dz) \quad \text{for all } g \in C(\Omega). \tag{4.10}$$

In terms of these notations, we rewrite (4.9) as

$$\text{for any } \nu \in \mathbb{M}, \text{ “}L_\nu(z^m) = 0 \text{ for } m = 0, 1, 2, \ldots\text{”} \implies \nu = 0. \tag{4.11}$$

By the Riesz representation theorem, *every* bounded linear functional L on $C(\Omega)$ can be represented as

$$L(g) = \int_\Omega g(z)\nu_1(dz) + i\int_\Omega g(z)\nu_2(dz) \quad \text{for all } g \in C(\Omega), \tag{4.12}$$

for a unique pair of measures ν_1 and ν_2 in $\mathbb{M}$ (Rudin, 1987). It follows that the necessary and sufficient condition (4.11) is equivalent to the condition:

$$\begin{aligned} &\text{For any bounded linear functional } L \text{ on } C(\Omega), \\ &\text{“}L(z^m) = 0 \text{ for } m = 0, 1, 2, \dots\text{”} \implies L = 0. \end{aligned} \tag{4.13}$$

The above condition is a statement about the sequence $\{1, z, z^2, \dots, \}$ in relation to the Banach space $C(\Omega)$ {p. 257 of (Davis, 1975)}. By Theorem 11.1.7 of (Davis, 1975), (4.13) is equivalent to the condition:

$$\text{“The linear span of the sequence } \{1, z, z^2, \dots, \} \text{ is dense in } C(\Omega).\text{”} \tag{4.14}$$

The above condition can be rephrased as: “Any $g \in C(\Omega)$ can be expanded in a uniformly convergent sequence of polynomials.” By a result of (Mergelyan, 1954) (see also (Lavrentieff, 1936)), we get the further equivalent condition:

$$\text{“The set } \Omega \text{ is nowhere dense and does not divide the plane.”} \tag{4.15}$$

Since the set Ω is a curve in the complex plane, it is always a nowhere dense set. This completes the proof. □

Proof of Theorem 4.4. Let $I = [-1.1/2d, 1.1/2d]$. Consider the two-point discrete distribution F, given by

$$F(t) = \begin{cases} 0 & \text{for } t < d, \\ 0.68 & \text{for } d \le t < 2.1d, \\ 1 & \text{for } t \ge 2.1d. \end{cases} \tag{4.16}$$

It follows that the average inter-sample spacing is $1.352d$. Also,

$$\eta(f) = 0.68e^{i2\pi fd} + 0.32e^{i2\pi f\times 2.1d}.$$

The plot of the imaginary part of $\eta(f)$ against the real part, for $f \in I$, is given in Figure 4.2. It can be verified that the graph does not divide the complex plane. The result follows from Theorem 4.3. □

Proof of Theorem 4.5. Here, the sampling process $\tau = \{t_j,\ j = \ldots, -2, -1, 0, 1, 2, \ldots\}$ is such that the inter-sample spacing $t_{m+1} - t_m$ for different values of m are independent and identically distributed having probability density function $\gamma(\cdot)$. Let β and μ_τ be the mean intensity and the reduced covariance measure, respectively, of the process τ. The measure μ_τ can be expressed as

$$\mu_\tau(B) = \beta\delta_0(B) + \int_B \beta[H(|u|) - \beta]du \text{ for each } B \in \mathbb{B}, \tag{4.17}$$

where $H(u)$ is the renewal density function, i.e,

$$H(u) = \sum_{m=1}^{\infty} \gamma^{(m)}(u).$$

Note that Assumption 4.4 implies that $\gamma(\cdot)$ is supported on $[d, \infty)$, and so $H(\cdot)$ is supported on $[d, \infty)$. Let l be and integer (greater than 1), such that $H(u) > 0$ for $u \geq ld$. (The positivity of $H(\cdot)$ over some semi-infinite interval follows from the additional assumption made in the statement of the theorem.)

Let $\mathbb{S}$ be the class of bandlimited spectra supported on $[-f_0, f_0]$. If the sampling scheme τ is not alias-free relative to the class of spectra $\mathbb{S}$, then there exist two zero mean, mean square continuous stationary stochastic processes X_1 and X_2 with different power spectral distributions $\Phi_1(\cdot)$ and $\Phi_2(\cdot)$ such that compound processes

$$Z_j = \left\{ Z_j(B) = \sum_{t_k \in B} X_j(t_k),\ B \in \mathbb{B} \right\},\ j = 1, 2,$$

have the covariance measures μ_{z_1} and μ_{z_2}, respectively, satisfying $\mu_{z_1} = \mu_{z_2}$. Here, for

$B \in \mathbb{B}$, the covariance measures are given by (see (4.4) and (4.17))

$$\mu_{z_j}(B) = \beta C_j(0)\delta_0(B) + \beta \int_B C_j(u)H(|u|)du, \;\; j = 1, 2,$$

where $C_1(\cdot)$ and $C_2(\cdot)$ are the covariance functions of the processes X_1 and X_2 respectively. In order that the covariance measures μ_{z_1} and μ_{z_2} are the same, the point masses at zero, as well as the absolutely continuous parts, must agree. The equality of the point masses requires

$$C_1(0) = C_2(0). \tag{4.18}$$

On the other hand, equality of the absolutely continuous parts means

$$C_1(u)H(|u|) = C_2(u)H(|u|) \text{ for } -\infty < u < \infty.$$

Since $H(u) > 0$ for the $u \geq ld$, we have

$$C_1(u) = C_2(u) \text{ for } |u| \geq ld. \tag{4.19}$$

If the processes X_1 and X_2 have spectra limited to the band $[-f_0, f_0]$, then the covariance function $C_j(\cdot)$ for $j = 1, 2$ can be expressed as (see (1.9))

$$C_j(u) = \sum_{m=-\infty}^{\infty} C_j(mT)\text{sinc}\left(\frac{\pi}{T}(u - mT)\right), \tag{4.20}$$

where $T = \frac{1}{2f_0}$.

Let $k = [ld/T]$, where $[u]$ represents the integer part of the real number u. It follows from (4.18)–(4.20) that

$$\begin{aligned} C_1(u) - C_2(u) = &\sum_{m=-k}^{k} \{C_1(mT) - C_2(mT)\}\,\text{sinc}\left(\frac{\pi}{T}(u - mT)\right) \\ &+ \sum_{|m|>k} \{C_1(mT) - C_2(mT)\}\,\text{sinc}\left(\frac{\pi}{T}(u - mT)\right) \\ = &\sum_{m=1}^{k} \{C_1(mT) - C_2(mT)\}\left(\text{sinc}\left(\frac{\pi}{T}(u - mT)\right) + \text{sinc}\left(\frac{\pi}{T}(u + mT)\right)\right). \end{aligned}$$

By using the fact that $\sin(k\pi + \theta) = (-1)^k \sin\theta$ for all integer k, we have for $\alpha =$

$u/T - [u/T] > 0$,

$$\begin{aligned}
&C_1(u) - C_2(u)\\
&= \sum_{m=1}^{k} \{C_1(mT) - C_2(mT)\} \left(\frac{\sin\left\{\left(-m + \left[\frac{u}{T}\right]\right)\pi + \alpha\pi\right\}}{\frac{\pi}{T}(u - mT)} + \frac{\sin\left\{\left(m + \left[\frac{u}{T}\right]\right)\pi + \alpha\pi\right\}}{\frac{\pi}{T}(u + mT)} \right)\\
&= (-1)^{[u/T]} \sin(\alpha\pi) \sum_{m=1}^{k} \{C_1(mT) - C_2(mT)\} \left(\frac{(-1)^{-m}}{\frac{\pi}{T}(u - mT)} + \frac{(-1)^{m}}{\frac{\pi}{T}(u + mT)} \right).
\end{aligned}$$

Since $(-1)^m = (-1)^{-m}$ for each integer m, we have

$$C_1(u) - C_2(u) = (-1)^{[u/T]} \frac{2uT}{\pi} \sin(\alpha\pi) \sum_{m=1}^{k} [(-1)^m \{C_1(mT) - C_2(mT)\}] \frac{1}{u^2 - m^2T^2}.$$

Let $v_m = [(-1)^m \{C_1(mT) - C_2(mT)\}]$. In view of (4.19), the above equation implies that

$$\sum_{m=1}^{k} \frac{v_m}{u^2 - m^2T^2} = 0 \tag{4.21}$$

for $u \in \{(ld, (k+1)T)\} \cup \left\{\cup_{j=k+1}^{\infty} (mT, (j+1)T)\right\}$.

Note that the function on the left hand side of (4.21) is a ratio of polynomials. The polynomial in the numerator has degree $2k-2$, while the denominator is bounded over the domain of the function. Thus, the ratio of the polynomials can be zero at most at $2k-2$ points. Therefore, the fact that this function assumes the value 0 everywhere on the interval $((k+1)T, (k+2)T)$ implies that the polynomial in the numerator is identically equal to zero. Thus, the ratio of the polynomials is identically zero. Hence,

$$\sum_{m=1}^{k} \frac{v_m}{u^2 - m^2T^2} = 0, \text{ for } u \in \bigcup_{j=0}^{\infty} (mT, (j+1)T).$$

By considering the limit of the left hand side as $u \downarrow mT$, it is found that $v_m = 0$ for $m = 1, \ldots, k$, that is,

$$C_1(mT) = C_2(mT) \text{ for } |m| = 1, \ldots, k.$$

According to (4.18), the above equality holds for $m = 0$, while (4.20) and (4.19) imply that it holds for $|m| = k+1, k+2, \ldots$. Thus, $C_1(mT) = C_2(mT)$ for all m. It follows from (4.20) that $C_1(u) = C_2(u)$ for each u, which contradicts the assumption that C_1 and C_2 are different. So the sampling scheme τ is alias-free for the class of the spectra $\mathbb{S}$. This completes the proof. □

4.5 Simulation study

In view of Theorems 4.1, 4.2 and 4.5, it is clear that non-aliased sampling (and consequently, consistent estimation) under the constraint of minimum separation between successive samples *may* be possible only in the case of bandlimited processes. In this context, performance of well-known estimators such as $\widehat{\phi}_\rho(\cdot)$ and $\widehat{\psi}_n(\cdot)$ given in (1.12) and (1.14), respectively, would be of interest.

The performance of the estimator $\widehat{\phi}_\rho(\cdot)$ under the constraint of minimum separation between successive samples is understood from standard theory. On the other hand, Poisson sampling is not implementable under this constraint. However, 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 amounts to a specific type of renewal process sampling. For this renewal process, the inter-sample spacings have the exponential distribution truncated from the left at d, the mean spacing being $d + \theta$. An interesting question is, how the estimator $\widehat{\psi}_n(\cdot)$ would perform under such imperfect Poisson sampling.

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

$$C(u) = \begin{cases} \sigma^2 \sum_{j=0}^{q-|l|} \xi_j \xi_{j+|l|} & \text{if } u = \frac{l}{2f_0}, |l| = 0, 1, \ldots, q, \\ 0 & \text{if } u = \frac{l}{2f_0}, |l| > q, q+1, \ldots, \\ \sum_{l=-\infty}^{\infty} C\left(\frac{l}{2f_0}\right) \frac{\sin(2\pi f_0 u - l\pi)}{(2\pi f_0 u - l\pi)} & \text{otherwise.} \end{cases} \tag{4.22}$$

This covariance function corresponds to a process limited to the frequency band $[-f_0, f_0]$, whose samples at regular intervals of length $\frac{1}{2f_0}$ 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 consider sampling with a stationary renewal process τ whose inter-sample spacing has the exponential distribution left-truncated at d, such that the mean is $d + \theta$. 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 estimator $\widehat{\psi}_n(\cdot)$ given in (1.14) as

$$\widehat{\psi}_n(f) = \frac{2}{n\beta} \sum_{l=1}^{n-1} \sum_{j=1}^{n-l} X(t_j)X(t_{j+l})K(b_n(t_{j+l} - t_j)) \cos(2\pi f(t_{j+l} - t_j)),$$

under the constraint of minimum separation, for the choices

$$\begin{aligned} f_0 &= 1, \\ \Xi(z) &= (1 + 1.2z)^8, \\ \sigma &= 1/20^2, \\ n &= 1000, \\ b_n &= 1/10, \\ \text{and } K(x) &= \begin{cases} \frac{1}{2}\{1 + \cos(\pi x)\} & \text{if } -1 \le x \le 1, \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

We first investigate how this estimator performs when $d > 0$. We conduct multiple simulation runs for each of the choices $d = 0$, 0.25, 0.5, 0.75, 1 and 2, together with $\theta = 1$. Figure 4.3 shows spectrum estimates from five typical simulation runs, along with the true power spectral density. The plots show how the estimator begins to perform poorly as one moves away from $d = 0$. For large values of d, the inter-sample spacing has a small coefficient of variation. Therefore, the sampled data resemble that from uniform sampling, which have the problem of aliasing. As a result, for large values of d, spurious peaks in greater numbers begin to show in the estimates. The estimator also assumes large negative values for some values of d.

Figure 4.4 shows the MSE of the estimate computed in each of the above cases from 500 simulated runs, along with the squared power spectral density. It is clear that the MSE around the peak of the power spectral density are of the same order as the squared power spectral density for large value of d, and the MSE at other frequencies is much

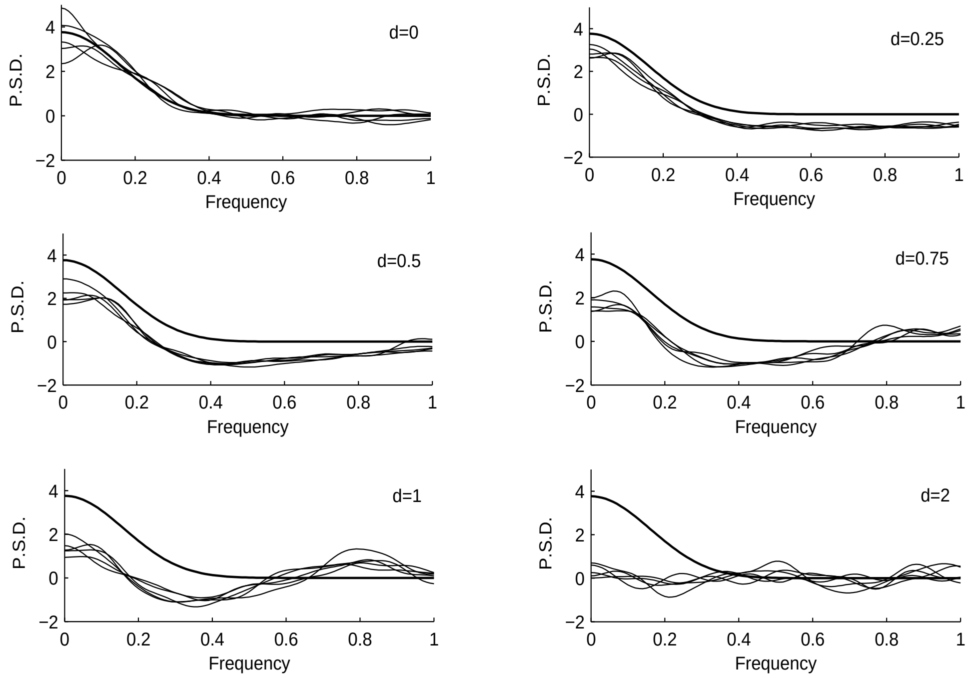

Figure 4.3: Estimates of the power spectral density for $\theta = 1$ and different values of d. The bold line represents the true power spectral density, while the thinner lines represent five typical estimates.

larger for $d > 0$ than for $d = 0$.

This simulation study indicates that the estimator, which is consistent in the absence of the constraint on the minimum inter-sample spacing, may perform poorly in the presence of the constraint.

The next question we try to answer is: Given the constraint $d = 1$ (so that uniform sampling at any feasible sampling rate would necessarily lead to aliasing), is there an appropriate choice of θ that would produce a reasonable estimate of the power spectral density? In order to answer this question, we again run multiple simulations for $\theta = 0$, 0.1, 0.2, 0.5, 1, 2, 5, and 10. In Figure 4.5, we present spectrum estimates from five typical simulation runs in each of these cases, together with the true power spectral density. For $\theta = 0$, i.e., the case of uniform sampling at sub-Nyquist rate, there is clear evidence of spurious peaks in the spectrum estimates. A similar occurrence is observed for small positive values of θ. On the other hand, large values of θ give rise to large variability in the estimates.

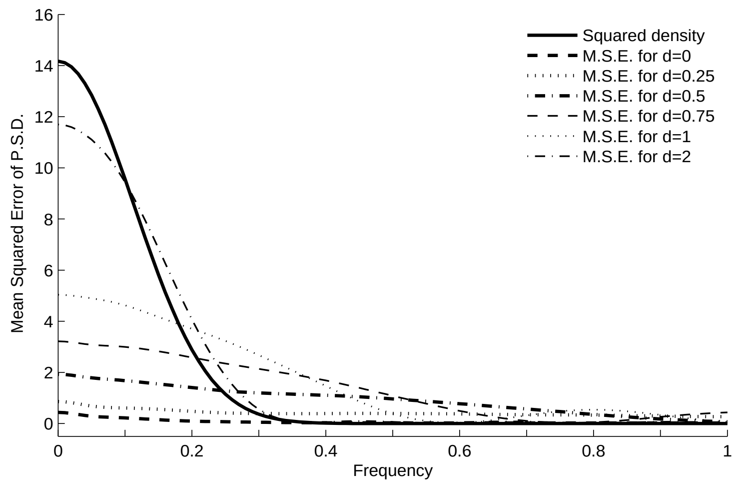

Figure 4.4: Plot of the true squared power spectral density and the MSE's of spectrum estimates (in log scale) based on 500 simulation runs for $\theta = 1$ and different values of d.

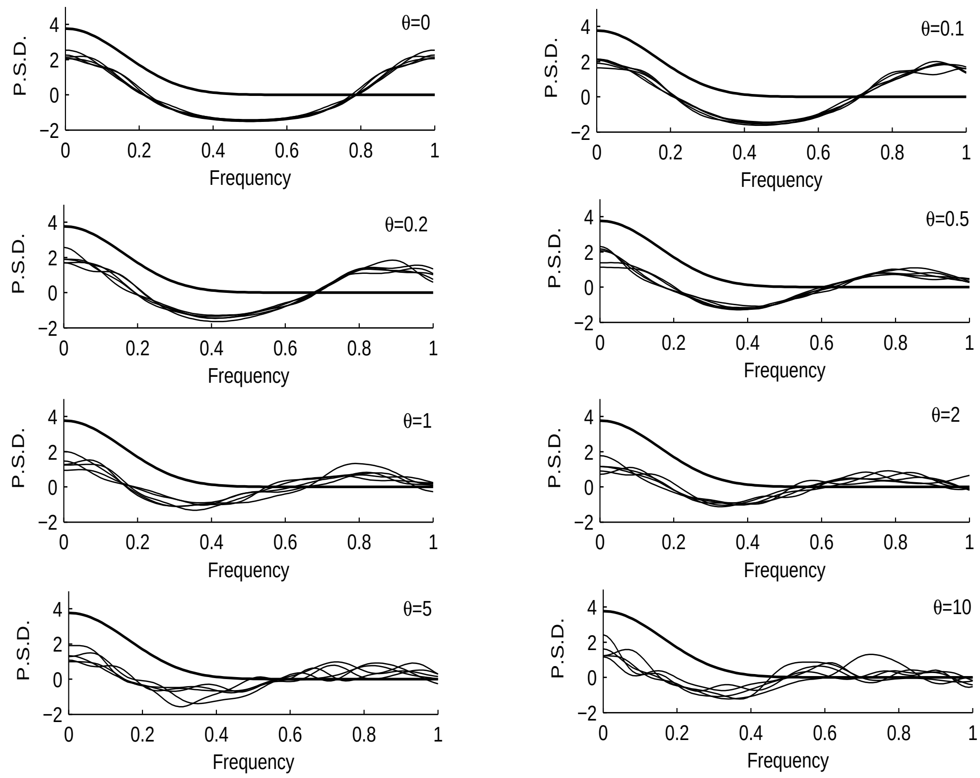

Figure 4.5: Estimates of the power spectral density for $d = 1$ and different values of θ. The bold line represents the true power spectral density, while the thinner lines represent five typical estimates.

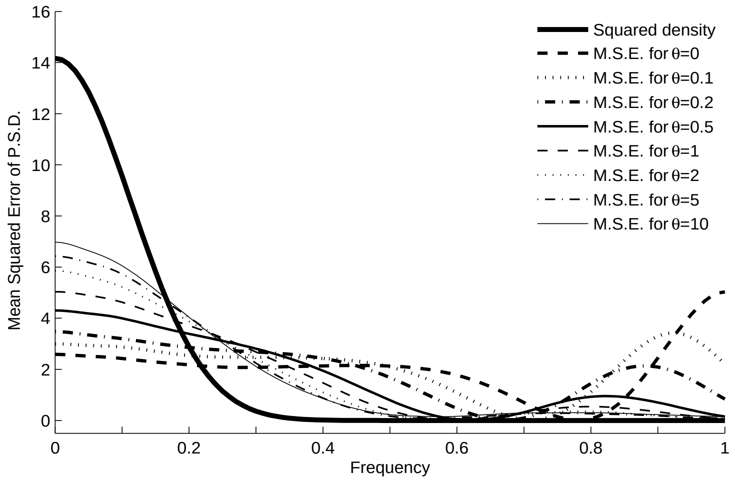

Figure 4.6: Plot of the true squared power spectral density and the MSE's of spectrum estimates (in log scale) based on 500 simulation runs for $d = 1$ and different values of θ.

Figure 4.6 shows MSE's of the estimates computed in each of the above cases from 500 simulated runs, along with the squared power spectral density. It transpires that irrespective of the trade-off between bias and variance observed in Figure 4.5, the MSE's in all the cases are comparable. The MSE is of the order of the squared value of the true power spectral density around the peak, while it is several orders of magnitude higher elsewhere.

These findings indicate that the estimator $\widehat{\psi}_n(\cdot)$ given in (1.14) does not perform well for any choice of θ under the conditions of the present simulation study.

4.6 Summary and discussion

The constraint of a specified minimum inter-sample spacing is a natural one, in view of technological and economic constraints. We have come across some interesting findings after formally incorporating this constraint in the study of aliasing in the context of spectrum estimation through stochastic sampling. The most important finding is that under this constraint, no point process sampling scheme is alias-free for the class of all spectra – according to any definition. It should be noted that the possibility of alias-free sampling, leading to consistent estimation of the power spectral density, has been a major argument put forward by some researchers in favour of point process sampling (in contrast with uniform sampling). This argument does not hold at all in the presence of the above constraint.

We have shown in Section 4.3 that when the inter-sample spacing is constrained to be larger than a threshold, renewal process sampling schemes are alias-free for suitably bandlimited spectra according to Definition 4.1. The range of bandwidths for alias-free renewal process sampling for some inter-sample spacing distributions is smaller than the corresponding range for regular sampling, while it is larger for some other distributions. On the other hand, according to Definition 4.2, all renewal process samplings schemes satisfying the conditions of Theorem 4.5 are alias-free for the class of spectra limited to any finite band.

The simulation study of Section 4.5 indicates that the estimator $\widehat{\psi}_n(\cdot)$ described in (1.14) performs poorly whenever Poisson sampling times are adjusted to ensure a minimum separation between successive samples. This finding is not unexpected, as the conditions under which Masry (1978b) proved the consistency of this estimator do not hold in the presence of this constraint. The constraint-adjusted sampling scheme constitutes a renewal process of a special kind. As we shall see in next chapter, a recipe for constructing consistent estimators for general renewal processes, given by Masry (1978a; see also (Brillinger, 1972)), does not work for such renewal processes (i.e., those which satisfy the constraint of a minimum separation between successive samples). This brings us back to square one as far as spectrum estimation (through sampling under this restriction) is concerned. A new approach will be required.

Chapter 5

Stochastically sampled bandlimited processes: Estimation under inter-sample spacing constraint

5.1 Introduction

We have seen in Chapter 4 that, under the constraint of a minimum separation between successive samples, no point process sampling scheme is alias-free for the class of non-bandlimited processes. This result implies that there can be two processes with different spectra, which cannot be distinguished from their respective point process samples, and hence the question of consistent spectrum estimation from these samples does not arise.

Theorem 4.4 and 4.5 open up the possiblility that, under the above constraint, additive random sampling (i.e., sampling at the renewal epochs of a stationary renewal process) can be alias-free for a class of spectra with a larger bandwidth than what is permitted by the Nyquist theorem – the bench-mark arising from uniform sampling. However, the results of Monte Carlo simulations reported in Section 4.5 indicate that the estimator (1.14) may not work well when the underlying Poisson sampling is modified as per the constraint.

We have also observed in Chapter 4 that the power spectral density of a continuous time process may be reconstructed from the covariance sequence c or the covariance measure μ_z of its stochastically sampled versions, under appropriate conditions, depending on the notion of alias-free sampling being used. These reconstruction methods have been used by Beutler (1970) and Masry (1978a) to propose spectrum estimators based on empirical versions of c and μ_z, respectively. One has to examine the appropriateness of these estimation approaches for arbitrarily bandlimited power spectral densities under the constraint of minimum separation between successive samples.

In this Chapter, we investigate these issues and conclude that the existing estimation approaches are not adequate for the problem at hand, for various reasons. Subsequently, we develop a new method of consistent estimation of an arbitrarily bandlimited power spectral density, using data obtained from additive random sampling under the said constraint.

In Section 5.2, we present the existing estimation approaches and discuss the difficulties under the constraint of minimum separation between successive samples. In Section 5.3, we develop a spectrum estimator based on constrained additive random sampling, and establish its consistency. The proofs of the theoretical results are given in Section 5.4. In Section 5.5, we study the performance of this estimator through Monte Carlo simulations. In the absence of a competing estimator, we contrast this performance with that of the estimator $\widehat{\psi}_n(\cdot)$ given in (1.14). We summarize the findings and provide some concluding remarks in Section 5.6.

5.2 Existing estimation approaches and their limitations

Consider a class of spectra having density supported on the closed and finite interval I. Given a sampling scheme that is alias-free relative to this class according to Definition 4.1, one would look for an estimator of the power spectral density based on estimated values of the sequence c. Beutler (1970) outlined a method of estimation based on the representation

$$\Phi(f_0) = \lim_{m\to\infty} \sum_{k=1}^{m} a_{km} c_k, \tag{5.1}$$

for a continuity point f_0 of the spectral distribution function, where a_{km}, $k = 1, \ldots, m$ are the coefficients such that the uniform convergence of the sequence of partial sums

$$\lim_{m\to\infty} \sum_{k=1}^{m} a_{km}\eta_k(f) = 1_{\{(-\infty,f_0)\cap I\}}(f) \tag{5.2}$$

happens everywhere except perhaps at f_0. Here, for $k = 1, 2, \ldots$, $\eta_k(\cdot)$ is the characteristic function of the distribution of the sum of k successive inter-sample spacings. As an example, for the inter-sample spacing distribution mentioned in the proof of Theorem 4.4, the characteristic function happens to be

$$\eta_k(f) = [0.68e^{i2\pi fd} + 0.32e^{i2\pi f\times 2.1d}]^k.$$

Note that the representation (5.2) is possible whenever the sampling scheme is alias-free according to Definition 4.1 (Beutler, 1970).

One can estimate the spectral distribution $\Phi(\cdot)$ by plugging in estimators of $c_1, c_2, \ldots, c_m$ in (5.1), and can subsequently obtain an estimator of the spectral density $\phi(\cdot)$. The coefficients $a_{1m}, \ldots, a_{mm}$ defined by (5.2), however, are attributes of the sampling scheme. These coefficients can be obtained numerically, either directly from (5.2) or by using a Gram-Schmidt orthogonalization of the characteristic functions as suggested in (Masry, 1978a). There is no closed form solution. In any case, the consistency of the plug-in estimator based on (5.1) has not been proved. Moreover, the largest spectral support I corresponding to this inter-sample spacing distribution is only marginally larger than the Nyquist limit $[-1/2d, 1/2d]$ achievable through uniform sampling. These facts prompt us to abandon Beutler's estimation approach.

We now explain the estimation approach based on the covariance measure μ_z (Masry, 1978a) and the difficulty of using this approach under the constraint of minimum separation between successive samples. We have seen in Section 4.2 that the covariance measure μ_z of the process Z has the representation (4.4), i.e.,

$$\mu_z(B) = \int_B C(u)[\beta^2 du + \mu_\tau(du)],$$

where β is the mean intensity and μ_τ is the reduced covariance measure of the sampling process τ. If τ is a stationary renewal process, then μ_τ has the simplified expression

$$\mu_\tau(B) = \beta\delta_0(B) + \beta\int_B (H(|u|) - \beta)du,$$

where $H(\cdot)$ is the renewal density defined over the positive real line as

$$H(u) = \sum_{l=1}^{\infty} \gamma^{(l)}(u), \tag{5.3}$$

$\gamma^{(l)}(\cdot)$ is the l-fold convolution of the inter-sample spacing density $\gamma(\cdot)$, and δ_0 is the degenerate measure having unit mass at the point 0. Consequently, the expression for μ_z simplifies to

$$\mu_z(du) = \beta C(u)H(|u|)du + \beta C(0)\delta_0(du). \tag{5.4}$$

If the renewal density is strictly positive, i.e., $H(u) > 0$ for $u > 0$, then

$$C(u)du = \frac{1}{\beta H(|u|)}\mu_z(du) - \frac{C(0)}{H(|u|)}\delta_0(du), \tag{5.5}$$

and the corresponding power spectral density is given as

$$\phi(f) = \int_{-\infty}^{\infty} e^{-i2\pi fu}C(u)du = \frac{1}{\beta}\int_{-\infty}^{\infty} \frac{1}{H(|u|)}e^{-i2\pi fu}\mu_z(du) - \frac{C(0)}{H(0)}, \tag{5.6}$$

provided the integral on the right hand side exists.

An estimator of $\phi(\cdot)$ may be obtained from the above expression by plugging in estimators of $C(0)$ and the covariance measure μ_z. In particular, μ_z can be estimated through its characteristic function $\phi_z(\cdot) = \int_{-\infty}^{\infty} e^{-i2\pi fu}\mu_z(du)$. Masry (1978a) suggested that $\phi_z(\cdot)$ may be estimated by

$$\hat{\phi}_z(f) = \int_{-\infty}^{\infty} \kappa_n(f-\nu)\hat{I}_n(\nu)d\nu,$$

where

$$\kappa_n(f) = \int_{-\infty}^{\infty} e^{-i2\pi fu}K(b_n u)du, \qquad \hat{I}_n(f) = \frac{\beta}{n}\left|\sum_{k=1}^{n} X(t_k)e^{-i2\pi f t_k}\right|^2,$$

$K(\cdot)$ is a suitable window function and b_n is the corresponding window width.

In the special case of Poisson sampling, we have $H(u) = \beta$ for all u, so that (5.6) simplifies to

$$\phi(f) = \frac{1}{\beta^2}\int_{-\infty}^{\infty} e^{-i2\pi f u}\mu_z(du) - \frac{C(0)}{\beta} = \frac{1}{\beta^2}\phi_z(f) - \frac{C(0)}{\beta},$$

and the corresponding plug-in spectrum estimator simplifies to

$$\frac{2}{\beta n}\sum_{j=1}^{n-1}\sum_{l=1}^{n-j} X(t_l)X(t_{j+l})K\left(b_n(t_j + l - t_l)\right)\cos(2\pi f(t_j + l - t_l)),$$

which is the estimator $\widehat{\psi}_n(\cdot)$, defined in (1.14).

In the special case of renewal process sampling subject to the constraint that the separation between successive samples is at least d units of time, the renewal density $H(u) = 0$ for $u \in [0, d]$. In this situation, a direct representation of $C(\cdot)$ similar to (5.5) is not possible. Consequently, there is no scope of a plug-in estimator, as above.

A new estimation strategy is developed in the next section.

5.3 Spectrum estimation under inter-sample spacing constraint

5.3.1 The estimator

Even though a direct representation of the covariance function $C(\cdot)$ similar to (5.5) for the entire real line is not possible, one can still use (5.5) for $u \in (d, \infty)$. We have seen in Theorem 4.5 that this part of the function contains complete information about $C(\cdot)$ over its entire domain. A possible way of recovering the missing information for the range $(0, d]$ is to use the representation of $C(\cdot)$ in terms of its values over a grid.

It had been mentioned in Chapter 1 (see (1.9)) that the covariance function $C(\cdot)$ of a continuous time, mean square continuous, stationary stochastic process X having power spectral density limited to a finite band $[-f_0, f_0]$, has the representation

$$C(u) = \sum_{l=-\infty}^{\infty} C(lT)\text{sinc}\left(\frac{\pi}{T}(u - lT)\right), \tag{5.7}$$

where $T = \frac{1}{2f_0}$. Therefore, in order to specify the function $C(\cdot)$ completely, one only needs to specify the sequence $\{C(lT), l = \ldots, -2, -1, 0, 1, 2, \ldots\}$. By using equation

(5.5), one can directly estimate the values $C(lT)$ for $l = 0$ and $l > J$, where $J = [d/T]$, the integer part of d/T. The remaining values, i.e., $C(T), \ldots, C(JT)$, can be expressed in terms of the left hand side of (5.7) and the known terms of the right hand side. Note that the left hand side of (5.7) can also be estimated directly for any $u > d$. Thus, the missing values satisfy the linear equations

$$\sum_{l=1}^{J} x_l(u)C(lT) = y(u) \quad \text{for } u > d, \tag{5.8}$$

where

$$\begin{aligned} x_l(u) &= \left\{\operatorname{sinc}\left(\frac{\pi}{T}(u - lT)\right) + \operatorname{sinc}\left(\frac{\pi}{T}(u + lT)\right)\right\}, \quad l = 1, 2, \ldots, J, & (5.9)\\ y(u) &= C(u) - \operatorname{sinc}(\pi u/T)C(0) \\ &\quad - \sum_{l=J+1}^{\infty} \left\{\operatorname{sinc}\left(\frac{\pi}{T}(u - lT)\right) + \operatorname{sinc}\left(\frac{\pi}{T}(u + lT)\right)\right\} C(lT), & (5.10) \end{aligned}$$

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

For direct estimation of $C(u)$ for $u = 0$ and $u > d$, we assume that the renewal density $H(\cdot)$ defined in (5.3) is strictly positive over the interval (d, ∞), and use the estimators

$$\begin{aligned} \widehat{C}_{nd}(0) &= \frac{1}{n}\sum_{k=1}^{n} X^2(t_k), & (5.11)\\ \widehat{C}_{nd}(u) &= \frac{1}{nH(u)}\sum_{k=1}^{n}\sum_{j=1}^{n} m_n W(m_n(u - t_k + t_j))X(t_k)X(t_j), \quad \text{for } u > d, & (5.12) \end{aligned}$$

where $W(\cdot)$ is a weight function and m_n is the smoothing parameter.

For indirect estimation of $C(T), \ldots, C(JT)$ from the linear equations (5.8), we define for $u > d$

$$\begin{aligned} y_n(u) &= \widehat{C}_{nd}(u) - \operatorname{sinc}(\pi u/T)\widehat{C}_{nd}(0) \\ &\quad - \sum_{l=J+1}^{L_n} \left\{\operatorname{sinc}\left(\frac{\pi}{T}(u - lT)\right) + \operatorname{sinc}\left(\frac{\pi}{T}(u + lT)\right)\right\} \widehat{C}_{nd}(lT), \end{aligned} \tag{5.13}$$

where L_n is a finite integer. We use $y_n(u)$ as an estimator of $y(u)$ defined in (5.10), with the infinite sum truncated at L_n. Substitution of this estimator on the right hand side of (5.8) gives a set of approximately linear equations in $C(T), \ldots, C(JT)$. This 'functional data' linear model (Ramsay et al., 2009) leads to the least squares estimator

$$\begin{pmatrix} \widehat{C}_{nd}(T) \\ \widehat{C}_{nd}(2T) \\ \vdots \\ \widehat{C}_{nd}(JT) \end{pmatrix} = \begin{pmatrix} \int_{u_1}^{u_2} x_1^2(u)du & \int_{u_1}^{u_2} x_1(u)x_2(u)du & \cdots & \int_{u_1}^{u_2} x_1(u)x_J(u)du \\ \int_{u_1}^{u_2} x_2(u)x_1(u)du & \int_{u_1}^{u_2} x_2^2(u)du & \cdots & \int_{u_1}^{u_2} x_2(u)x_J(u)du \\ \vdots & \vdots & \ddots & \vdots \\ \int_{u_1}^{u_2} x_J(u)x_1(u)du & \int_{u_1}^{u_2} x_J(u)x_2(u)du & \cdots & \int_{u_1}^{u_2} x_J^2(u)du \end{pmatrix}^{-1} \times \begin{pmatrix} \int_{u_1}^{u_2} x_1(u)y_n(u)du \\ \int_{u_1}^{u_2} x_2(u)y_n(u)du \\ \vdots \\ \int_{u_1}^{u_2} x_J(u)y_n(u)du \end{pmatrix}, \tag{5.14}$$

where the interval $[u_1, u_2]$ is a suitable sub-interval of (d, ∞).

In summary, the estimator of $C(\cdot)$ is

$$\widehat{C}_{nd}(u) = \begin{cases} \frac{1}{n}\sum_{j=1}^{n} X^2(t_j) & \text{if } u = 0, \\ \frac{1}{nH(u)}\sum_{k=1}^{n}\sum_{j=1}^{n} m_n W(m_n(u - t_k + t_j))X(t_k)X(t_j) & \\ & \text{if } u > d, \\ \text{as defined in (5.14)} & \text{if } u = lT \text{ for some integer } l \text{ and } 0 < u < d, \\ \widehat{C}_{nd}(0)\text{sinc}\left(\frac{\pi}{T}u\right) + \sum_{j=1}^{L_n} \widehat{C}_{nd}(jT)\left\{\text{sinc}\left(\frac{\pi}{T}(u - jT)\right) + \text{sinc}\left(\frac{\pi}{T}(u + jT)\right)\right\} & \\ & \text{if } u \neq lT \text{ for any integer } l \text{ and } 0 < u < d, \\ \widehat{C}_{nd}(-u) & \text{if } u < 0. \end{cases} \tag{5.15}$$

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

$$\widehat{\phi}_{nd}(f) = T\widehat{C}_{nd}(0) + 2T\sum_{l=1}^{n-1} \widehat{C}_{nd}(lT)K(b_n l)\cos(2\pi f lT)1_{[-f_0, f_0]}(f), \tag{5.16}$$

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

5.3.2 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 need to make some assumptions regarding the underlying process, the weight function used for estimating its covariance function, the smoothing parameter used for this weight function and the additive random sampling scheme.

Assumption 5.1. *The function $h_0(\cdot)$ defined over the real line as $h_0(u) = \sup_{t\geq|u|} |C(t)|$ is integrable.*

Assumption 5.2. *The inter-sample spacings density $\gamma(\cdot)$ has a finite mean and its support contains an interval with d at its left endpoint.*

Note that Assumption 5.2 ensures that $H(\cdot)$ is positive over (d, ∞), as required in (5.12).

Assumption 5.3. *The weight function $W(\cdot)$ used in (5.12) is compactly supported, even, continuous and square integrable, with $\int_{-\infty}^{\infty} W(v)dv = 1$.*

Assumption 5.4. *The smoothing parameter m_n used in (5.12) is such that $m_n \to \infty$ as $n \to \infty$.*

Assumption 5.5. *The truncation parameter L_n used in (5.13) is such that $L_n \to \infty$ as $n \to \infty$.*

Theorem 5.1. *Under Assumptions 5.1–5.5, the bias of the estimator $\widehat{C}_{nd}(\cdot)$ converges pointwise to zero as the sample size n goes to infinity.*

A few additional conditions are needed for proving convergence of the variance of the covariance estimator.

Assumption 5.5A. *The truncation parameter L_n used in (5.13) is such that $L_n \to \infty$ and $\frac{m_n(\log L_n)^2}{n} \to 0$ as $n \to \infty$.*

Assumption 5.6. *The fourth moment $E|X(t)|^4$ exists for every t, and the fourth order cumulant function cum $[X(t+t_1), X(t+t_2), X(t+t_3), X(t)]$ 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 $[0, \infty)$.

Remark 5.1. *Assumption 5.5A is stronger than Assumption 5.5. It also implies that $\frac{m_n}{n} \to 0$ as $n \to \infty$.*

Theorem 5.2. *Under Assumptions 5.1–5.4, 5.5A and 5.6, the variance of the estimator $\widehat{C}_{nd}(\cdot)$ converges pointwise to zero as the sample size n goes to infinity.*

Theorems 5.1 and 5.2 together imply that under Assumptions 5.1–5.4, 5.5A and 5.6, $\widehat{C}_{nd}(u)$ is a consistent estimator of the covariance function $C(u)$ for any lag u.

We now turn to the convergence of the bias of the spectrum estimator $\widehat{\phi}_{nd}(f)$ given in (5.16), for which we need to make further assumptions.

Assumption 5.7. *The covariance averaging kernel $K(\cdot)$ used in (5.16) is an even, continuous, integrable and square integrable function with $K(0) = 1$, and is bounded by a nondecreasing function over $(0, \infty)$.*

Assumption 5.8. *The kernel bandwidth b_n used in (5.16) is such that $b_n \to 0$ as $n \to \infty$.*

Theorem 5.3. *Under Assumptions 5.1–5.5, 5.7 and 5.8, the bias of the estimator $\widehat{\phi}_{nd}(\cdot)$ converges pointwise to zero as the sample size n goes to infinity.*

In order to prove the convergence of the variance of the spectrum estimator, we replace Assumption 5.8 by a stronger assumption.

Assumption 5.8A. *The kernel bandwidth b_n used in (5.16) is such that $b_n \to 0$, $nb_n \to \infty$ and $\frac{m_n}{b_n} \to 0$ as $n \to \infty$.*

Theorem 5.4. *Under Assumptions 5.1–5.4, 5.5A, 5.6, 5.7 and 5.8A, the variance of the estimator* $\widehat{\phi}_{nd}(\cdot)$ *converges pointwise to zero as the sample size* n *goes to infinity.*

Theorem 5.3 and 5.4 show that, under Assumptions 5.1–5.4, 5.5A, 5.6, 5.7 and 5.8A, $\widehat{\phi}_{nd}(f)$ is a consistent estimator of the power spectral density $\phi(f)$ for any frequency f.

5.4 Proofs

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.

Proof of Theorem 5.1. It is enough to prove that the bias of the estimator $\widehat{C}_{nd}(u)$ converges to 0 for $u \geq 0$. Observe that

$$E[\widehat{C}_{nd}(0)] = \frac{1}{n}\sum_{j=1}^{n} E[X^2(t_j)] = C(0), \tag{5.17}$$

i.e., $\widehat{C}_{nd}(0)$ is an unbiased estimator of $C(0)$. For $u > d$, using (5.12) we have

$$\begin{aligned}
E[\widehat{C}_{nd}(u)] =& \frac{1}{nH(u)}\sum_{k=1}^{n}\sum_{j=1}^{n} m_n E[E\left\{W(m_n(u - t_k + t_j))X(t_k)X(t_j)|t_k,\ k = 1, \ldots, n\right\}] \\
=& \frac{1}{nH(u)}\sum_{k=1}^{n}\sum_{j=1}^{n} m_n E[W(m_n(u - t_k + t_j))C(t_k - t_j)].
\end{aligned}$$

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

$$\begin{aligned}
&E[\widehat{C}_{nd}(u)] \\
&\quad = \frac{m_n W(m_n u)C(0)}{H(u)} \\
&\qquad + \frac{1}{H(u)}\int_0^\infty m_n\{W(m_n(u+v)) + W(m_n(u-v))\}C(v)\left\{\frac{1}{n}\sum_{1\leq k<j\leq n}\gamma^{(j-k)}(v)\right\}dv \\
&\quad = \frac{m_n W(m_n u)C(0)}{H(u)} + \frac{1}{H(u)}\int_0^\infty m_n\{W(m_n(u+v)) + W(m_n(u-v))\}C(v)H_n(v)dv,
\end{aligned}$$

where

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

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

$$E[\widehat{C}_{nd}(u)]$$
$$= \frac{m_n W(m_n u) C(0)}{H(u)} + \frac{1}{H(u)} \int_{-\infty}^{\infty} W(v) C\left(u - \frac{v}{m_n}\right) \left\{ H_n\left(u - \frac{v}{m_n}\right) + H_n\left(-u + \frac{v}{m_n}\right) \right\} dv. \tag{5.18}$$

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. Further, by using Assumptions 5.1 and 5.2, we have the dominance

$$\begin{aligned} &\left| W(v) C\left(u - \frac{v}{m_n}\right) \left\{ H_n\left(u - \frac{v}{m_n}\right) + H_n\left(-u + \frac{v}{m_n}\right) \right\} \right| \\ &\quad \leq \ |W(v)| h_0(0) \left\{ H\left(u - \frac{v}{m_n}\right) + H\left(-u + \frac{v}{m_n}\right) \right\} \\ &\quad \leq \ 2|W(v)| h_0(0) M_2, \end{aligned}$$

and Assumption 5.3 ensures that the bounding function is integrable. From Assumption 5.4, we have the convergence

$$\begin{aligned} &\lim_{n\to\infty} W(v) C\left(u - \frac{v}{m_n}\right) \left\{ H_n\left(u - \frac{v}{m_n}\right) + H_n\left(-u + \frac{v}{m_n}\right) \right\} \\ &\quad = \ W(v) C(u) H(u). \end{aligned}$$

By applying the DCT, we have

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

Now, we consider the case of $\widehat{C}_{nd}(u)$, when $u = T, 2T, \ldots, JT$. In order to compute expectation of the indirect estimator $(\widehat{C}_{nd}(T), \ldots, \widehat{C}_{nd}(JT))$ given in (5.14), 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

$$\begin{aligned} E[y_n(u)] &= E[\widehat{C}_{nd}(u)] - \operatorname{sinc}(\pi u/T)E[\widehat{C}_{nd}(0)] \\ &\quad - \sum_{l=J+1}^{L_n} E[\widehat{C}_{nd}(lT)] \left\{\operatorname{sinc}\Big(\frac{\pi}{T}(u-lT)\Big)+\operatorname{sinc}\Big(\frac{\pi}{T}(u+lT)\Big)\right\}. \end{aligned}$$

We have already proved that $E[\widehat{C}_{nd}(u)] \to C(u)$ as $n \to \infty$ for $u > d$ and $E[\widehat{C}_{nd}(0)] = C(0)$. From (5.18), we have

$$\begin{aligned} &\sum_{l=J+1}^{L_n} E[\widehat{C}_{nd}(lT)] \left\{\operatorname{sinc}\left(\frac{\pi}{T}(u-lT)\right)+\operatorname{sinc}\left(\frac{\pi}{T}(u+lT)\right)\right\} \\ &= \sum_{l=J+1}^{L_n} \left[\frac{m_n W(m_n lT) C(0)}{H(lT)}\right. \\ &\qquad \left. + \frac{1}{H(lT)} \int_{-\infty}^{\infty} W(v) C\left(lT - \frac{v}{m_n}\right) \left\{H_n\left(lT - \frac{v}{m_n}\right) + H_n\left(-lT + \frac{v}{m_n}\right)\right\} dv\right] \\ &\qquad \times \left\{\operatorname{sinc}\left(\frac{\pi}{T}(u-lT)\right)+\operatorname{sinc}\left(\frac{\pi}{T}(u+lT)\right)\right\}. \end{aligned}$$

Choose n sufficiently large so that $m_n > a/d$. Then we have

$$\begin{aligned} &\sum_{l=J+1}^{L_n} E[\widehat{C}_{nd}(lT)] \left\{\operatorname{sinc}\left(\frac{\pi}{T}(u-lT)\right)+\operatorname{sinc}\left(\frac{\pi}{T}(u+lT)\right)\right\} \\ &= \sum_{l=J+1}^{L_n} \left[\frac{1}{H(lT)} \int_{-\infty}^{\infty} W(v) C\left(lT - \frac{v}{m_n}\right) \left\{H_n\left(lT - \frac{v}{m_n}\right) + H_n\left(-lT + \frac{v}{m_n}\right)\right\} dv \right. \\ &\qquad \left. \times \left\{\operatorname{sinc}\Big(\frac{\pi}{T}(u-lT)\Big)+\operatorname{sinc}\Big(\frac{\pi}{T}(u+lT)\Big)\right\}\right] \\ &= \int_{-\infty}^{\infty} W(v) \left[\sum_{l=J+1}^{L_n} \frac{1}{H(lT)} C\left(lT - \frac{v}{m_n}\right) \left\{H_n\left(lT - \frac{v}{m_n}\right) + H_n\left(-lT + \frac{v}{m_n}\right)\right\}\right. \\ &\qquad \left. \times \left\{\operatorname{sinc}\Big(\frac{\pi}{T}(u-lT)\Big)+\operatorname{sinc}\Big(\frac{\pi}{T}(u+lT)\Big)\right\}\right] dv. \end{aligned} \tag{5.19}$$

By using Assumptions 5.1 and 5.2, we have the dominance

$$\begin{aligned}
&\Bigg| W(v)\sum_{l=J+1}^{L_n}\frac{1}{H(lT)}C\left(lT-\frac{v}{m_n}\right)\left\{H_n\left(lT-\frac{v}{m_n}\right)+H_n\left(-lT+\frac{v}{m_n}\right)\right\} \\
&\qquad\times\left\{\mathrm{sinc}\Big(\frac{\pi}{T}(u-lT)\Big)+\mathrm{sinc}\Big(\frac{\pi}{T}(u+lT)\Big)\right\}\Bigg| \\
&\qquad\le\ |W(v)|\frac{2M_2}{M_3}\sum_{J+1}^{L_n}h_0\left(lT-\frac{v}{m_n}\right)\ \le\ |W(v)|\frac{2M_2}{M_3}\times 2\sum_{l=1}^{\infty}h_0(lT),
\end{aligned}$$

and the integrability of the bound is guaranteed by Assumption 5.3. From Assumptions 5.4 and 5.5, the integrand of (5.19) converges pointwise as

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

Thus, by using the representation (5.7) and the DCT, we have

$$\begin{aligned}
&\lim_{n\to\infty}E[y_n(u)] \\
&= C(u)-\mathrm{sinc}(\pi u/T)C(0)-\sum_{l=J+1}^{\infty}C(lT)\Big\{\mathrm{sinc}\Big(\frac{\pi}{T}(u-lT)\Big)+\mathrm{sinc}\Big(\frac{\pi}{T}(u+lT)\Big)\Big\} \\
&= \sum_{l=1}^{J}x_l(u)C(lT). \qquad (5.20)
\end{aligned}$$

Thus for $j=1,\dots,J$, we have

$$\lim_{n\to\infty}E\left[\int_{u_1}^{u_2}x_j(u)y_n(u)du\right]=\int_{u_1}^{u_2}x_j(u)\lim_{n\to\infty}E[y_n(u)]du=\int_{u_1}^{u_2}x_j(u)\sum_{l=1}^{J}x_l(u)C(lT). \qquad (5.21)$$

By using (5.21) and (5.14), we have

$$\lim_{n\to\infty}E\begin{pmatrix}\widehat{C}_{nd}(T)\\ \widehat{C}_{nd}(2T)\\ \vdots\\ \widehat{C}_{nd}(JT)\end{pmatrix}=\begin{pmatrix}C(T)\\ C(2T)\\ \vdots\\ C(JT)\end{pmatrix}. \qquad (5.22)$$

Consider $\widehat{C}_{nd}(u)$, when $0 < u < d$ and $u \neq lT$ for any integer l. From (5.15), we have

$$\begin{aligned}
&E[\widehat{C}_{nd}(u)] \\
&= E\left[\widehat{C}_{nd}(0)\right] \operatorname{sinc}\left(\frac{\pi}{T}u\right) + \sum_{j=1}^{J} E\left[\widehat{C}_{nd}(jT)\right]\left\{\operatorname{sinc}\left(\frac{\pi}{T}(u-jT)\right) + \operatorname{sinc}\left(\frac{\pi}{T}(u+jT)\right)\right\} \\
&\qquad + \sum_{j=J+1}^{L_n} E\left[\widehat{C}_{nd}(jT)\right]\left\{\operatorname{sinc}\left(\frac{\pi}{T}(u-jT)\right) + \operatorname{sinc}\left(\frac{\pi}{T}(u+jT)\right)\right\}.
\end{aligned}$$

By using an argument similar to that used in proving (5.20), it can be shown that

$$\lim_{n\to\infty} E[\widehat{C}_{nd}(u)] = C(u) \qquad \text{if } u \neq lT \text{ for any integer } l \text{ and } 0 < u < d.$$

This completes the proof. □

Proof of Theorem 5.2. It suffices to prove that the variance of the estimator $\widehat{C}_{nd}(u)$ converges to 0 for $u \geq 0$. Observe that

$$Var[\widehat{C}_{nd}(0)] = \frac{1}{n^2}\sum_{k_1=1}^{n}\sum_{k_2=1}^{n} E[X(t_{k_1})X(t_{k_2})] = \frac{C(0)}{n} + \frac{1}{n^2}\int_0^\infty C(v) \sum_{1\leq k_1<k_2\leq n} \gamma^{(k_2-k_1)}(v)dv.$$

Thus, we have, from Assumption 5.2

$$nVar[\widehat{C}_{nd}(0)] \leq C(0) + M_2 \int_0^\infty C(v)dv.$$

Thus we have,

$$\lim_{n\to\infty} Var[\widehat{C}_{nd}(0)] = 0. \tag{5.23}$$

By using Assumption 5.6, we have for $u > d$

$$\begin{aligned}
E[\widehat{C}_{nd}^2(u)] =& \frac{m_n^2}{n^2H^2(u)} \sum_{k_1=1}^{n}\sum_{k_2=1}^{n}\sum_{k_3=1}^{n}\sum_{k_4=1}^{n} E\Big[W\left(m_n(u - t_{k_1} + t_{k_2})\right) W\left(m_n(u - t_{k_3} + t_{k_4})\right) \\
&\qquad \times X(t_{k_1})X(t_{k_2})X(t_{k_3})X(t_{k_4})\Big] \\
=& \frac{m_n^2}{n^2H^2(u)} \sum_{k_1=1}^{n}\sum_{k_2=1}^{n}\sum_{k_3=1}^{n}\sum_{k_4=1}^{n} E\Big[W\left(m_n(u - t_{k_1} + t_{k_2})\right) W\left(m_n(u - t_{k_3} + t_{k_4})\right) \\
&\times \Big\{C(t_{k_1} - t_{k_2})C(t_{k_3} - t_{k_4}) + C(t_{k_1} - t_{k_3})C(t_{k_2} - t_{k_4}) \\
&\quad + C(t_{k_1} - t_{k_4})C(t_{k_2} - t_{k_3}) + Q(t_{k_1} - t_{k_4}, t_{k_2} - t_{k_4}, t_{k_3} - t_{k_4})\Big\}\Big].
\end{aligned}$$

Thus, we have for $u > d$,

$$\begin{aligned}
&Var[\widehat{C}_{nd}(u)] \\
&= \frac{m_n^2}{n^2H^2(u)} \sum_{k_1=1}^{n}\sum_{k_2=1}^{n}\sum_{k_3=1}^{n}\sum_{k_4=1}^{n} E\Big[W\left(m_n(u-t_{k_1}+t_{k_2})\right)W\left(m_n(u-t_{k_3}+t_{k_4})\right) \\
&\qquad \times\Big\{C(t_{k_1}-t_{k_3})C(t_{k_2}-t_{k_4})+C(t_{k_1}-t_{k_4})C(t_{k_2}-t_{k_3}) \\
&\qquad\quad +Q(t_{k_1}-t_{k_4},t_{k_2}-t_{k_4},t_{k_3}-t_{k_4})\Big\}\Big] \\
&= I_1(u)+I_2(u)+I_3(u). \qquad (5.24)
\end{aligned}$$

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

$$\begin{aligned}
I(u) &= M_1\frac{m_n}{n}\frac{1}{nH^2(u)} \\
&\quad \times\sum_{k_1=1}^{n}\sum_{k_2=1}^{n}\sum_{k_3=1}^{n}\sum_{k_4=1}^{n} E\Big[\left|m_nW\left(m_n(u-t_{k_1}+t_{k_2})\right)C(t_{k_1}-t_{k_3})C(t_{k_2}-t_{k_4})\right|\Big].
\end{aligned} \qquad (5.25)$$

It follows from Lemma 5.1 proved below that $\frac{n}{m_n}I(u)$ is bounded from above by a constant that does not depend on u. Lemma 5.2 proved below indicates that $\frac{n}{m_n}I_3(u)$ is bounded from above by a constant that does not depend on u. Therefore, for $u > d$, $Var(\widehat{C}_{nd}(u))$ converges to 0 as $n \to \infty$.

We now consider the indirect estimators $\widehat{C}_{nd}(T), \ldots, \widehat{C}_{nd}(JT)$. From (5.14), it is enough to show that the variance covariance matrix of the vector $\Big(\int_{u_1}^{u_2} x_1(u)y_n(u)du,$ $\int_{u_1}^{u_2} x_2(u)y_n(u)du, \;\ldots,\; \int_{u_1}^{u_2} x_J(u)y_n(u)du\Big)$ converges to zero.

For $j, j' \in \{1, 2, \ldots J\}$, we compute

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

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

$$
\begin{aligned}
&Var[y_n(u)] \\
&= Var\left[\widehat{C}_{nd}(u) - \mathrm{sinc}(\pi u/T)\widehat{C}_{nd}(0) - \sum_{l=J+1}^{L_n} \widehat{C}_{nd}(lT)\left\{\mathrm{sinc}\left(\frac{\pi}{T}(u-lT)\right) + \mathrm{sinc}\left(\frac{\pi}{T}(u+lT)\right)\right\}\right] \\
&\leq 3Var[\widehat{C}_{nd}(u)] + 3Var[\mathrm{sinc}(\pi u/T)\widehat{C}_{nd}(0)] \\
&\quad + 3Var\left[\sum_{l=J+1}^{L_n} \widehat{C}_{nd}(lT)\left\{\mathrm{sinc}\left(\frac{\pi}{T}(u-lT)\right) + \mathrm{sinc}\left(\frac{\pi}{T}(u+lT)\right)\right\}\right].
\end{aligned}
$$

We have already shown that the first two terms are uniformly bounded from above by $\frac{m_n}{n}M$ for some constant M. We now consider the third term.

$$
\begin{aligned}
&Var\left[\sum_{l=J+1}^{L_n} \widehat{C}_{nd}(lT)\left\{\mathrm{sinc}\left(\frac{\pi}{T}(u-lT)\right) + \mathrm{sinc}\left(\frac{\pi}{T}(u+lT)\right)\right\}\right] \\
&= \sum_{l=J+1}^{L_n}\sum_{l'=J+1}^{L_n} Cov(\widehat{C}_{nd}(lT), \widehat{C}_{nd}(l'T))\left\{\mathrm{sinc}\left(\frac{\pi}{T}(u-lT)\right) + \mathrm{sinc}\left(\frac{\pi}{T}(u+lT)\right)\right\} \\
&\qquad \times \left\{\mathrm{sinc}\left(\frac{\pi}{T}(u-l'T)\right) + \mathrm{sinc}\left(\frac{\pi}{T}(u+l'T)\right)\right\} \\
&\leq M\frac{m_n}{n}\left[\sum_{l=J+1}^{L_n}\left|\left\{\mathrm{sinc}\left(\frac{\pi}{T}(u-lT)\right) + \mathrm{sinc}\left(\frac{\pi}{T}(u+lT)\right)\right\}\right|\right]^2 \\
&\leq M'\frac{m_n}{n}(\log(L_n))^2, \qquad (5.26)
\end{aligned}
$$

where M' is a constant which does not depend on u.

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

$$
\begin{aligned}
&Cov\left(\int_{u_1}^{u_2} x_j(u)y_n(u)du, \int_{u_1}^{u_2} x_{j'}(u)y_n(u)du\right) \\
&\qquad = \int_{u_1}^{u_2}\int_{u_1}^{u_2} x_j(u)x_{j'}(v)Cov(y_n(u), y_n(v))dudv \\
&\qquad \leq M'\frac{m_n(\log(L_n))^2}{n}\left(\int_{u_1}^{u_2}|x_j(u)|du\right)\left(\int_{u_1}^{u_2}|x_{j'}(v)|dv\right).
\end{aligned}
$$

By using Assumption 5.5A, we have

$$
\lim_{n\to\infty} Cov\left(\int_{u_1}^{u_2} x_j(u)y_n(u)du, \int_{u_1}^{u_2} x_{j'}(u)y_n(u)du\right) = 0.
$$

Thus, we have

$$\lim_{n\to\infty} Cov(\widehat{C}_{nd}(jT), \widehat{C}_{nd}(j'T)) \quad = \quad 0 \quad \text{for } j, j' = 1, \ldots, J. \tag{5.27}$$

Now, we consider $Var(\widehat{C}_{nd}(u))$, when $u \neq lT$ for any integer l and $0 < u < d$. By using (5.15), we have

$$\begin{aligned} Var[\widehat{C}_{nd}(u)] \quad \leq \quad & 3Var\left[\widehat{C}_{nd}(0)\text{sinc}\left(\frac{\pi}{T}u\right)\right] \\ & +3Var\left[\sum_{j=1}^{J} \widehat{C}_{nd}(jT)\left\{\text{sinc}\left(\frac{\pi}{T}(u-jT)\right)+\text{sinc}\left(\frac{\pi}{T}(u+jT)\right)\right\}\right] \\ & +3Var\left[\sum_{j=J+1}^{L_n} \widehat{C}_{nd}(jT)\left\{\text{sinc}\left(\frac{\pi}{T}(u-jT)\right)+\text{sinc}\left(\frac{\pi}{T}(u+jT)\right)\right\}\right]. \end{aligned}$$

By using (5.23), (5.26) and (5.27), we have

$$\lim_{n\to\infty} Var[\widehat{C}_{nd}(u)] = 0 \qquad \text{if } u \neq lT \text{ for any integer } l \text{ and } 0 < u < d.$$

This completes the proof. □

Lemma 5.1. Under the Assumptions of Theorem 5.2, the function $I(\cdot)$ defined in (5.25) is uniformly bounded from above by $\frac{m_n}{n}M$, where M is a constant.

Proof of Lemma 5.1. We partition the range of summation as

$$\{(k_1, k_2, k_3, k_4) : 1 \leq k_1, k_2, k_3, k_4 \leq n\} = \bigcup_{j=1}^{24} S_{1,j} \bigcup_{j=1}^{36} S_{2,j} \bigcup_{j=1}^{8} S_{3,j} \bigcup S_4,$$

where $S_{1,j}$ for $j = 1, \ldots, 24$ are sets of quadruples of indices having different types of strict order among themselves ($k_1 < k_2 < k_3 < k_4$, $k_1 < k_2 < k_4 < k_3$, 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 ($k_1 < k_2 < k_3 = k_4$, $k_1 < k_3 = k_4 < k_2$, 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 ($k_1 < k_2 = k_3 = k_4$, $k_2 = k_3 = k_4 < k_1$, and 6 other arrangements), and S_4 is the set $\{(k_1, k_2, k_3, k_4) : 1 \leq k_1 = k_2 = k_3 = k_4 \leq n\}$.

Consider $S_{1,1} = \{(k_1, k_2, k_3, k_4) : 1 \leq k_1 < k_2 < k_3 < k_4 \leq n\}$. By using the transformation $t_{k_2} - t_{k_1} = \vartheta_{k_2-k_1}$, $t_{k_3} - t_{k_2} = \vartheta_{k_3-k_2}$ and $t_{k_4} - t_{k_3} = \vartheta_{k_4-k_3}$, and by making use of the fact that the transformed random variables are independent, we have

$$\begin{aligned}
&\frac{M_1}{nH^2(u)}\sum_{S_{1,1}} E\left[\,|m_n W\left(m_n(u - t_{k_1} + t_{k_2})\right) C(t_{k_1} - t_{k_3}) C(t_{k_2} - t_{k_4})|\,\right] \\
&= \frac{M_1}{nH^2(u)}\sum_{S_{1,1}} E\left[\,|m_n W\left(m_n(u - \vartheta_{k_2-k_1})\right) C(\vartheta_{k_3-k_2} + \vartheta_{k_2-k_1}) C(\vartheta_{k_4-k_3} + \vartheta_{k_3-k_2})|\,\right] \\
&= \frac{M_1}{nH^2(u)}\sum_{S_{1,1}} \int_0^\infty \int_0^\infty \int_0^\infty |m_n W\left(m_n(u - v_1)\right) C(v_2 + v_1) C(v_3 + v_2)| \\
&\qquad\qquad \times \gamma^{(k_2-k_1)}(v_1)\gamma^{(k_3-k_2)}(v_2)\gamma^{(k_4-k_3)}(v_3)\,dv_1 dv_2 dv_3 \\
&= \frac{M_1}{H^2(u)} \int_0^\infty \int_0^\infty \int_0^\infty |m_n W\left(m_n(u - v_1)\right) C(v_2 + v_1) C(v_3 + v_2)| \\
&\qquad\qquad \left\{\times \frac{1}{n}\sum_{S_{1,1}} \gamma^{(k_2-k_1)}(v_1)\gamma^{(k_3-k_2)}(v_2)\gamma^{(k_4-k_3)}(v_3)\right\} dv_1 dv_2 dv_3.
\end{aligned}$$

Now, note that

$$\frac{1}{n}\sum_{S_{1,1}} \gamma^{(k_2-k_1)}(v_1)\gamma^{(k_3-k_2)}(v_2)\gamma^{(k_4-k_3)}(v_3) \leq H(v_1)H(v_2)H(v_3) \leq M_2^3.$$

Thus, we have

$$\begin{aligned}
&\frac{M_1}{nH^2(u)}\sum_{S_{1,1}} E\left[\,|m_n W\left(m_n(u - t_{k_1} + t_{k_2})\right) C(t_{k_1} - t_{k_3}) C(t_{k_2} - t_{k_4})|\,\right] \\
&\leq \frac{M_1 M_2^3}{M_3^2} \int_0^\infty \int_0^\infty \int_0^\infty |m_n W\left(m_n(u - v_1)\right) C(v_2 + v_1) C(v_3 + v_2)|\, dv_1 dv_2 dv_3. \\
&= M_4 \int_0^\infty m_n |W\left(m_n(u - v_1)\right)| \left[\int_0^\infty |C(v_2 + v_1)| \left\{\int_0^\infty |C(v_3 + v_2)| dv_3\right\} dv_2\right] dv_1 \\
&\leq M_5 \int_0^\infty m_n |W\left(m_n(u - v_1)\right)| \left[\int_0^\infty |C(v_2 + v_1)| dv_2\right] dv_1 \\
&\leq M_6 \int_0^\infty m_n |W\left(m_n(u - v_1)\right)| dv_1 \quad \leq \quad M_7,
\end{aligned}$$

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

By using similar arguments, one can establish the boundedness of

$$\frac{M_1}{nH^2(u)}\sum_{S_{1,j}} E\left[\,|m_n W\left(m_n(u - t_{k_1} + t_{k_2})\right) C(t_{k_1} - t_{k_3}) C(t_{k_2} - t_{k_4})|\,\right]$$

for the partitions $S_{1,j}$, $2 \le j \le 22$. A slightly different argument is needed for $S_{1,23} = \{(k_1,k_2,k_3,k_4) : 1 \le k_1 < k_4 < k_3 < k_2 \le n\}$ and $S_{1,24} = \{(k_1,k_2,k_3,k_4) : 1 \le k_2 < k_3 < k_4 < k_1 \le n\}$. Consider the case of $S_{1,24}$. By using the transformation $\vartheta_{k_4-k_1} = t_{k_4} - t_{k_1}$, $\vartheta_{k_3-k_4} = t_{k_3} - t_{k_4}$ and $\vartheta_{k_2-k_3} = t_{k_2} - t_{k_3}$ and using the fact that $\vartheta_{k_4-k_1}$, $\vartheta_{k_3-k_4}$ and $\vartheta_{k_2-k_3}$ are independent random variables, we have

$$
\begin{aligned}
& \frac{M_1}{nH^2(u)} \sum_{S_{1,23}} E\big[\,|m_n W\left(m_n(u - t_{k_1} + t_{k_2})\right) C(t_{k_1} - t_{k_3}) C(t_{k_2} - t_{k_4})|\,\big] \\
= \;& \frac{M_1}{nH^2(u)} \sum_{S_{1,23}} E\Big[\,|m_n W\left(m_n(u + \vartheta_{k_4-k_1} + \vartheta_{k_3-k_4} + \vartheta_{k_2-k_3})\right)| \\
& \qquad \times |C(\vartheta_{k_4-k_1} + \vartheta_{k_3-k_4}) C(\vartheta_{k_3-k_4} + \vartheta_{k_2-k_3})|\Big] \\
= \;& \frac{M_1}{H^2(u)} \int_0^\infty \int_0^\infty \int_0^\infty |m_n W\left(m_n(u + v_1 + v_2 + v_3)\right) C(v_1 + v_2)| \\
& \qquad \times |C(v_2 + v_3)| \left\{ \frac{1}{n} \sum_{S_{1,23}} \gamma^{(k_4-k_1)}(v_1) \gamma^{(k_3-k_4)}(v_2) \gamma^{(k_2-k_3)}(v_3) \right\} dv_1 dv_2 dv_3 \\
\le \;& \frac{M_1 M_2^3}{M_3^2} \int_0^\infty \int_0^\infty \int_0^\infty |m_n W\left(m_n(u + v_1 + v_2 + v_3)\right) C(v_1 + v_2)| \\
& \qquad \times |C(v_2 + v_3)| dv_1 dv_2 dv_3 \\
\le \;& \frac{M_1 M_2^3}{M_3^2} \int_0^\infty m_n |W\left(m_n(u + v_1')\right| \\
& \qquad \times \left[\int_0^\infty |C(v_1' - v_3)| \left\{ \int_0^\infty |C(v_2 + v_3)| dv_2 \right\} dv_3 \right] dv_1' \\
\le \;& M_4 \int_0^\infty m_n |W\left(m_n(u + v_1')\right)| \left[\int_0^\infty |C(v_1' - v_3)| dv_3 \right] dv_1' \\
\le \;& M_5 \int_0^\infty m_n |W\left(m_n(u + v_1')\right)| dv_1' \;\; \le \;\; M_6,
\end{aligned}
$$

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 $S_{2,1} = \{(k_1,k_2,k_3,k_4) : k_1 < k_2 < k_3 = k_4\}$, and the transformation $\vartheta_{k_2-k_1} = t_{k_2} - t_{k_1}$ and $\vartheta_{k_3-k_2} = t_{k_3} - t_{k_2}$. Then we have

$$\begin{aligned}
&\frac{M_1}{nH^2(u)}\sum_{S_{2,1}} E\left[\,|m_n W\left(m_n(u-t_{k_1}+t_{k_2})\right) C(t_{k_1}-t_{k_3})C(t_{k_2}-t_{k_4})|\,\right] \\
&= \frac{M_1}{nH^2(u)}\sum_{S_{2,1}} E\left[|m_n W\left(m_n(u+\vartheta_{k_2-k_1})\right) C(\vartheta_{k_2-k_1}+\vartheta_{k_3-k_2})C(\vartheta_{k_3-k_2})|\right] \\
&= \frac{M_1}{H^2(u)}\int_0^\infty\int_0^\infty |m_n W\left(m_n(u+v_1)\right) C(v_1+v_2)C(v_2)| \\
&\qquad\qquad \times\left\{\frac{1}{n}\sum_{S_{2,1}}\gamma^{(k_2-k_1)}(v_1)\gamma^{(k_3-k_2)}(v_2)\right\}dv_1dv_2 \\
&\le \frac{M_1}{H^2(u)}\int_0^\infty\int_0^\infty |m_n W\left(m_n(u+v_1)\right) C(v_1+v_2)C(v_2)|\, H(v_1)H(v_2)dv_1dv_2 \\
&\le \frac{M_1M_2^2C(0)}{M_3^2}\int_0^\infty m_n|W\left(m_n(u+v_1)\right)|\left\{\int_0^\infty |C(v_2)|dv_2\right\}dv_1 \quad \le \quad M_4,
\end{aligned}$$

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 $k_1 \neq k_2$. A slightly different argument is needed in the cases of the six sets with $k_1 = k_2$. We show the calculations for $S_{2,31} = \{(k_1,k_2,k_3,k_4) : k_1 = k_2 < k_3 < k_4\}$, as a representative of these six sets. By using the transformation $\vartheta_{k_3-k_2} = t_{k_3} - t_{k_2}$ and $\vartheta_{k_4-k_3} = t_{k_4} - t_{k_3}$, we have

$$\begin{aligned}
&\frac{M_1}{nH^2(u)}\sum_{S_{2,31}} E\left[|m_n W\left(m_n(u-t_{k_1}+t_{k_2})\right) C(t_{k_1}-t_{k_3})C(t_{k_2}-t_{k_4})|\right] \\
&= \frac{M_1}{nH^2(u)}\sum_{S_{2,31}} E\left[|m_n W\left(m_n u\right) C(\vartheta_{k_3-k_2})C(\vartheta_{k_3-k_2}+\vartheta_{k_4-k_3})|\right] \\
&= \frac{M_1}{H^2(u)}\cdot m_n|W\left(m_n u\right)|\int_0^\infty\int_0^\infty |C(v_1)C(v_1+v_2)| \\
&\qquad\qquad \times\left\{\frac{1}{n}\sum_{S_{2,31}}\gamma^{(k_3-k_2)}(v_1)\gamma^{(k_4-k_3)}(v_2)\right\}dv_1dv_2 \\
&\le \frac{M_1}{H^2(u)}\cdot m_n|W\left(m_n u\right)|\int_0^\infty\int_0^\infty |C(v_1)C(v_1+v_2)|H(v_1)H(v_2)dv_1dv_2 \\
&\le \frac{M_1M_2^2}{M_3^2}\cdot m_n|W\left(m_n u\right)|\int_0^\infty |C(v_1)|\left\{\int_0^\infty |C(v_1+v_2)|dv_2\right\}dv_1 \\
&\le M_4\cdot m_n|W\left(m_n u\right)|,
\end{aligned}$$

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 $S_{3,1} = \{(k_1, k_2, k_3, k_4) : 1 \le k_1 = k_2 = k_3 < k_4 \le n\}$ and $S_{3,2} = \{(k_1, k_2, k_3, k_4) : 1 \le k_1 < k_2 = k_3 = k_4 \le n\}$, each case being a representative of the calculations needed in three other cases. We proceed with the sum over $S_{3,1}$ as follows:

$$
\begin{aligned}
&\frac{M_1}{nH^2(u)} \sum_{S_{3,1}} E\left[|m_n W\left(m_n(u - t_{k_1} + t_{k_2})\right) C(t_{k_1} - t_{k_3}) C(t_{k_2} - t_{k_4})| \right] \\
&= \frac{M_1}{H^2(u)} \int_0^\infty |m_n W\left(m_n u\right) C(0) C(v)| \left\{ \frac{1}{n} \sum_{S_{3,1}} \gamma^{(k_4 - k_3)}(v) \right\} dv \\
&\le \frac{M_1}{H^2(u)} m_n |W\left(m_n u\right)| C(0) \int_0^\infty |C(v)|\, H(v) dv \\
&\le \frac{M_1 M_2}{M_3^2} m_n |W\left(m_n u\right)| C(0) \int_0^\infty |C(v)|\, dv = M_4 \times m_n |W\left(m_n u\right)|
\end{aligned}
$$

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,

$$
\begin{aligned}
&\frac{M_1}{nH^2(u)} \sum_{S_{3,2}} E\left[|m_n W\left(m_n(u - t_{k_1} + t_{k_2})\right) C(t_{k_1} - t_{k_3}) C(t_{k_2} - t_{k_4})| \right] \\
&= \frac{M_1}{H^2(u)} \int_0^\infty |m_n W\left(m_n(u + v)\right) C(v) C(0)| \left\{ \frac{1}{n} \sum_{S_{3,2}} \gamma^{(k_2 - k_1)}(v) \right\} dv \\
&\le \frac{M_1}{H^2(u)} C^2(0) \int_0^\infty m_n \left|W\left(m_n(u + v)\right)\right| H(v) dv \\
&\le \frac{M_1 M_2}{M_3^2} C^2(0) \int_0^\infty |W(v)|\, dv \quad \le M_4
\end{aligned}
$$

for some constant M_4.

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

$$
\begin{aligned}
&\frac{M_1}{nH^2(u)} \sum_{S_4} E\left[\, |m_n W\left(m_n(u - t_{k_1} + t_{k_2})\right) C(t_{k_1} - t_{k_3}) C(t_{k_2} - t_{k_4})| \,\right] \\
&\le \frac{M_1}{nM_3^2} m_n W(m_n u) C^2(0).
\end{aligned}
$$

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

Lemma 5.2. Under the Assumptions of Theorem 5.2, the function $I_3(\cdot)$ defined in (5.24) is uniformly bounded from above by $\frac{m_n}{n}M$, where M is a constant.

Proof of Lemma 5.2. It follows from Assumption 5.6 that $I_3(u)$ is bounded as

$$\begin{aligned}
&\frac{n}{m_n}|I_3(u)| \\
&\leq \frac{M_1}{nM_3^2}\sum_{k_1=1}^{n}\sum_{k_2=1}^{n}\sum_{k_3=1}^{n}\sum_{k_4=1}^{n} E\big[m_n|W\left(m_n(u-t_{k_1}+t_{k_2})\right)|g(t_{k_2}-t_{k_4})g(t_{k_3}-t_{k_4})\big].
\end{aligned}$$

A similar argument as in Lemma 5.1 shows that $\frac{n}{m_n}|I_3(u)|$ is bounded uniformly from above. □

Proof of Theorem 5.3. From (5.16), for $f \in [-f_0, f_0]$, we have

$$\begin{aligned}
E[\widehat{\phi}_{nd}(f)] =& TE[\widehat{C}_{nd}(0)] + 2T\sum_{l=1}^{J} E[\widehat{C}_{nd}(lT)]K(b_n l)\cos(2\pi f lT) \\
&+ 2T\sum_{l=J+1}^{n-1} E[\widehat{C}_{nd}(lT)]K(b_n l)\cos(2\pi f lT).
\end{aligned}$$

Note that the second term in the above expression is a finite sum. By using Theorem 5.1 and Assumption 5.7 and 5.8, we have

$$\begin{aligned}
\lim_{n\to\infty} E[\widehat{\phi}_{nd}(f)] &= TC(0) + 2T\sum_{l=1}^{J} C(lT)\cos(2\pi f lT) \\
&\quad + \lim_{n\to\infty} 2T\sum_{l=J+1}^{n-1} E[\widehat{C}_{nd}(lT)]K(b_n l)\cos(2\pi f lT). \qquad (5.28)
\end{aligned}$$

By using the exact expectation of $E(\widehat{C}_{nd}(lT))$ for $l > J$, from the expression (5.18), we have

$$\begin{aligned}
&2T\sum_{l=J+1}^{n-1} E[\widehat{C}_{nd}(lT)]K(b_n l)\cos(2\pi f lT) \\
&= 2T\sum_{l=J+1}^{n-1} K(b_n l)\cos(2\pi f lT)\Bigg[\frac{m_n W(m_n lT)C(0)}{H(lT)} \\
&\qquad + \frac{1}{H(lT)}\int_{-\infty}^{\infty} W(v)C\left(lT-\frac{v}{m_n}\right)\left\{H_n\left(lT-\frac{v}{m_n}\right)+H_n\left(-lT+\frac{v}{m_n}\right)\right\}dv\Bigg].
\end{aligned}$$

For sufficiently large n such that $m_n > a/d$, we have

$$\begin{aligned}
&2T\sum_{l=J+1}^{n-1} E[\widehat{C}_{nd}(lT)]K(b_nl)\cos(2\pi f lT)\\
&= 2T\sum_{l=J+1}^{n-1} K(b_nl)\cos(2\pi flT)\Bigg[\frac{1}{H(lT)}\int_{-\infty}^{\infty} W(v)C\left(lT-\frac{v}{m_n}\right)\\
&\qquad\times\left\{H_n\left(lT-\frac{v}{m_n}\right)+H_n\left(-lT+\frac{v}{m_n}\right)\right\}dv\Bigg]\\
&= 2T\int_{-\infty}^{\infty} W(v)\Bigg[\sum_{l=J+1}^{n-1} K(b_nl)\cos(2\pi flT)\frac{1}{H(lT)}C\left(lT-\frac{v}{m_n}\right)\\
&\qquad\times\left\{H_n\left(lT-\frac{v}{m_n}\right)+H_n\left(-lT+\frac{v}{m_n}\right)\right\}\Bigg]dv.
\end{aligned}\tag{5.29}$$

From Assumptions 5.1, 5.2 and 5.7, the integrand of (5.29) is dominated as

$$\begin{aligned}
&|W(v)|\sum_{l=J+1}^{n-1}\Bigg|K(b_nl)\cos(2\pi flT)\frac{1}{H(lT)}C\left(lT-\frac{v}{m_n}\right)\\
&\qquad\times\left\{H_n\left(lT-\frac{v}{m_n}\right)+H_n\left(-lT+\frac{v}{m_n}\right)\right\}\Bigg|\\
&\leq |W(v)|\sup|K(\cdot)|\frac{2M_2}{M_3}\sum_{l=J+1}^{n-1} h_0\left(lT-\frac{v}{m_n}\right)\\
&\leq |W(v)|\sup K(\cdot)\frac{2M_2}{M_3}\times 2\sum_{l=0}^{\infty} h_0\left(lT\right).
\end{aligned}$$

Integrability of the bounding function is ensured from Assumption 5.2. From Assumptions 5.4 and 5.8, we have pointwise convergence as

$$\begin{aligned}
&\lim_{n\to\infty} W(v)\sum_{l=J+1}^{n-1} K(b_nl)\cos(2\pi flT)\frac{1}{H(lT)}C\left(lT-\frac{v}{m_n}\right)\\
&\qquad\times\left\{H_n\left(lT-\frac{v}{m_n}\right)+H_n\left(-lT+\frac{v}{m_n}\right)\right\}\\
&= W(v)\sum_{l=J+1}^{\infty}\cos(2\pi flT)C(lT).
\end{aligned}$$

By applying the DCT, we have

$$\lim_{n\to\infty} 2T\sum_{l=J+1}^{n-1} E[\widehat{C}_{nd}(lT)]K(b_nl)\cos(2\pi flT) = 2T\sum_{l=J+1}^{\infty} C(lT)\cos(2\pi flT).$$

From (5.28), we have

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

This completes the proof. □

Proof of Theorem 5.4. From (5.16), we have

$$\begin{aligned} Var[\widehat{\phi}_{nd}(f)] \quad &\leq \quad 3\times T^2 Var[\widehat{C}_{nd}(0)] + 3\times 4T^2 Var\left[\sum_{l=1}^{J}\widehat{C}_{nd}(lT)K(b_n l)\cos(2\pi f lT)\right] \\ &\quad + 3\times 4T^2 Var\left[\sum_{l=J+1}^{n-1}\widehat{C}_{nd}(lT)K(b_n l)\cos(2\pi f lT)\right]. \end{aligned}$$

From Theorem 5.2, the first term on the right hand side goes to zero as $n\to\infty$. The second term is a constant multiple of

$$\begin{aligned} &Var\left[\sum_{l=1}^{J}\widehat{C}_{nd}(lT)K(b_n l)\cos(2\pi f lT)\right] \\ &= \sum_{l=1}^{J}\sum_{l'=1}^{J} Cov(\widehat{C}_{nd}(lT),\widehat{C}_{nd}(l'T))K(b_n l)\cos(2\pi f lT)K(b_n l')\cos(2\pi f l'T). \end{aligned}$$

Observe that the above sum is over a finite number of terms. From Theorem 5.2, we have

$$\lim_{n\to\infty} Var\left[\sum_{l=1}^{J}\widehat{C}_{nd}(lT)K(b_n l)\cos(2\pi f lT)\right] = 0.$$

Now, consider the third term. We have from Theorem 5.2

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

where M' is a constant. Note that from Assumption 5.7, we have

$$\lim_{n\to\infty} \sum_{l=J+1}^{n-1} |K(b_n l)| b_n \leq \int_0^\infty |K(v)| dv.$$

Thus we have, from Assumption 5.4 and 5.8A,

$$\lim_{n\to\infty} Var\left[\sum_{l=J+1}^{n-1} \widehat{C}_{nd}(lT) K(b_n l) \cos(2\pi f l T)\right]$$
$$\leq \lim_{n\to\infty} M' \frac{m_n}{b_n} \frac{1}{n b_n} \int_0^\infty |K(v)| dv = 0.$$

This completes the proof. □

5.5 Simulation study

We consider the continuous time stationary stochastic process X limited to the frequency band $[-f_0, f_0]$, sampled as per the additive random sampling scheme mentioned in Section 4.5, where the inter-sample spacing has the exponential distribution left-truncated at d and the mean spacing is $d + \theta$. 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 estimator $\widehat{\phi}_{nd}(\cdot)$, and compare this performance with that of the Poisson sampling based estimator $\widehat{\psi}_n(\cdot)$ given in (1.14), regardless of the fact that Poisson sampling has not been used. We also generate uniformly spaced samples of the same continuous time process at the Nyquist rate when this is feasible under the constraint, and if so, compare the performance of the above estimators with the uniform sampling based estimator $\widehat{\phi}_\rho(\cdot)$ given in (1.12).

We continue to choose the following values of the parameters:

$$\begin{aligned} f_0 &= 1, \\ \Xi(z) &= (1 + 1.2z)^8, \\ \sigma &= 1/20^2, \\ \theta &= 1. \end{aligned}$$

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

$$W(x) = \begin{cases} \frac{1}{2}\{1 + \cos(\pi x)\} & \text{if } -1 \leq x \leq 1, \\ 0 & \text{otherwise,} \end{cases}$$

and

$$K(x) = \begin{cases} \frac{1}{2}\{1 + \cos(\pi x)\} & \text{if } -1 \leq x \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$

We first assume that the minimum separation between successive samples is $d = 0.25$. For this value of d, uniform sampling at the Nyquist rate (which happens to be equal to 2) is feasible. We run simulations for sample sizes $n = 100$ and $n = 1000$. For the estimators $\widehat{\phi}_{nd}(\cdot)$, $\widehat{\phi}_\rho(\cdot)$ and $\widehat{\psi}_n(\cdot)$ and both sample sizes, we use $b_n = 0.1$. We use $\rho = 2$ for $\widehat{\phi}_\rho(\cdot)$. Also, for $\widehat{\phi}_{nd}(\cdot)$ and both sample sizes, we use $L_n = 20$, $u_1 = d + 0.25$ and $u_2 = 6$. Finally, we use $m_n = 2$ for $n = 100$ and $m_n = 2.5$ for $n = 1000$.

Figure 5.1 shows the average of the estimates of $\widehat{\phi}_{nd}(f)$, $\widehat{\phi}_\rho(f)$ and $\widehat{\psi}_n(f)$ computed from 500 Monte Carlo simulation runs, along with the true density, for $f \in [0, f_0]$. In order to highlight the inconsistency of $\widehat{\psi}_n(f)$ in the present situation, we also include in this figure the plot of the limiting expected value of this estimator, which turns out to be

$$2\int_d^\infty C(u)\cos(2\pi f u)du.$$

The plot indicates that both $\widehat{\phi}_{nd}(f)$ and $\widehat{\phi}_\rho(f)$ converges to the true power spectral density while $\widehat{\psi}_n(f)$ converges to the wrong function, which is not even positive over its entire range. Figure 5.2 indicates that the MSE of the estimator $\widehat{\phi}_{nd}(f)$ is much smaller than that of $\widehat{\psi}_n(f)$, but larger than that of $\widehat{\phi}_\rho(f)$.

We now assume that the minimum inter-sample spacing is $d = 0.75$. For this value of d, uniform sampling of the process at the Nyquist rate is not feasible, and therefore we confine our comparison to $\widehat{\phi}_{nd}(\cdot)$ and $\widehat{\psi}_n(\cdot)$. We run the simulations using the parameters as chosen in the case of $d = 0.25$. Figure 5.3 shows the average of the estimates of $\widehat{\phi}_{nd}(f)$ and $\widehat{\psi}_n(f)$ computed from 500 Monte Carlo simulation runs, along with the true density and the limiting expected value of the estimator $\widehat{\psi}_n(f)$, for $f \in [0, f_0]$. Again the plots suggest that the $\widehat{\psi}_n(f)$ converges to the wrong function. The empirical bias of the

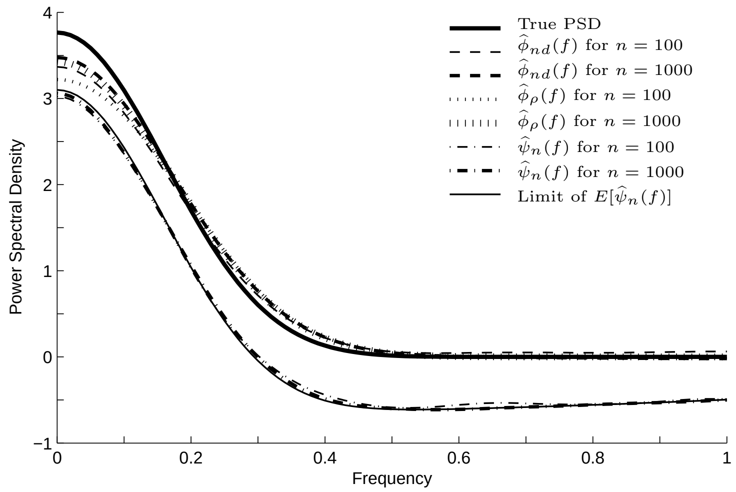

Figure 5.1: Average of the estimates of $\widehat{\phi}_{nd}(f)$, $\widehat{\phi}_{\rho}(f)$ and $\widehat{\psi}_n(f)$ for minimum separation $d = 0.25$ based on sample sizes 100 and 1000 computed from 500 Monte Carlo simulation runs, along with True PSD and $\lim_{n\to\infty} E[\widehat{\psi}_n(f)]$.

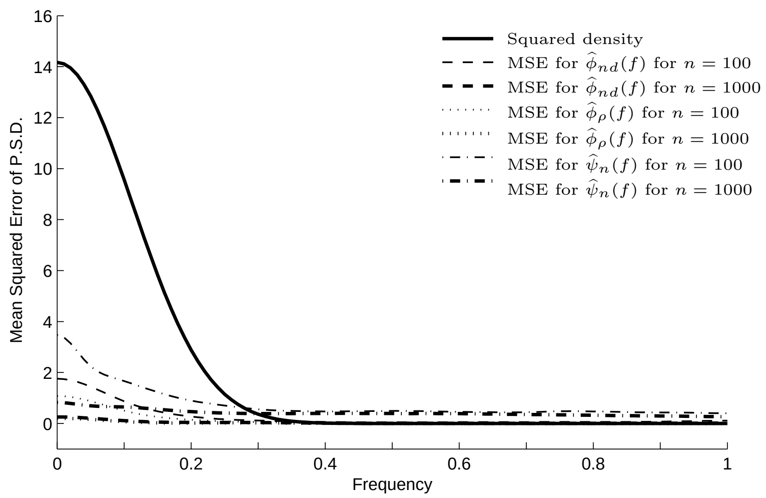

Figure 5.2: Empirical MSE of $\widehat{\phi}_{nd}(f)$, $\widehat{\phi}_{\rho}(f)$ and $\widehat{\psi}_n(f)$ for minimum separation $d = 0.25$ based on sample sizes 100 and 1000 computed from 500 Monte Carlo simulation runs along with squared True PSD.

estimator $\widehat{\phi}_{nd}(f)$ is smaller for larger n. Figure 5.4, which shows the empirical MSE of the two estimators along with the squared power spectral density, indicates that the MSE of $\widehat{\phi}_{nd}(f)$ is smaller, and it reduces with sample size, while that of $\widehat{\psi}_n(f)$ 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

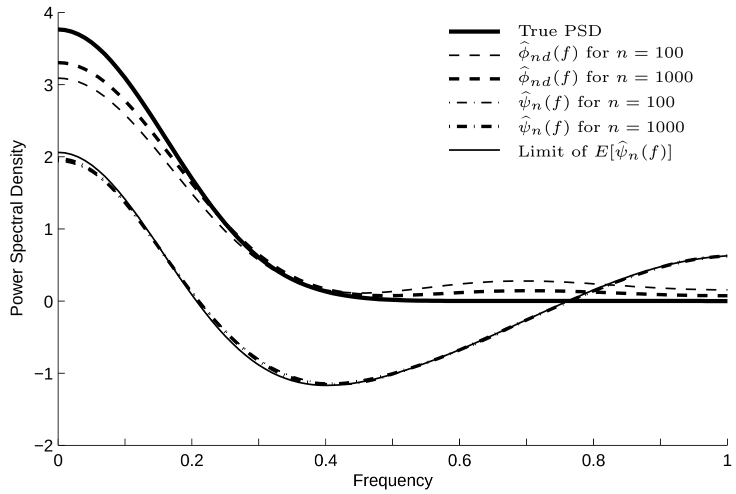

Figure 5.3: Average of the estimates of $\widehat{\phi}_{nd}(f)$ and $\widehat{\psi}_n(f)$ for minimum separation $d = 0.75$ based on sample sizes 100 and 1000 computed from 500 Monte Carlo simulation runs, along with True PSD and $\lim_{n\to\infty} E[\widehat{\psi}_n(f)]$.

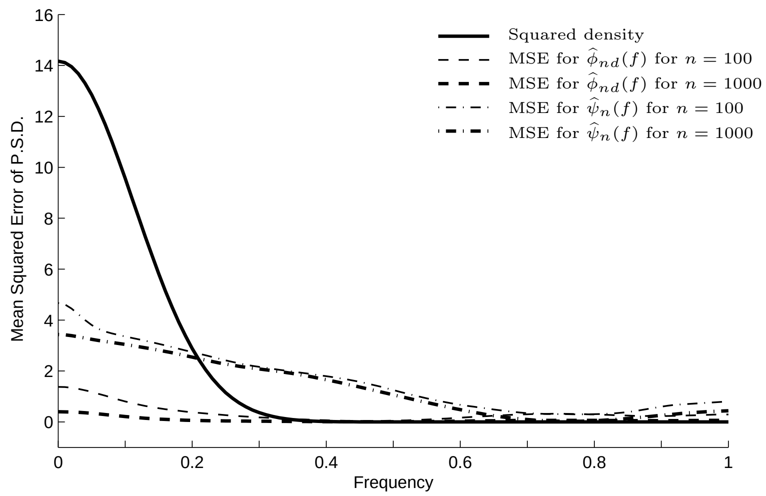

Figure 5.4: Empirical MSE of $\widehat{\phi}_{nd}(f)$ and $\widehat{\psi}_n(f)$ for minimum separation $d = 0.75$ based on sample sizes 100 and 1000 computed from 500 Monte Carlo simulation runs along with squared True PSD.

with the parameters chosen as in the previous case. Figures 5.5 and 5.6 depict the empirical average and the MSE, respectively, for the two estimators, computed from 500 simulation runs. Figure 5.5 indicates convergence of $\widehat{\psi}_n(f)$ to a wrong function, which is different from the asymptotic mean curves shown in Figures 5.1 and 5.3. For $n = 100$, the estimator $\widehat{\phi}_{nd}(f)$ has larger MSE than that of $\widehat{\psi}_n(f)$ for some values of

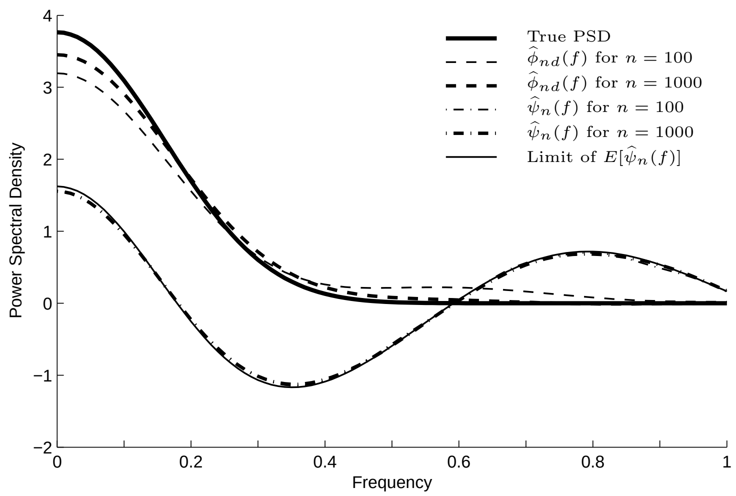

Figure 5.5: Average of the estimates of $\widehat{\phi}_{nd}(f)$ and $\widehat{\psi}_n(f)$ for minimum separation $d = 1$ based on sample sizes 100 and 1000 computed from 500 Monte Carlo simulation runs, along with True PSD and $\lim_{n\to\infty} E[\widehat{\psi}_n(f)]$.

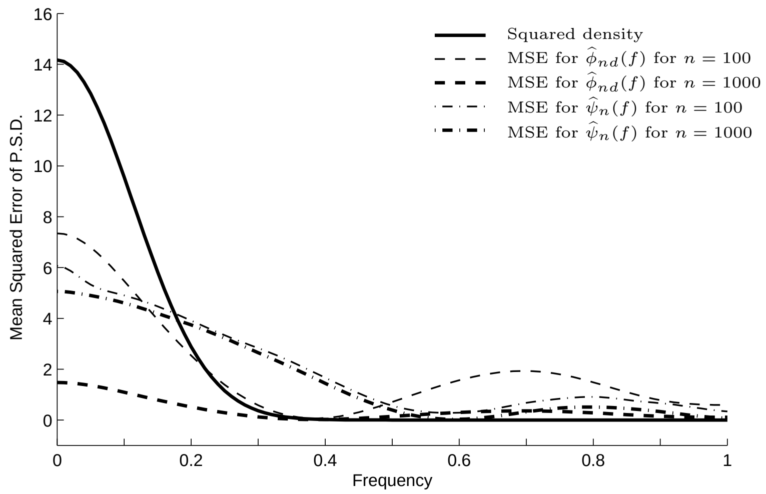

Figure 5.6: Empirical MSE of $\widehat{\phi}_{nd}(f)$ and $\widehat{\psi}_n(f)$ for minimum separation $d = 1$ based on sample sizes 100 and 1000 computed from 500 Monte Carlo simulation runs along with squared True PSD.

f. However, for $n = 1000$, the MSE of $\widehat{\phi}_{nd}(f)$ reduces sharply. On the other hand, the MSE of $\widehat{\psi}_n(f)$ saturates to a certain positive value, determined by the asymptotic bias.

The MSE of $\widehat{\phi}_{nd}(f)$ is seen to be larger for higher values of d. This finding can be explained by the fact that in the case $d = 1$, the estimator $\widehat{\phi}_{nd}(\cdot)$ 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 $d = 0.75$ and no indirect estimation at all in the case of $d = 0.25$. Diminishing 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 $\widehat{\phi}_{nd}(f)$ for $n = 1000$. This pattern is observed to continue for $d = 2$, but even higher sample sizes are needed for satisfactory performance. These graphs are not shown here.

5.6 Summary and discussion

This chapter 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 theoretical studies and Monte Carlo simulations reported in this chapter demonstrate that it is possible to judiciously use additive random sampling to surpass the Nyquist limit with the help of large sample size. It also illustrates an exclusive advantage of stochastic sampling over uniform sampling.

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 (5.7). It appears that 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 5.5 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 poor.

The convergence rate of MSE of an estimator is often used as a measure of performance of the estimator. In a subsequent work, the optimal rate of convergence of the MSE of the covariance estimator (5.15) and of the spectrum estimator (5.16) have been established. The optimal rates of convergence for both the estimators turn out to be a polynomial rate (see Srivastava and Sengupta (2011b) for details).

Chapter 6

Uniform vs. Stochastic Sampling: Present Status and Future Work

6.1 Introduction

In this thesis, we have looked at the problem of estimation of the power spectral density of a continuous time, mean square continuous, stationary stochastic process based on finitely many samples of the process. When one has control over the sampling mechanism, selection of the sampling scheme becomes an important issue. The two most common forms of sampling are uniform and stochastic sampling. It is well known that uniform sampling is relatively easy to implement, and the data produced from such sampling are amenable to analysis through a rich collections of methods. On the other hand, stochastic sampling is sometimes advocated by citing certain weaknesses of uniform sampling. The findings of Chapters 2–5 of this thesis should be useful in comparing the two sampling schemes from various angles, and making an informed choice. In this chapter, we attempt to summarize all the information relevant for this comparison, including results that have been obtained in this thesis and those which were already known. We also identify the gaps in the existing body of knowledge, which can be filled through future research.

Section 6.2 contains a summary of the information available on this matter till date, and Section 6.3 outlines the areas of possible further work.

6.2 Uniform vs. stochastic sampling: Present status

6.2.1 Estimation of possibly non-bandlimited spectral density

It has been known for over three decades that a possibly non-bandlimited power spectral density can be consistently estimated through stochastic sampling. Masry (1978a) gave a recipe for constructing a consistent estimator, and Masry (1978b) made the detailed asymptotic calculations for the estimator (1.14) based on Poisson sampling. On the other hand, it had been wrongly argued that consistent estimation through uniform sampling is not possible (Shapiro and Silverman, 1960; Beutler, 1970; Masry, 1978a). The 'shrinking asymptotic' arguments given in Chapter 2 prove that the smoothed periodogram estimator (2.1) is actually consistent for a possibly non-bandlimited power spectral density, under conditions that are similar to those used in proving the consistency of the estimator (1.14). Moreover, the two estimators have about the same optimal rate of convergence. The estimator (2.1) has larger bias, but smaller variance, in comparison with the estimator (1.14). Both the estimators are asymptotically normal, subject to appropriate scaling, and can be used to construct asymptotic confidence intervals that shrink to the true power spectral density (see Chapter 3 and Lii and Masry (1994)).

6.2.2 Spectrum estimation with restriction on average sampling rate

A restriction on the sampling rate makes the asymptotic arguments of Chapters 2 and 3 inapplicable. Consistent estimation through uniform sampling is possible only if the underlying power spectral density is known to have bandwidth smaller than half of the largest permissible sampling rate. However, no such condition is necessary for stochastic sampling. In particular, the estimator (1.14) based on Poisson sampling is consistent and asymptotically normal, subject to appropriate scaling.

If the underlying power spectral density is suitably bandlimited so as to permit consistent estimation through uniform sampling, then the smoothed periodogram estimator (1.12) is consistent and asymptotically normal, subject to appropriate scaling. It has the same rate of convergence as (1.14). First order calculations produce identical constants for the bias of the two estimators, but the constant for the variance is smaller in the case of (1.12) (Parzen, 1957; Masry, 1978b).

6.2.3 Spectrum estimation with restriction on minimum inter-sample spacing

If there is a restriction on the minimum spacing between successive samples, and the underlying power spectral density is possibly non-bandlimited, then neither uniform sampling nor point process sampling can lead to a consistent estimator. The limitation of uniform sampling in this regard is well known (Kay, 1999), while the limitation of point process sampling has been proved in Section 4.3.

Now suppose that the restriction on inter-sample spacing continues to be there, but the underlying power spectral density is bandlimited. Consistent estimation through uniform sampling is possible only if the maximum permissible rate of uniform sampling happens to be the same as the Nyquist rate or higher. In particular, the smoothed periodogram estimator (1.12) based on uniform sampling is consistent in such a case. On the other hand, irrespective of the magnitude of the minimum permissible inter-sample spacing, consistent estimation through additive random sampling is possible. Even though the estimator (1.14) based on Poisson sampling is not feasible in this case, the estimator (5.16) developed in Chapter 5 based on *additive random sampling satisfying Assumption 5.2* is consistent, under appropriate conditions. An empirical study (see Section 5.6) indicates that when the permissible inter-sample spacing is smaller than the reciprocal of the Nyquist rate (i.e., the estimator (1.12) is consistent), the MSE of the estimator (5.16) may be larger than that of the estimator (1.12).

6.2.4 Discrete power spectral distribution: Estimation of sinusoids

We have considered in this thesis power spectral distributions that have a density. Estmation of power spectral distributions consisting of mass at only a countable number of points is also an important problem. In this case, the underlying process is a linear combination of sinusoids. The locations of the point masses on the frequency axis (i.e., the frequencies of the sinusoids) as well as the magnitudes of those masses are to be estimated.

Many methods based on uniform sampling have been developed for this problem.

Usually a finite (and specified) number of sinusoids is assumed. These estimators have the limitation that if there is a restriction on the maximum sampling rate, but the sinusoids of the underlying process are not necessarily restricted to have frequencies smaller than half of this maximum rate, then the estimators cannot be consistent.

Stochastic sampling can be useful in this situation. Isokawa (1983) considered the case of a single sinusoid, proposed a general estimator based on stationary orderly point process sampling. For this estimator, the frequency of the sinusoid is estimated by locating the global maximum of a function that can be called an extension of the periodogram for non-uniformly sampled data. The amplitude parameters are subsequently estimated through a least square fit. This estimator, which can be generalized to the case of multiple sinusoids in the same manner as its uniform sampling counterparts (Walker, 1971; Hannan, 1973) are generalized, is consistent under appropriate conditions, as long as the underlying sinusoid has frequency restricted to a maximum (but possibly unknown) limit. It is also asymptotically normal, subject to appropriate scaling. Restrictions on average sampling rate or minimum inter-sample spacing do not invalidate the asymptotic results.

6.3 Uniform vs. stochastic sampling: Future work

The asymptotic arguments involving a sampling rate that goes to infinity in tune with the sample size ('shrinking' asymptotics) can be useful for a number of problems related to those considered in Chapters 2 and 3. Some potential areas of application are indicated below.

A. The results of Chapter 3, can be built upon in order to obtain asymptotic confidence bands that shrink to the true power spectral density, which may possibly be non-bandlimited.

B. One can try and establish consistency of *parametric* estimators of the power spectral density of a continuous time process based on uniformly spaced samples. Such asymptotic calculations may potentially be used to justify and/or fine-tune multi-resolution methods of spectrum estimation (Eldar et al., 1997).

C. For the problem of estimation of a power spectral distribution function arising from a linear combination of sinusoids, the possibility of an underlying sinusoid having arbitrarily large frequency is often avoided for technical reasons. The above arguments can be used to remove this difficulty for estimators based on uniformly sampled data.

Our work on stochastic sampling also gives rise to new research questions.

In Chapter 5, we have established the consistency of the estimator (5.16) based on renewal process sampling under the constraint of a minimum spacing between successive samples. One can look into the asymptotic distribution of this estimator under suitable scaling. The simulation results reported in Section 5.5 indicate that the estimator may have larger MSE than the smoothed periodogram estimator based on uniform sampling, when the latter is feasible under the specified minimum inter-sample spacing. Of course, the estimator (5.16) is meant to be used only where the bandwidth of the underlying process is larger than half of the maximum rate of uniform sampling, i.e., when there is no competing estimator. It is not clear whether the large MSE of the estimator (5.16), indicated by the simulation results, is due the inherent difficulty of the estimation problem under the given constraint, or it is due to a poor strategy for estimation. This question can be looked into.

Appendix

The simulations reported in Sections 2.5, 3.5, 4.5 and 5.5 involve generating samples from a continuous time AR process and a continuous time bandlimited process. Here, we give a detailed description of the methods used to generate samples from such processes.

A.1 Generating samples from a continuous time AR process

The spectral density (2.36) is that of an AR(4) process. A general AR(P) process $X = \{X(t), \ -\infty < t < \infty\}$ having the spectral density

$$\phi(f) = \sigma^2 \prod_{j=1}^{P} \frac{1}{(2\pi f)^2 + \alpha_j^2}$$

can be represented in terms of stochastic integral as (Hoel et al., 1972)

$$X(t) = \int_{-\infty}^{t} h(t-s) dB(s),$$

where $\{B(t), \ -\infty < t < \infty\}$ is the Brownian motion with variance σ^2 and $h(\cdot)$ is the 'impulse response' function given by

$$h(t) = \sum_{j=1}^{P} c_j e^{-\alpha_j t} 1_{[0,\infty)}(t).$$

Here, the constants c_j, $j = 1, \ldots, P$, are solutions to the system of linear equations:

$$h(0) = h^{(1)}(0) = h^{(2)}(0) = \cdots = h^{(P-2)}(0) = 0; \quad h^{(P-1)}(0) = 1,$$

$h^j(0)$ being the jth derivative of $h(\cdot)$ evaluated at 0.

In view of the above representation, we aim at simulating the process $X_0 = \{X_0(t),\ 0 < t < \infty\}$ given by

$$X_0(t) = \int_0^t h(t-s)dB(s).$$

The process X_0 is not stationary. However, as t becomes large, the variance of the difference between the processes X and X_0 becomes small. We find out the value of t such that $Var(X_0(t) - X(t))$ is less than a threshold (we choose the threshold as 10^{-9}), and consider the path of the simulated process X_0 for larger values of t.

Given the time point $t_0 = 0 < t_1 < t_2 < \ldots < t_n$, the sampled process $X_0(t_k)$ for $k = 1, 2, \ldots, n$ can be expressed as

$$\begin{aligned} X_0(t_{k+1}) &= \sum_{j=1}^{P} c_j \Big\{ e^{-\alpha_j(t_{k+1}-t_k)} \int_0^{t_k} e^{-\alpha_j(t_k-s)} dB(s) + \int_{t_k}^{t_{k+1}} e^{-\alpha_j(t_{k+1}-s)} dB(s) \Big\} \\ &= \sum_{j=1}^{P} c_j \Bigg[e^{-\alpha_j(t_{k+1}-t_k)} \left\{ \sum_{l=1}^{k} e^{-\alpha_j(t_k-t_l)} \int_{t_{l-1}}^{t_l} e^{-\alpha_j(t_l-s)} dB(s) \right\} \\ &\qquad + \int_{t_k}^{t_{k+1}} e^{-\alpha_j(t_{k+1}-s)} dB(s) \Bigg] \\ &= \sum_{j=1}^{P} c_j \left\{ \sum_{l=1}^{k} e^{-\alpha_j(t_k-t_l)} T_{l,j} + T_{k+1,j} \right\}, \end{aligned} \tag{A.1}$$

where $T_{k,j} = \int_{t_{k-1}}^{t_k} e^{-\alpha_j(t_k-s)} dB(s)$ for $j = 1, \ldots, P$ and $k = 1, \ldots, n$. Note that the random vectors $T_k = (T_{k,1}, T_{k,2}, \ldots, T_{k,P})$ for $k = 1, \ldots, n$ are independent normal. The P-dimensional random vectors T_k have mean 0 and the covariance between $T_{k,j}$ and $T_{k,j'}$, for $j, j' = 1, \ldots, P$, is

$$\begin{aligned} &\sigma^2 \int_{t_{k-1}}^{t_k} \int_{t_{k-1}}^{t_k} e^{-\alpha_j(t_k-s_1)} e^{-\alpha_{j'}(t_k-s_2)} ds_1 ds_2 \\ &= \frac{\sigma^2}{\alpha_j \alpha_{j'}} \left(1 - e^{-\alpha_j(t_k-t_{k-1})}\right) \left(1 - e^{-\alpha_{j'}(t_k-t_{k-1})}\right). \end{aligned}$$

Once n samples of T_k are generated, $X_0(t_k)$ for $k = 1, \ldots, n$ can be generated from (A.1).

The bivariate process described in (3.28) consists of components that are sums of AR(1) processes, which can be generated in a similar manner.

A.2 Generating samples from a continuous time bandlimited process

The process described by the covariance function (4.22) is bandlimited. From the spectral representation of a general process X having bandwidth f_0 and spectral density $\phi(\cdot)$, we have

$$X(t) = \int_{-f_0}^{f_0} \cos(2\pi ft)dY_1(f) + \int_{-f_0}^{f_0} \sin(2\pi ft)dY_2(f), \tag{A.2}$$

where $Y_j = \{Y_j(f),\ -f_0 < f < f_0\}$ for $j = 1, 2$ are mutually uncorrelated, orthogonal increment stochastic processes with mean 0 and variance given as $Var(dY_j(f)) = \phi(f)df$ (Priestley, 1983).

Consider the process $X_0 = \{X_0(t),\ 0 < t < \infty\}$ defined as

$$X_0(t) = \sum_{k=1}^{m} \cos(2\pi s_k t)\sqrt{(s_k - s_{k-1})}Y_{01}(s_k) + \sum_{k=1}^{m} \sin(2\pi s_k t)\sqrt{(s_k - s_{k-1})}Y_{02}(s_k), \tag{A.3}$$

where the frequencies $s_0, s_1, \ldots, s_m$ form a grid satisfying the relations $s_0 = -f_0 < s_1 < s_2 < \ldots < s_m = f_0$, and $Y_{01}(s_k)$ and $Y_{02}(s_k)$ for $k = 1, 2 \ldots, m$ are independent random variables with mean 0 and variance $\phi(s_k)$. Note that X_0 is a stationary stochastic process with mean 0 and

$$\begin{aligned} Cov(X_0(u), X_0(v)) = &\sum_{k=1}^{m} \cos(2\pi s_k u)\cos(2\pi s_k v)(s_k - s_{k-1})Var(Y_{01}(s_k)) \\ &+ \sum_{k=1}^{m} \sin(2\pi s_k u)\sin(2\pi s_k v)(s_k - s_{k-1})Var(Y_{02}(s_k)) \\ = &\sum_{k=1}^{m} \cos(2\pi s_k(u - v))(s_k - s_{k-1})\phi(s_k). \end{aligned}$$

Further,

$$\lim_{\max(s_k - s_{k-1}) \to 0} Cov(X_0(u), X_0(v)) = \int_{-f_0}^{f_0} \cos(2\pi f(u - v))\phi(f)df = Cov(X(u), X(v)).$$

Thus, by choosing the grid to be sufficiently fine, the difference between the covariance functions of X and X_0 can be made arbitrary small.

We assume that $Y_{01}(s_k)$ and $Y_{02}(s_k)$ for $k = 1, 2 \ldots, m$ are normal, and use (A.3) to generate the samples of the process X_0 (in lieu of X) at any time point t.

Bibliography

Benedetto, J. J. and Ferreira, P. J. S. G. (2001). *Modern Sampling Theory*. Birkhäuser, Boston.

Beutler, F. J. (1970). Alias-free randomly timed sampling of stochastic processes. *IEEE Trans. Inf. Theory*, IT-16(2):147–152.

Beutler, F. J. and Leneman, O. A. Z. (1966). The theory of stationary point processes. *Acta Math.*, 116(1):159–197.

Brillinger, D. R. (1972). The spectral analysis of stationary interval functions. *Proc. Sixth Berkeley Symp. Prob. Statist.*, 1:483–513.

Brillinger, D. R. (2001). *Time Series Data Analysis and Theory*. SIAM, Philadelphia.

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

Bronez, T. P. (1988). Spectral estimation of irregularly sampled multidimensional processes by generalized prolate spheroidal sequences. *IEEE Trans. Acoustics Speech Signal Processing*, 36(12):1862–1873.

Chen, H., Simpson, D. G., and Ying, Z. (2000). Infill asymptotics for a stochastic process model with measurement error. *Statist. Sinica*, 10:141–156.

Constantine, A. G. and Hall, P. (1994). Characterizing surface smoothness via estimation of effective fractal dimension. *J. Roy. Statist. Soc. Ser. B*, 56(1):97–113.

Costain, J. K. and Çoruh, C. (2004). *Basic Theory of Exploration Seismology*. Elsevier, Amsterdam.

Daley, D. J. and Vere-Jones, D. (2002). *An Introduction to the Theory of Point Proceses.* Springer, New York.

Davis, P. J. (1975). *Interpolation and Approximation.* Dover, New York.

Eldar, Y., Lindenbaum, M., Porat, M., and Zeevi, Y. Y. (1997). The farthest point strategy for progressive image sampling. *IEEE Trans. Image Processing*, 6(9):1305–1315.

French, A. S. and Holden, A. V. (1971). Alias-free sampling of neuronal spike trains. *Biological Cybernetics*, 8(5):165–171.

Fuentes, M. (2002). Spectral methods for nonstationary spatial processes. *Biometrika*, 89(1):197–210.

Hall, P., Fisher, N. I., and Hoffmann, B. (1994). On the nonparametric estimation of covariance functions. *Ann. Statist.*, 22(4):2115–2134.

Hannan, E. J. (1973). The estimation of frequency. *J. Appl. Prob.*, 10(3):510–519.

Hasselmann, K., Munk, W., and MacDonald, G. (1963). *Bispectra of Ocean Waves – Proc. Symposium on Time Series Analysis.* John Wiley, New York.

Higgins, J. R. (1996). *Sampling Theory in Fourier and Signal Analysis.* Oxford University Press, Oxford.

Hoel, P. G., Port, S. C., and Stone, C. J. (1972). *Introduction to Stochastic Processes.* Houghton Mifflin, Boston.

Hung, W. (2002). *International Handbook of Earthquake and Engineering Seismology.* Academic Press, New York.

Ishimaru, A. and Chen, Y. (1965). Thinning and broadbanding antenna arrays by unequal spacings. *IEEE Trans. Antenna Propagation*, AP-13(1):208–215.

Isokawa, Y. (1983). Estimation of frequency by random sampling. *Ann. Ins. Statist. Math.*, 35(1):201–213.

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

Lahiri, S. N. (1999). Asymptotic distribution of the empirical spatial cumulative distribution function predictor and prediction bands based on a subsampling method. *Probab. Theory Related Fields*, 114(1):55–84.

Landau, H. J. (1967). Sampling, data transmission, and the nyquist rate. *Proc. IEEE*, 55(10):1701–1706.

Lavrentieff, M. (1936). *Sur les Fonctions d'une Variable Complexe Représentables par des Séries de Polynômes.* Hermann and Cie, Paris.

Lehr, M. and Lii, S. K. (1997). Wavelet spectral density estimation under irregular sampling. *Conference Record of the Thirty-First Asilomar Conference on Signals, Systems and Computers*, 2:1117–1121.

Leonov, V. and Shiryayev, A. N. (1959). On a method of calculation of semi-invariants. *Theory Probab. Appl.*, 4(3):319–329.

Lii, K. S. and Masry, E. (1994). Spectral estimation of continuous-time stationary processes from random sampling. *Stochastic Process. Appl.*, 52(1):39–64.

Lim, C. Y. and Stein, M. (2008). Properties of spatial cross-periodograms using fixed-domain asymptotics. *J. Multivariate Anal.*, 99(9):1962–1984.

MacDonald, G. J. F. and Ness, N. (1961). A study of the free oscillations of the earth. *J. Geophys Res.*, 66(8):1865–1911.

Marvasti, F. A. (2001). *Nonuniform Sampling.* Kluwer Plenum, New York.

Masry, E. (1978a). Alias-free sampling: An alternative conceptualization and its applications. *IEEE Trans. Inf. Theory*, IT-24(3):173–183.

Masry, E. (1978b). Poisson sampling and spectral estimation of continuous-time processes. *IEEE Trans. Inf. Theory*, IT-24(2):173–183.

Matson, P., Johnson, L., Billow, C., Miller, J., and Pu, R. (1994). Seasonal patterns and remote spectral estimation of canopy chemistry across the oregon transect. *Ecological Applications*, 4(2):280–298.

Mergelyan, S. N. (1954). Uniform approximations of functions of a complex variable. *Amer. Math. Soc. Trans. Ser. 1 Series Approximation*, 3:295–291.

Mitchel, D. P. (1987). Generating antialiased images at low sampling density. *Comp. Graphics*, 21(4):65–72.

Moore, M. I., Visser, A. W., and Shirtcliffe, T. L. G. (2008). Experiances with the brillinger spectral estimator applied to simulated irregularly observed process. *J. Time Series Anal.*, 8(4):433–442.

Munson, D., OBrien, J., and Jenkins, W. (1983). A tomographic formulation of spotlight-mode synthetic aperture radar. *Proc. IEEE*, 71(8):917–925.

Naidu, P. S. (1999). *Modern Spectrum Analysis of Time Series*. CRC-Press, Florida.

Nobach, H., Mller, E., and Tropea, C. (1998). Efficient estimation of power spectral density from laser doppler anemometer data. *Experiments in Fluids*, 24(5–6):499–509.

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

Ozbek, A. and Ferber, R. (2005). Multidimensional filtering of irregularly sampled seismic data. *Proceedings of the 13th European Signal Processing Conference*.

Parzen, E. (1957). On consistent estimates of the spectrum of a stationary time series. *Ann. Math. Statist.*, 28(2):329–348.

Pillai, S. U. and Shim, T. I. (1993). *Spectrum Estimation and System Identification*. Springer, New York.

Politis, D. N., Romano, J. P., and Lai, T. L. (1992). Bootstrap confidence bands for spectra and cross-spectra. *IEEE Trans. Signal Processing*, 40(5):1206–1215.

Priestley, M. B. (1983). *Spectral Analysis and Time Series: Vol. II.* Academic Press, New York.

Ramsay, J. O., Hooker, G., and Graves, S. (2009). *Functional Data Analysis with R and Matlab.* Spriger, New York.

Resnick, S. I. (1987). *Extreme Values Regular Variation and Point Processes.* Springer-Verlag, New York.

Roberts, J. B. and Gaster, M. (1980). On estimation of the spectra from randomly sampled signals: a method of reducing variability. *Proc. Roy. Soc. Ser. A*, 371(1745):235–258.

Roughan, M. (2006). A comparision of poisson and uniform sampling for active measurements. *IEEE Trans. Selected Areas Communication*, 24(12):2299–2312.

Rudin, W. (1987). *Real and Complex Analysis.* McGraw-Hill, New York.

Shannon, C. E. (1949). Comunication in presence of noise. *Proc. IRE*, 37:10–21.

Shapiro, H. S. and Silverman, R. A. (1960). Alias-free sampling of random noise. *J. Soc. Indust. Appl. Math*, 8(2):225–248.

Shorak, G. R. and Wellner, J. A. (1986). *Empirical Processes with Applications to Statistics.* John Wiley, New York.

Srivastava, R. and Sengupta, D. (2010). Consistent estimation of non-bandlimited spectral density from uniformly spaced samples. *IEEE Trans. Inf. Theor.*, 56(8):3642–3659.

Srivastava, R. and Sengupta, D. (2011a). Asymptotic confidence intervals of spectra and cross-spectra of continuous time processes based on uniformly spaced samples. *Technical report*, ASD/2011/2.

Srivastava, R. and Sengupta, D. (2011b). Bandlimited spectrum estimation under the constraint of a minimum inter-sample spacing. *Technical report*, ASD/2011/1.

Srivastava, R. and Sengupta, D. (2011c). Effect of inter-sample spacing constraint on spectrum estimation with irregular sampling. *IEEE Trans. Inf. Theor.*, to appear.

Stein, M. L. (1995). Fixed-domain asymptotics for spatial periodograms. *J. Amer. Statist. Assoc.*, 90(432):1277–1288.

Stoica, P., Li, P. J., and He, H. (2009). Spectral analysis of nonuniformly sampled data: A new approach versus the periodogram. *IEEE Trans. Signal Processing*, IT-24(2):173–183.

Stoica, P. and Sandgren, N. (2006). Spectral analysis of irregularly-sampled data: Paralleling the regularly-sampled data approach. *Digital Signal Processing*, 16(6):712–734.

Swanepoel, J. W. H. and van Wyk, J. W. J. (1986). The bootstrap applied to power spectral density function estimation. *Biometrika*, 73(1):135–141.

Tarczynski, A. and Allay, N. (2004). Spectral analysis of randomly sampled signals: Suppression of aliasing and sampler jitter. *IEEE Trans. Signal Processing*, 52(12):3324–3334.

Tummers, M. J. and Passchier, D. M. (2000). Estimation of the spectral density function from randomly sampled lda data. *10th International Symposium on Applications of Laser Techniques to Fluid Mechanics.*

Venkataramani, R. and Bresler, Y. (2001). Optimal sub-nyquist nonuniform sampling and reconstruction for multiband signals. *IEEE Trans. Signal Processing*, 49(10):2301–2313.

Walker, A. M. (1971). On the estimation of a harmonic component in a time series with stationary independent residuals. *Biometrika*, 58(1):21–36.

Whittaker, E. T. (1915). On the functions which are represented by the expansion of the interpolation theory. *Proc. Roy. Soc. Edingburgh*, Sect. A(35):181–194.

Wolf, T., Cai, Y., Kelly, P., and Gong, W. (2007). Stochastic sampling for internet traffic measurement. *Proc. of 10th IEEE Global Internet Symposium*, pages 31–36.

Zhang, H. and Zimmerman, D. L. (2005). Towards reconciling two asymptotic frameworks in spatial statistics. *Biometrika*, 92(4):921–936.