A simple test for white noise in functional time series

Pramita Bagchi, Vaidotas Characiejus, Holger Dette
Ruhr-Universität Bochum
Fakultät für Mathematik
44780 Bochum
Germany

December 14, 2016

Abstract

We propose a new procedure for white noise testing of a functional time series. Our approach is based on an explicit representation of the L^2-distance between the spectral density operator and its best (L^2-)approximation by a spectral density operator corresponding to a white noise process. The estimation of this distance can be easily accomplished by sums of periodogram kernels and it is shown that an appropriately standardized version of the estimator is asymptotically normal distributed under the null hypothesis (of functional white noise) and under the alternative. As a consequence we obtain a very simple test (using the quantiles of the normal distribution) for the hypothesis of a white noise functional process. In particular the test does neither require the estimation of a long run variance (including a fourth order cumulant) nor resampling procedures to calculate critical values. Moreover, in contrast to all other methods proposed in the literature our approach also allows to test for "relevant" deviations from white noise and to construct confidence intervals for a measure which measures the discrcepancy of the underlying process from a functional white noise process.

1 Introduction

The problem of testing for white noise in dependent data is of particular importance because these tests are frequently used to check the adequacy of a postulated parametric model. The seminal work on this problem can be found in the papers of Box and Pierce (1970); Ljung and Box (1978); Pierce (1972) who proposed *portmanteau tests* to check the goodness of fit of an ARMA model. They operate in the time domain and are based on a sum of squared correlations with fixed lag truncation number [see also Dette and Spreckelsen (2000); Mokkadem (1997) for some more recent references]. The asymptotic properties of the different test statistics considered in these papers are usually derived under the assumption of independent and identically distributed (i.i.d.) innovations and several authors point out that these tests are are not reliable if the innovations are uncorrelated but not independent [see Romano and Thombs (1996) and Francq et al. (2005) among others].

An alternative to the tests operating in the time domain are frequency domain tests, which are based on a comparison between the spectral density corresponding to the process of the innovations and the spectral density of a white noise. For example, Hong (1996) proposed to use an L^2-distance between a kernel-based spectral density estimator and the spectral density of the noise under the null hypothesis to construct a test statistic and this approach has been more recently further developed by Shao (2011). We also refer to Dette and Spreckelsen (2003); Paparoditis (2000) for some results testing more general hypotheses by investigating distances between a parametric and non-parametric spectral density estimate. Other authors propose to use normalized cumulated deviations between a non-parametrically and a parametrically estimated spectral density [see for example Deo (2000)]. All methods mentioned in this and the previous paragraph require the specification of a regularization parameter (lag number or bandwidth). Dette et al. (2011) proposed to estimate the L^2-distance between the unknown density directly using sums of (squared) periodograms. The corresponding test statistic does not require regularization and under the null hypothesis and the additional assumption of a linear moving averaging process with Gaussian innovations its asymptotic distribution is a centered normal distribution with an easily estimable variance. As a consequence a very simple test for white noise can be proposed with attractive finite sample properties.

Due to the increasing demand in analyzing data providing information about curves, surfaces or anything else varying over a continuum many of these methods have been recently further developed to be applicable for functional data. For a general review on Functional data analysis (FDA) with dependent observations we refer the interested readers to the monograph by Horváth and Kokoszka (2012). A test for the hypotheses of white noise of a sequence of functional observations in the time domain has been proposed by Gabrys and Kokoszka (2007). They combine principle components for functional data analysis with a "classical" portmanteau test. More recently, Horváth et al. (2013) considered an alternative portmanteau test which is based on the sum of the L^2-norms of the empirical covariance kernels. As the validity of these tests is only justified under the i.i.d. assumption of the innovation process (and therefore not robust to dependent white noise), Zhang (2016) proposed a spectral domain test using a cumulative distance between the periodogram function and its integral with respect to the frequency [for an early result in the one-dimensional case we also refer to Dahlhaus (1988)]. This author proved weak convergence of an appropriately standardized version of this process and derived a Cramer von Mises type statistic with non-pivotal limiting null distribution. To solve this problem a bootstrap procedure is introduced to generate critical values.

The present paper is devoted to an alternative test in the spectral domain for the hypothesis of white noise functional data. Our approach is based on a direct estimate of the L^2-distance between (unknown) spectral density operator and its best approximation by an operator corresponding to functional white noise process. This distance can be estimated directly by sums of periodogram kernels (thus we do not estimate the spectral density kernel, but just real valued functionals of it). We show that that the corresponding test statistic is asymptotically normal distributed such that critical values can easily be obtained. The main advantage of our approach is its simplicity as it neither requires regularization nor bootstrap in its implementation. In particular the latter fact makes it computationally very efficient for functional data. Morever, we also demonstrate by means of a simulation study that the new test is very competitive to an alternative procedure which has recently been proposed in the literature [see Zhang (2016)].

The corresponding model is introduced in Section 2. Section 3 is devoted to the new distance, its estimate and the corresponding asymptotic theory. We also note that our approach (as it is based on a distance) provides a measure of deviation from a functional white noise for which we provide an explicit (and simple) confidence interval. Other statistical applications are also discussed in this section. In Section 4 we investigate the finite sample properties of the new test and compare it with the alternative test proposed by Zhang (2016). Finally, the proofs of the main results are given in Section 5 and Section 6.

2 Notations and preliminaries

Let $L^p([0,1]^k,\mathbb{C})$ with $p \geq 1$ and $k \geq 1$ denote the Banach space of measurable functions $f:[0,1]^k \to \mathbb{C}$ whose absolute value raised to the p-th power has finite integral. The norm of $L^p([0,1]^k,\mathbb{C})$ is defined by

$$\|f\|_p \equiv \Big(\int_{[0,1]^k} |f(x)|^p \, \mathrm{d}x\Big)^{1/p} < \infty.$$

Note that the equality of the $L^p([0,1]^k,\mathbb{C})$ elements is understood in the sense of the norm of their difference being zero. The real and the imaginary parts of the complex number x are denoted by $\operatorname{Re} x$ and $\operatorname{Im} x$ respectively. $\overline{x}$ denotes the complex conjugate of $x \in \mathbb{C}$ and i is the imaginary unit, i.e. $i = \sqrt{-1}$. $L^2([0,1]^k,\mathbb{C})$ is also a Hilbert space with the inner product given by

$$\langle f,g\rangle \equiv \int_{[0,1]^k} f(x)\overline{g(x)} \, \mathrm{d}x$$

for $f,g \in L^2([0,1]^k,\mathbb{C})$. $L^p([0,1]^k,\mathbb{R})$ denotes the corresponding space of real-valued functions.

Suppose that $\{X_t\}_{t\in\mathbb{Z}}$ is a functional time series such that X_t is a random element of $L^2([0,1],\mathbb{R})$ for each $t \in \mathbb{Z}$. We assume that $\{X_t\}_{t\in\mathbb{Z}}$ is strictly stationary in the sense that for any finite set of indices $I \subset \mathbb{Z}$ and any $s \in \mathbb{Z}$, the joint law of $\{X_t, t \in I\}$ coincides with that of $\{X_{t+s}, t \in I\}$. If $\mathbb{E}\|X_0\|_2 < \infty$, there exists a unique function $\mu \in L^2([0,1],\mathbb{R})$ such that $\mathbb{E}\langle f, X_0\rangle = \langle f,\mu\rangle$ for any $f \in L^2([0,1],\mathbb{R})$. It follows that $\mu(\tau) = \mathbb{E}X_0(\tau)$ for almost all $\tau \in [0,1]$. For all $t,s \in \mathbb{Z}$ and $\tau,\sigma \in [0,1]$, we define the autocovariance kernel $r_t \in L^2([0,1]^2,\mathbb{R})$ at lag $t \in \mathbb{Z}$ as

$$r_t(\tau,\sigma) = \mathbb{E}[(X_{t+s}(\tau)-\mu(\tau))(X_s(\sigma)-\mu(\sigma))]$$

provided that $\mathbb{E}\|X_0\|_2^2 < \infty$. The autocovariance operator $\mathscr{R}_t : L^2([0,1],\mathbb{R}) \to L^2([0,1],\mathbb{R})$ at lag t is defined as the integral operator induced by the kernel r_t, i.e.,

$$\mathscr{R}_t h(\tau) = \int_0^1 r_t(\tau,\sigma)h(\sigma)\,\mathrm{d}\sigma$$

for each $h \in L^2([0,1],\mathbb{R})$ and $\tau \in [0,1]$.

The joint cumulant of a real- or complex-valued random vector $(\ \xi_1 \quad \dots \quad \xi_n\ )^{\mathrm{T}}$ with $\mathrm{E}|\xi_i|^n < \infty$ for $i = 1,\dots,n$ is denoted by $\mathrm{cum}(\xi_1,\dots,\xi_n)$. The cumulant of order $p \geq 1$ of a real- or complex-valued random variable ξ with $\mathrm{E}|\xi|^p < \infty$ is denoted by $\mathrm{cum}_p(\xi)$. We quantify the dependence among the observations $\{X_t\}_{t\in\mathbb{Z}}$ using the cumulant kernel of the series. The pointwise definition of the k-th order cumulant kernel is given by

$$\mathrm{cum}\big(X_{t_1}(\tau_1),\dots,X_{t_k}(\tau_k)\big) = \sum_{\nu=\nu_1\cup\cdots\cup\nu_p} (-1)^{p-1}(p-1)!\prod_{l=1}^{p}\mathbb{E}\Big[\prod_{j\in\nu_l} X_{t_j}(\tau_j)\Big],$$

where the sum extends over all unordered partitions of $\{1,2,\dots,k\}$. The cumulant kernel of order k is an element of $L^2([0,1]^k,\mathbb{R})$ under the assumption of $\mathbb{E}\|X_0\|_2^k < \infty$. We also introduce the cumulant spectral density of order k defined by

$$\begin{aligned} &f_{\omega_1,\dots,\omega_k}(\tau_1,\dots,\tau_k) \\ &\quad = \frac{1}{(2\pi)^{k-1}} \sum_{t_1,\dots,t_{k-1}=-\infty}^{\infty} \exp\Big(-i\sum_{j=1}^{k-1}\omega_j t_j\Big)\,\mathrm{cum}\big(X_{t_1}(\tau_1),\dots,X_{t_{k-1}}(\tau_{k-1}),X_0(\tau_k)\big), \end{aligned} \quad (2.1)$$

where the series converges in L^2 under the cumulant mixing condition

$$\text{(B)} \qquad \sum_{t_1,\dots,t_{k-1}=-\infty}^{\infty} \big\|\mathrm{cum}\big(X_{t_1},\dots,X_{t_{k-1}},X_0\big)\big\|_2 < \infty.$$

The cumulant spectral density of order k is uniformly bounded in $\omega_1,\dots,\omega_k$.

Next we introduce some notations for operators. Let H_1 and H_2 be two separable Hilbert spaces. For any operator A from H_1 to H_2, the Hermitian adjoint of A is denoted by A^*. A bounded linear operator $A: H_1 \to H_2$ is a Hilbert-Schmidt operator if

$$|||A|||_2^2 \equiv \sum_{i=1}^{\infty}\|Ae_i\|^2 < \infty,$$

where $\|\cdot\|$ is the norm of the space H_2 and $\{e_i\}_{i\geq 1}$ is any orthonormal basis of H_1. The space of the Hilbert-Schmidt operators is also a Hilbert space with the inner product defined by

$$\langle A,B\rangle_{\mathrm{HS}} \equiv \sum_{i=1}^{\infty}\langle Ae_i, Be_i\rangle$$

for two Hilbert-Schmidt operators A and B, where $\langle\cdot,\cdot\rangle$ is the inner product of the space H_2. Again, this definition is independent of the choice of the basis $\{e_i\}_{i\geq 1}$. A bounded linear operator $A: L^2([0,1]^k,\mathbb{C}) \to L^2([0,1]^k,\mathbb{C})$ is a Hilbert-Schmidt operator if and only if there exists a kernel $k_A \in L^2([0,1]^{2k},\mathbb{C})$ such that

$$Af(x) = \int_{[0,1]^k} k_A(x,y)f(y)\,\mathrm{d}y$$

almost everywhere in $[0,1]^k$ for each $f \in L^2([0,1]^k,\mathbb{C})$ (see Theorem 6.11 of Weidmann (1980)). Furthermore,

$$|||A|||_2^2 = \|k_A\|_2^2 = \int_{[0,1]^k}\int_{[0,1]^k} |k_A(x,y)|^2 \, \mathrm{d}x\mathrm{d}y, \tag{2.2}$$

$$\langle A,B\rangle_{\mathrm{HS}} = \int_{[0,1]^k}\int_{[0,1]^k} k_A(x,y)\overline{k_B(x,y)} \, \mathrm{d}x\mathrm{d}y \tag{2.3}$$

for two Hilbert-Schmidt operators $A: L^2([0,1]^k,\mathbb{C}) \to L^2([0,1]^k,\mathbb{C})$ and $B: L^2([0,1]^k,\mathbb{C}) \to L^2([0,1]^k,\mathbb{C})$ with the kernels k_A and k_B respectively. The adjoint operator A^* is induced by the kernel $k_A^*(x,y) = k_A(y,x)$. A kernel $k_A : [0,1]^{2k} \to \mathbb{C}$ is called a Hilbert-Schmidt kernel if $k_A \in L^2([0,1]^{2k},\mathbb{C})$.

We briefly review the definitions of the spectral density kernel and the spectral density operator that were introduced by Panaretos and Tavakoli (2013a). The spectral density kernel f_ω at frequency $\omega \in \mathbb{R}$ is defined as

$$f_\omega = \frac{1}{2\pi}\sum_{t\in\mathbb{Z}} \exp(-i\omega t) r_t, \tag{2.4}$$

where the series converges in $L^2([0,1]^2,\mathbb{C})$ provided that

$$\sum_{t\in\mathbb{Z}} \|r_t\|_2 = \sum_{t\in\mathbb{Z}} \left\{\int_0^1\int_0^1 |r_t(\tau,\sigma)|^2 \mathrm{d}\tau\mathrm{d}\sigma\right\}^{1/2} < \infty.$$

The spectral density kernel is uniformly bounded and uniformly continuous in ω with respect to $\|\cdot\|_2$ (see Proposition 2.1 of Panaretos and Tavakoli (2013a). The corresponding spectral density operator $\mathscr{F}_\omega : L^2([0,1],\mathbb{R}) \mapsto L^2([0,1],\mathbb{C})$, induced by the spectral density kernel through right integration, is a self-adjoint and non-negative definite for all $\omega \in \mathbb{R}$.

3 White noise testing

We want to test if the time series is a white noise, i.e., the spectral density operator does not depend on the frequency $\omega \in \mathbb{R}$. Formally, we write

$$H_0 : \mathscr{F}_\omega \equiv \mathscr{F} \;\; a.e. \quad \text{vs.} \quad H_a : \mathscr{F}_\omega \neq \mathscr{F} \quad \text{on a set of positive Lebesgue measure} \tag{3.1}$$

for some operator $\mathscr{F} : L^2([0,1],\mathbb{R}) \to L^2([0,1],\mathbb{C})$. Following Dette et al. (2011) we propose to measure deviations from white noise by an L^2 distance and consider the problem of approximating $\mathscr{F}_\omega$ by a constant self-adjoint Hilbert-Schmidt operator $\mathscr{F}$ (corresponding to a white noise functional process) by the distance function

$$M^2(\mathscr{F}) = \int_{-\pi}^{\pi} |||\mathscr{F}_\omega - \mathscr{F}|||_2^2 d\omega. \tag{3.2}$$

Let us define the kernel $\tilde{f} : [0,1]^2 \to \mathbb{C}$ by setting

$$\tilde{f}(\tau,\sigma) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f_\omega(\tau,\sigma)\mathrm{d}\omega \tag{3.3}$$

for each $\tau, \sigma \in [0,1]$. We show that the kernel $\tilde{f}$ is a symmetric, positive definite Hilbert-Schmidt kernel (i.e. $\|\tilde{f}\|_2 < \infty$) and the distance $M^2(\mathscr{F})$ attains its minimum at the Hilbert-Schmidt integral operator $\widetilde{\mathscr{F}} : L^2([0,1],\mathbb{R}) \to L^2([0,1],\mathbb{C})$ defined by

$$\widetilde{\mathscr{F}} h(\tau) = \int_0^1 \tilde{f}(\tau,\sigma) h(\sigma)\, \mathrm{d}\sigma. \tag{3.4}$$

First we establish the properties of the kernel $\tilde{f}$.

Proposition 3.1. *The kernel $\tilde{f}$ defined by* (3.3) *is symmetric (i.e. $\tilde{f}(\tau,\sigma) = \overline{\tilde{f}(\sigma,\tau)}$), positive definite and $\|\tilde{f}\|_2 < \infty$.*

Proof. The assertion that $\tilde{f}$ is symmetric and positive definite follows from the fact f_ω is symmetric and positive definite. Using the Cauchy-Schwarz inequality and the boundedness of the spectral density kernel (see Proposition 2.1 from Panaretos and Tavakoli (2013a)), we establish that

$$\|\tilde{f}\|_2^2 \le \frac{1}{2\pi} \int_0^1 \int_0^1 \int_{-\pi}^{\pi} \left| f_\omega(\tau,\sigma) \right|^2 \mathrm{d}\omega \mathrm{d}\tau \mathrm{d}\sigma \le \sup_{\omega \in [-\pi,\pi]} \|f_\omega\|_2^2 < \infty,$$

which shows that $\tilde{f}$ is a Hilbert-Schmidt kernel. □

We derive an explicit expression for the distance $M^2(\mathscr{F})$, which shows that the minimum of $M^2(\mathscr{F})$ in the class of all Hilbert-Schmidt operators $\mathscr{F} : L^2([0,1],\mathbb{R}) \to L^2([0,1],\mathbb{C})$ is attained at the operator $\widetilde{\mathscr{F}}$ defined by (3.4).

Theorem 3.1. *Suppose that $\mathscr{F} : L^2([0,1],\mathbb{R}) \to L^2([0,1],\mathbb{C})$ is a Hilbert-Schmidt operator. Then*

$$M^2(\mathscr{F}) = \int_{-\pi}^{\pi} |||\mathscr{F}_\omega - \widetilde{\mathscr{F}}|||_2^2 \mathrm{d}\omega + \int_{-\pi}^{\pi} |||\widetilde{\mathscr{F}} - \mathscr{F}|||_2^2 \mathrm{d}\omega,$$

where M^2 is the distance function defined by (3.2) *and $\widetilde{\mathscr{F}}$ is the operator defined by* (3.4). *In particular, $M^2(\mathscr{F})$ is minimized at $\widetilde{\mathscr{F}}$.*

Proof. The fact that the Hilbert-Schmidt operator norm is induced by the Hilbert-Schmidt inner product yields

$$|||\mathscr{F}_\omega - \mathscr{F}|||_2^2 = |||\mathscr{F}_\omega - \widetilde{\mathscr{F}}|||_2^2 + \langle \mathscr{F}_\omega - \widetilde{\mathscr{F}}, \widetilde{\mathscr{F}} - \mathscr{F} \rangle_{\mathrm{HS}} + \langle \widetilde{\mathscr{F}} - \mathscr{F}, \mathscr{F}_\omega - \widetilde{\mathscr{F}} \rangle_{\mathrm{HS}} + |||\widetilde{\mathscr{F}} - \mathscr{F}|||_2^2.$$

Using expression (2.3) for the Hilbert-Schmidt inner product and changing the order of integration, we obtain

$$\int_{-\pi}^{\pi} \langle \mathscr{F}_\omega - \widetilde{\mathscr{F}}, \widetilde{\mathscr{F}} - \mathscr{F} \rangle_{\mathrm{HS}} \mathrm{d}\omega = \int_0^1 \int_0^1 \int_{-\pi}^{\pi} [f_\omega(x,y) - \tilde{f}(x,y)] \mathrm{d}\omega [\tilde{f}(y,x) - \overline{f(x,y)}] \mathrm{d}x \mathrm{d}y = 0.$$

The interchange of the order of integration is justified by noticing that

$$\int_{-\pi}^{\pi} \int_0^1 \int_0^1 |[f_\omega(x,y) - \tilde{f}(x,y)][\tilde{f}(y,x) - \overline{f(x,y)}]| \mathrm{d}x \mathrm{d}y \mathrm{d}\omega \le 2\pi |||\widetilde{\mathscr{F}} - \mathscr{F}^*|||_2 \sup_{\omega \in [-\pi,\pi]} |||\mathscr{F}_\omega - \widetilde{\mathscr{F}}|||_2.$$

A similar argument shows that

$$\int_{-\pi}^{\pi} \langle \widetilde{\mathscr{F}} - \mathscr{F}, \mathscr{F}_\omega - \widetilde{\mathscr{F}} \rangle_{\mathrm{HS}} \mathrm{d}\omega = 0,$$

which completes the proof. □

The next lemma gives us an expression of the minimal distance $M^2(\widetilde{\mathscr{F}})$ in terms of the spectral density kernel f_ω.

Lemma 3.1. *The minimal distance $M^2(\widetilde{\mathscr{F}})$ is given by*

$$M^2(\widetilde{\mathscr{F}}) = \int_0^1 \int_0^1 \int_{-\pi}^{\pi} |f_\omega(\tau,\sigma)|^2 \mathrm{d}\omega \mathrm{d}\tau \mathrm{d}\sigma - \frac{1}{2\pi} \int_0^1 \int_0^1 \left| \int_{-\pi}^{\pi} f_\omega(\tau,\sigma) \mathrm{d}\omega \right|^2 \mathrm{d}\tau \mathrm{d}\sigma, \tag{3.5}$$

where f_ω is the spectral density kernel defined by (2.4).

Proof. Using (2.2) and changing the order of integration it follows that

$$M^2(\widetilde{\mathscr{F}}) = \int_{-\pi}^{\pi} |||\mathscr{F}_\omega - \widetilde{\mathscr{F}}|||_2^2 \mathrm{d}\omega = \int_0^1 \int_0^1 \int_{-\pi}^{\pi} |f_\omega(\tau,\sigma) - \tilde{f}(\tau,\sigma)|^2 \mathrm{d}\omega \mathrm{d}\tau \mathrm{d}\sigma. \tag{3.6}$$

Since

$$|f_\omega(\tau,\sigma) - \tilde{f}(\tau,\sigma)|^2 = |f_\omega(\tau,\sigma)|^2 - f_\omega(\tau,\sigma)\overline{\tilde{f}(\tau,\sigma)} - \tilde{f}(\tau,\sigma)\overline{f_\omega(\tau,\sigma)} + |\tilde{f}(\tau,\sigma)|^2,$$

we obtain

$$\begin{aligned} \int_{-\pi}^{\pi} |f_\omega(\tau,\sigma) - \tilde{f}(\tau,\sigma)|^2 \mathrm{d}\omega &= \int_{-\pi}^{\pi} |f_\omega(\tau,\sigma)|^2 \mathrm{d}\omega - \int_{-\pi}^{\pi} f_\omega(\tau,\sigma) \mathrm{d}\omega \overline{\tilde{f}(\tau,\sigma)} \\ &\quad - \tilde{f}(\tau,\sigma) \int_{-\pi}^{\pi} \overline{f_\omega(\tau,\sigma)} d\omega + \int_{-\pi}^{\pi} |\tilde{f}(\tau,\sigma)|^2 d\omega \\ &= \int_{-\pi}^{\pi} |f_\omega(\tau,\sigma)|^2 \mathrm{d}\omega - \frac{1}{2\pi} \left| \int_{-\pi}^{\pi} f_\omega(\tau,\sigma) \mathrm{d}\omega \right|^2, \end{aligned}$$

and the assertion follows from equation (3.6). □

For the estimation of the minimal distance $M_0^2 := M^2(\widetilde{\mathscr{F}})$ we avoid a direct estimation of the spectral density operator and propose to use the sums of periodograms. More precisely, we consider the functional discrete Fourier transform (fDFT) of the data $\{X_t\}_{t=0}^{T-1}$ defined as

$$\widetilde{X}_\omega^{(T)}(\tau) := \frac{1}{\sqrt{2\pi T}} \sum_{t=0}^{T-1} X_t(\tau) \exp(-i\omega t) \tag{3.7}$$

and consider the periodogram kernel

$$p_\omega^{(T)}(\tau,\sigma) := [\widetilde{X}_\omega^{(T)}(\tau)][\overline{\widetilde{X}_\omega^{(T)}(\sigma)}] = \widetilde{X}_\omega^{(T)}(\tau)\widetilde{X}_{-\omega}^{(T)}(\sigma).$$

The estimator of M_0^2 is then defined by

$$\widehat{M_T^2} = 2\pi \int_0^1 \int_0^1 (S_{T,2}(\tau,\sigma) - S_{T,1}(\tau,\sigma)\overline{S_{T,1}(\tau,\sigma)}) \mathrm{d}\tau \mathrm{d}\sigma$$

where

$$S_{T,1} = \frac{1}{T}\sum_{k=1}^{\lfloor T/2\rfloor}(p_{\omega_k}^{(T)} + \bar{p}_{\omega_k}^{(T)}) \quad \text{and} \quad S_{T,2} = \frac{2}{T}\sum_{k=2}^{\lfloor T/2\rfloor} p_{\omega_k}^{(T)}\bar{p}_{\omega_{k-1}}^{(T)} \tag{3.8}$$

with

$$\omega_k = \frac{2\pi k}{T}$$

for $k = 1, \ldots, \lfloor T/2\rfloor$.

The definition of $\widehat{M_T^2}$ is motivated by Proposition 2.6 and Theorem 2.7 of Panaretos and Tavakoli (2013a) which states

$$\mathbb{E}(p_{\omega_k}^{(T)}(\tau,\sigma)) \approx f_{\omega_k}(\tau,\sigma) \quad \text{and} \quad \mathrm{Cov}(p_{\omega_k}^{(T)}(\tau_1,\sigma_1), p_{\omega_l}^{(T)}(\tau_2,\sigma_2)) \approx 0$$

in L^2 for $k, l \in \{1,2,\ldots,\lfloor T/2\rfloor\}$ and $k \neq l$. Therefore using the fact that $\overline{f_\omega} = f_{-\omega}$ we have

$$\mathbb{E}(S_{T,1}(\tau,\sigma)) \approx \frac{1}{T}\sum_{k=1}^{\lfloor T/2\rfloor}(f_{\omega_k}(\tau,\sigma) + \overline{f_{\omega_k}(\tau,\sigma)}) \approx \frac{1}{2\pi}\int_{-\pi}^{\pi} f_\omega(\tau,\sigma)d\omega$$

$$\mathbb{E}(S_{T,2}(\tau,\sigma)) \approx \frac{2}{T}\sum_{k=1}^{\lfloor T/2\rfloor} f_{\omega_k}(\tau,\sigma)\overline{f_{\omega_{k-1}}(\tau,\sigma)} \approx \frac{2}{2\pi}\int_0^{\pi}|f_\omega(\tau,\sigma)|^2 d\omega = \frac{1}{2\pi}\int_{-\pi}^{\pi}|f_\omega(\tau,\sigma)|^2 d\omega\,.$$

This heuristically motivates the approximation $\mathbb{E}(\widehat{M_T^2}) = M_0^2 = M^2(\widetilde{\mathscr{F}})$ and the use of $\widehat{M_T^2}$ as an estimator of the minimal distance M_0^2. The next theorem is our main result and formalizes these heuristic arguments. It shows that $\widehat{M_T^2}$ is a consistent estimator of M_0^2 and that an appropriately standardized version of $\widehat{M_T^2}$ is asymptotically normal distributed under the null hypothesis and the alternative. The proof is complicated and therefore deferred to Section 5 and 6.

Theorem 3.2. *Suppose that $\{X_k\}_{k\in\mathbb{Z}}$ is a strictly stationary time series with values in $L^2([0,1],\mathbb{R})$, $\mathbb{E}\|X_0\|_2^k < \infty$ for each $k \geq 1$,*

(i) $\sum_{t_1,t_2,t_3=-\infty}^{\infty}\mathbb{E}[\|X_{t_1}X_{t_2}\|_1\|X_{t_3}X_{t_0}\|_1] < \infty$, *where fg denotes a pointwise product of two functions $f : [0,1] \to \mathbb{R}$ and $g : [0,1] \to \mathbb{R}$ defined by $fg(\tau) = f(\tau)g(\tau)$ for each $\tau \in [0,1]$,*

(ii) $\sum_{t_1,t_2,\ldots,t_{k-1}=-\infty}^{\infty}(1+|t_j|)\|\mathrm{cum}(X_{t_1},\ldots,X_{t_{k-1}},X_0)\|_2 < \infty$ *for $j = 1,2,\ldots,k-1$ and all $k \geq 1$.*

Then

$$\sqrt{T}(\widehat{M_T^2} - M_0^2) \xrightarrow{d} N(0,v^2) \quad \textit{as} \quad T \to \infty,$$

where the asymptotic variance v^2 is given by

$$\begin{aligned} v^2 &= 16\pi \int_{[0,1]^4}\int_{-\pi}^{\pi} f_\omega(\tau_1,\sigma_1)f_\omega(\sigma_1,\tau_2)f_\omega(\tau_2,\sigma_2)f_\omega(\sigma_2,\tau_1)d\omega d\tau_1 d\sigma_1 d\tau_2 d\sigma_2 \\ &+ 4\pi \int_{[0,1]^4}\int_{-\pi}^{\pi}|f_\omega(\tau_1,\sigma_1)f_\omega(\tau_2,\sigma_2)|^2 d\omega d\tau_1 d\sigma_1 d\tau_2 d\sigma_2 \end{aligned}$$

$$+ \quad 8\pi \int\limits_{[0,1]^4} \int\limits_{[-\pi,\pi]^2} f_{\omega_1}(\tau_1,\sigma_1) f_{\omega_2}(\tau_2,\sigma_2) f_{\omega_1,-\omega_1,\omega_2}(\sigma_1,\tau_1,\sigma_2,\tau_2) d\omega_1 d\omega_2 \tau_1 d\tau_2 d\sigma_1 d\sigma_2$$
$$- \quad 16 \int\limits_{[0,1]^4} \int\limits_{[-\pi,\pi]^2} f_{\omega_1}(\tau_1,\sigma_1) f_{\omega_2}(\sigma_1,\tau_2) f_{\omega_2}(\tau_2,\sigma_2) f_{\omega_2}(\sigma_2,\tau_1) d\omega_1 d\omega_2 d\tau_1 d\sigma_1 d\tau_2 d\sigma_2$$
$$- \quad 4 \int\limits_{[0,1]^4} \int\limits_{[-\pi,\pi]^3} f_{\omega_1}(\tau_1,\sigma_1) f_{\omega_2}(\tau_2,\sigma_2) f_{\omega_3,-\omega_3,\omega_2}(\sigma_1,\tau_1,\sigma_2,\tau_2) d\omega_1 d\omega_2 d\omega_3 d\tau_1 d\tau_2 d\sigma_1 d\sigma_2$$
$$+ \quad \frac{4}{\pi} \int\limits_{[0,1]^4} \int\limits_{[-\pi,\pi]^3} f_{\omega_1}(\tau_1,\sigma_1) f_{\omega_2}(\tau_2,\sigma_2) f_{\omega_3}(\tau_1,\sigma_2) f_{\omega_3}(\tau_2,\sigma_1) d\omega_1 d\omega_2 d\omega_3 d\tau_1 d\sigma_1 d\tau_2 d\sigma_2$$
$$+ \quad \frac{2}{\pi} \int\limits_{[0,1]^4} \int\limits_{[-\pi,\pi]^4} f_{\omega_1}(\tau_1,\sigma_1) f_{\omega_2}(\tau_2,\sigma_2) f_{\omega_3,-\omega_3,\omega_4}(\sigma_1,\tau_1,\sigma_2,\tau_2) d\omega_1 d\omega_2 d\omega_3 d\omega_4 d\tau_1 d\tau_2 d\sigma_1 d\sigma_2$$

Moreover, under the null hypothesis the asymptotic variance simplifies to

$$v_{H_0}^2 = 8\pi^2 \left(\int_{[0,1]^2} \left| f_0(\tau,\sigma) \right|^2 d\tau d\sigma \right)^2. \tag{3.9}$$

Remark 3.1. Note that the hypotheses in (3.1) can be rewritten as

$$H_0 : M_0^2 = 0 \qquad \text{vs.} \qquad H_a : M_0^2 > 0.$$

Therefore Theorem 3.2 provides a very simple test for these hypotheses by rejecting the null hypothesis H_0 for large values of $\widehat{M_T^2}$. However, although $\widehat{M_T^2}$ will be real with a probability converging to 1 as $T \to \infty$, the statistic $S_{T2}(\tau,\sigma)$ and consequently the estimator $\widehat{M_T^2}$ may be a complex number for a finite sample size. In fact Theorem 3.2 can be interpreted as

$$\mathrm{Re}\left(\sqrt{T}(\widehat{M_T^2} - M_0^2)\right) \xrightarrow{d} \mathcal{N}(0, v^2) \quad \text{as} \quad T \to \infty$$

and

$$\mathrm{Im}\left(\sqrt{T}(\widehat{M_T^2} - M_0^2)\right) \xrightarrow{p} 0 \quad \text{as} \quad T \to \infty.$$

Therefore we test the hypotheses (3.1) by rejecting the null hypothesis of a functional white noise process whenever

$$\mathrm{Re}\left(\widehat{M_T^2}\right) \; > \; \frac{\widehat{v_{H_0}}}{\sqrt{T}} u_{1-\alpha} \; , \tag{3.10}$$

where $u_{1-\alpha}$ denotes the $(1-\alpha)$-quantile of the standard normal distribution and $\widehat{v_{H_0}}$ is the square root of an appropriate estimator of the asymptotic variance under the null hypothesis. It is also worthwhile to mention that in the simulation study conducted in Section 4 we did not observe any case, where $\widehat{M_T^2}$ is in fact a complex number (more precisely, the imaginary part is smalller than the numerical precision).

Because Theorem 3.2 is also valid under the alternative the test (3.10) is obviously consistent. Moreover, Theorem 3.2 also provides a simple approximation of the power of the test, that is

$$\mathbb{P}\Big(Re\big(\widehat{M_T^2}\big) > \widehat{v_{H_0}} u_{1-\alpha} \Big) \approx \Phi\Big(\sqrt{T}\,\frac{M_0^2}{v_{H_1}} - \frac{v_{H_0}}{v_{H_1}}\, u_{1-\alpha}\Big), \tag{3.11}$$

where v_{H_0} and v_{H_1} denote the (asymptotic) standard deviation of $\sqrt{T}\widehat{M_T^2}$ under the null hypothesis and alternative, respectively, and Φ is the distribution function of the standard normal distribution.

Remark 3.2. Under the null hypothesis the variance does not involve fourth order cumulants. Note that $2\pi\mathbb{E}(S_{T2}(\tau,\sigma)) = \int_{-\pi}^{\pi} |f_\omega(\tau,\sigma)|^2 d\omega$, and therefore a consistent estimator of the standard deviation under the null hypothesis is given by

$$\widehat{v_{H_0}} = 2\pi \int_0^1 \int_0^1 S_{T,2}(\tau,\sigma)\, d\tau\, d\sigma. \tag{3.12}$$

Remark 3.3. Besides the simple test for the classical hypotheses of the form (3.1), Theorem 3.2 has further important statistical applications, which will be briefly discussed in this remark.

(a) In applications it is often reasonable to work under the white noise assumption even in cases where the errors show only slight deviations from white noise. In this case a test for the "classical" hypothesis (3.1) is not useful as it rejects the null hypothesis even for small values of M_0^2 if the sample size is sufficiently large. Moreover, if the null hypothesis in (3.1) is not rejected there is no control of the type I error.
In order to address these problems we propose to formulate hypotheses in terms of the L^2-distance M_0^2 and consider *precise hypotheses* as introduced by Berger and Delampady (1987), that is

$$H_\Delta : M_0^2 \leq \Delta \qquad \text{vs.} \qquad K_\Delta : M_0^2 > \Delta\,, \tag{3.13}$$

$$H_\Delta : M_0^2 \geq \Delta \qquad \text{vs.} \qquad K_\Delta : M_0^2 < \Delta\,, \tag{3.14}$$

where Δ is a pre-specified constant. For $\Delta > 0$ we call the alternative in (3.13) *relevant deviation* from white noise and note that the case $\Delta = 0$ in (3.13) corresponds to the "classical" hypothesis (3.1). The alternative in (3.14) is called *similarity* to white noise and of particular importance. Hypotheses of the type (3.14) are useful if one wants to control the type one error when one works under the assumption of a functional white noise error process. Precise hypotheses of the form (3.13) and (3.14) have nowadays been considered in various fields of statistical inference including medical, pharmaceutical, chemistry or environmental statistics [see Chow and Liu (1992) or McBride (1999) among others].
In contrast to other methods, the approach motivated by Theorem 3.2 can be easily used to construct a test for hypotheses of this type. For the sake of brevity we restrict ourselves to the hypothesis of *similarity* to white noise. Then it is easy to see that

an asymptotic level α test for the hypothesis (3.3) is obtained by rejecting the null hypothesis, whenever

$$Re\big(\widehat{M_T^2}\big) - \Delta < \frac{\hat{v}_{H_1}}{\sqrt{T}} u_\alpha\,, \tag{3.15}$$

where $\hat{v}^2_{H_1}$ denotes an estimator of the variance in (3.9) and u_α is the $(1-\alpha)$-quantile of the standard normal distribution. Note that this procedure allows for accepting the null hypothesis of an "approximate" white noise at controlled type I error.

(b) In a a similar way an application of Theorem 3.2 shows that the interval

$$\Big[\max\Big\{0, Re\big(\widehat{M_T^2}\big) - \frac{\hat{v}_{H_1}}{\sqrt{T}} u_{1-\alpha/2}\Big\}, Re\big(\widehat{M_T^2}\big) + \frac{\hat{v}_{H_1}}{\sqrt{T}} u_{1-\alpha/2}\Big]$$

is an asymptotic confidence interval for the deviation M^2 from a white noise functional process.

4 Finite sample properties

In this section, we investigate the finite sample performance of the method proposed in this paper by means of a simulation study. We have calculated the rejection probabilities of the test (3.10) for the sample sizes T = 128, 256, 512 and 1024, where the number of Monte Carlo replications is always 1000. For the sake of a comparison our simulation setup is similar to that of Zhang (2016) who proposed an alternative test for a functional white noise process.

Under the null hypothesis, we simulate i.i.d. data from a standard Brownian motion, Brownian bridge and data from the FARCH(1) process defined as,

$$X_t(\tau) = \epsilon_t(\tau)\sqrt{\tau + \int_0^1 c_\psi \exp\left(\frac{\tau^2+\sigma^2}{2}\right) X_{t-1}^2(\sigma)d\sigma}, \qquad t = 1,2,\ldots,\ \ \tau \in [0,1],$$

where $\{\epsilon_t\}_{t\in\mathbb{Z}}$ is a sequence of i.i.d. standard Brownian motions and $c_\psi = 0.3418$. As explained in Zhang (2016) observations from the FARCH(1) process are uncorrelated but dependent. Therefore most of the white noise testing methods (especially non spectra-based procedures) have a large type I error under this setup. The data are generated on a grid of 1000 equispaced points in $[0,1]$. The kernels S_{T2} and S_{T1} are computed at the 1000×1000 equispaced grid points on $[0,1]^2$. The integrals of the kernels are estimated by averaging of the function values on the grid points. The asymptotic variance of the test statistic under the null hypothesis is estimated by (3.12).

The corresponding results are presented in Table 1. We observe a very good approximation of the nominal level in all cases under considerations (even for the sample size T = 128). For the sake of comparison we also display in Table 1 the simulated level of the bootstrap test proposed in Zhang (2016) (numbers in brackets). This author used a block bootstrap procedure to generate critical values, which requires the specification of

Table 1: *Empirical rejection probabilities (in percentage) of the test* (3.10) *under the null hypothesis. The numbers in brackets give the corresponding results of the test of Zhang (2016).*

	Brownian Motion			Brownian Bridge			FARCH(1)		
T	10%	5%	1%	10%	5%	1%	10%	5%	1%
128	9.5 (11.0)	4.8 (4.2)	1.1 (0.8)	10.8 (11.0)	5.3 (5.4)	0.8 (1.1)	11.1 (10.7)	5.7 (5.9)	0.8 (0.9)
256	9.6 (10.0)	5.1 (4.2)	1.3 (0.9)	10.3 (9.5)	5.4 (4.8)	0.9 (0.7)	10.9 (11.1)	5.5 (5.2)	0.7 (0.9)
512	10.1 (9.9)	5.1 (4.7)	0.8 (0.6)	9.7 (10.3)	5.1 (5.9)	1.0 (1.3)	10.9 (11.1)	5.3 (4.9)	0.8 (0.7)
1024	9.8 (10.0)	4.9 (4.9)	0.9 (0.8)	9.9 (9.9)	5.2 (5.1)	0.8 (1.1)	10.5 (9.8)	5.2 (4.8)	0.7 (1.2)

the block length as a regularization parameter. This parameter was chosen by the minimum volatility index method as described in Section 2.2 (page 79) of Zhang (2016). For this choice we also observe a very good approximation of the nominal level in all cases under consideration.

Under the alternatives, we simulate data from the FAR(1) model

$$X_t(\tau) - \mu(\tau) = \rho(X_{t-1} - \mu)(\tau) + \epsilon_t(\tau), \quad t = 1, 2, \dots \tag{4.1}$$

where ρ denotes an integral operator acting on $L^2[0,1]$ defined by

$$\rho(x)(\tau) = \int_0^1 \mathcal{K}(\tau,\sigma)x(\sigma)d\sigma\ , \ \ x \in L^2[0,1], \tag{4.2}$$

for some kernel $\mathcal{K} \in L^2([0,1]^2)$, and $\{\epsilon_t(\tau)\}_{t\in\mathbb{Z}}$ is a sequence of i.i.d. mean zero innovations in $L^2[0,1]$. For our simulations we use four different FAR(1) models where the innovations are either Brownian motions or Brownian bridges and the kernel in the integral operator (4.2) is either the Gaussian kernel

$$\mathcal{K}_g(\tau,\sigma) = c_g \exp\big((\tau^2 + \sigma^2)/2\big) \tag{4.3}$$

or the Wiener kernel

$$\mathcal{K}_w(\tau,\sigma) = c_w \min(\tau,\sigma), \tag{4.4}$$

where the constants c_g and c_w were chosen such that the corresponding Hilbert-Schmidt norm is equal 0.3. The corresponding results of the new test are presented in Table 2. We observe very good rejection probabilities in all considered models. A comparison with the procedure of Zhang (2016) shows that the power of both tests is very similar. Only for small sample sizes we observe small differences between both procedures. While the test

Table 2: *Empirical rejection probabilities (in percentage) of the test* (3.10) *under the alternative. The model is a* FAR(1) *model with i.i.d. innovations* ϵ_t *from a Brwonian motion or Brwonian bridge and two different integral opertors are considered. The numbers in brackets give the corresponding results of the test of Zhang (2016).*

ϵ_t	Brownian motion					
$\mathscr{K}$	(4.3)			(4.4)		
T	10%	5%	1%	10%	5%	1%
128	82.6 (86.1)	80.7 (83.7)	65.9 (58.5)	87.6 (89.9)	82.4 (83.1)	66.9 (59.7)
256	99.0 (99.6)	98.2 (99.2)	98.2 (99.0)	99.4 (99.9)	98.3 (99.5)	94.2 (98.6)
512	99.8 (99.7)	99.6 (99.5)	99.6 (99.0)	99.9 (99.9)	99.9 (99.8)	99.6 (99.1)
1024	100.0 (100.0)	99.9 (100.0)	99.7 (99.8)	100.0 (100.0)	100.0 (99.8)	99.8 (99.5)
ϵ_t	Brownian bridge					
$\mathscr{K}$	(4.3)			(4.4)		
T	10%	5%	1%	10%	5%	1%
128	80.1 (79.2)	77.4 (68.3)	60.1 (54.4)	87.6 (80.2)	79.9 (65.8)	61.2 (58.1)
256	100.0 (100.0)	97.0 (98.2)	95.5 (97.2)	99.9 (100.0)	98.3 (99.1)	98.1 (98.8)
512	100.0 (100.0)	99.3 (98.7)	99.3 (98.1)	100.0 (100.0)	100.0 (100.0)	98.8 (99.1)
1024	100.0 (100.0)	100.0 (100.0)	100.0 (99.4)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)

of Zhang (2016) has slightly larger power for a FAR(1) model with a Gaussian kernel and i.i.d innovations (in most cases), the new test proposed in this paper usually yields better results for a FAR(1) model with a Wiener kernel and i.i.d innovations.

Our numerical study can be summarized as follows. The new test proposed in this paper exhibits similar properties as the block bootstrap test suggested in Zhang (2016). The latter approach uses resampling, which is computationally expensive for functional data. Moreover, it requires the specification of the length of the blocks for the bootstrap, and the results may depend on this regularization. In contrast the new test does not need a regularization parameter and critical values can be directly obtained from the table of the standard normal distribution. Moreover, the method can be easily extended to test precise hypotheses of the form (3.13) or (3.14) and our results can be used to provide confidence intervals for a measure of deviation from white noise.

Acknowledgments. This work has been supported in part by the Collaborative Research Center "Statistical modeling of nonlinear dynamic processes" (SFB 823, Teilprojekt C1, C3) of the German Research Foundation (DFG).

References

Berger, J. O. and Delampady, M. (1987). Testing precise hypotheses. *Statistical Science*, 2(3):317–335.

Billingsley, P. (1995). *Probability and Measure*. Wiley Series in Probability and Statistics. Wiley.

Box, G. E. P. and Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive integrated moving average time series models. *Journal of the American Statistical Association*, 65:1509–1526.

Brillinger, D. R. (2001). *Time series: data analysis and theory*, volume 36. Siam.

Chow, S.-C. and Liu, P.-J. (1992). *Design and Analysis of Bioavailability and Bioequivalence Studies*. Marcel Dekker, New York.

Cremers, H. and Kadelka, D. (1986). On weak convergence of integral functionals of stochastic processes with applications to processes taking paths in L_p^E. *Stochastic Processes and their Applications*, 21(2):305–317.

Dahlhaus, R. (1988). Empirical spectral processes and their applications to time series analysis. *Stochastic Processes and their Applications*, 30:69–83.

Deo, R. S. (2000). Spectral tests for the martingale hypothesis under conditional heteroscedasticity. *Journal of Econometrics*, 99:291–315.

Dette, H., Kinsvater, T., and Vetter, M. (2011). Testing non-parametric hypotheses for stationary processes by estimating minimal distances. *Journal of Time Series Analysis*, 32(5):447–461.

Dette, H. and Spreckelsen, I. (2000). A test for randomness against ARMA alternatives. *Stochastic Processes and their Applications*, 89:131–139.

Dette, H. and Spreckelsen, I. (2003). A note on a specification test for time series models based on spectral density estimation. *Scandinavian Journal of Statistics*, 30:481–491.

Francq, C., Roy, R., and Zakoïan, J. M. (2005). Diagnostic checking in ARMA models with uncorrelated errors. *Journal of the American Statistical Association*, 100:532–544.

Gabrys, R. and Kokoszka, P. (2007). Portmanteau test of independence for functional observations. *Journal of the American Statistical Association*, 102:1338–1348.

Hong, Y. (1996). Consistent testing for serial correlation of unknown form. *Econometrica*, 64:837–864.

Horváth, L., Hušková, M., and Rice, G. (2013). Test of independence for functional data. *Journal of Multivariate Analysis*, 117:100–119.

Horváth, L. and Kokoszka, P. (2012). *Inference for Functional Data with Applications*. Springer, New York.

Lehmann, E. L. and Casella, G. (1998). *Theory of point estimation*. Springer Science & Business Media.

Ljung, G. M. and Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65:297–303.

McBride, G. B. (1999). Equivalence tests can enhance environmental science and management. *Australian* & *New Zealand Journal of Statistics*, 41(1):19–29.

Mokkadem, A. (1997). A measure of information and its applications to test for randomness against ARMA alternatives and to goodness-of-fit test. *Stochastic Processes and their Applications*, 72:145–159.

Panaretos, V. M. and Tavakoli, S. (2013a). Fourier analysis of stationary time series in function space. *The Annals of Statistics*, 41(2):568–603.

Panaretos, V. M. and Tavakoli, S. (2013b). Supplement to "Fourier analysis of stationary time series in function space".

Paparoditis, E. (2000). Spectral density based goodness-of-fit tests for time series models. *Scandinavian Journal of Statistics*, 27:143–176.

Pierce, D. A. (1972). Residual correlations and diagnostic checking in dynamic-disturbance time series models. *Journal of the American Statistical Association*, 67(339):636–640.

Romano, J. L. and Thombs, L. A. (1996). Inference for autocorrelations under weak assumptions. *Journal of the American Statistical Association*, 91:590–600.

Shao, X. (2011). Testing for white noise under unknown dependence and its applications to goodness-of-fit for time series models. *Econometric Theory*, 27:312–343.

Weidmann, J. (1980). *Linear operators in Hilbert spaces*. Graduate texts in mathematics. Springer-Verlag.

Zhang, X. (2016). White noise testing and model diagnostic checking for functional time series. *Journal of Econometrics*, 194:76–95.

5 Proof of Theorem 3.2

We prove that the random elements $\{I_T\} = \{I_T\}_{T\geq 1}$ with values in $L^1([0,1]^2,\mathbb{C})$ defined by

$$I_T(\tau,\sigma) = \sqrt{T}\Big[2\pi[S_{T,2}(\tau,\sigma)-S_{T,1}(\tau,\sigma)\overline{S_{T,1}(\tau,\sigma)}]-\int_{-\pi}^{\pi}|f_\omega(\tau,\sigma)|^2 d\omega+\frac{1}{2\pi}\Big|\int_{-\pi}^{\pi} f_\omega(\tau,\sigma)d\omega\Big|^2\Big]$$

for each $(\tau,\sigma)\in[0,1]^2$ converge in distribution to a zero mean Gaussian random element $\mathscr{G}$ with values in $L^1([0,1]^2,\mathbb{R})$ and the covariance kernel ν^2 given by formula (6.2) in Proposition 6.2 of the following section, where S_{T_1} and $S_{T,2}$ are defined by (3.8). The proof is based on Theorem 2 of Cremers and Kadelka (1986), which states that I_T converges in distribution to $\mathscr{G}$ as $T\to\infty$ provided that the following three conditions are fulfilled:

(I) the finite dimensional distributions (fdds) of I_T converge to the fdds of $\mathscr{G}$ a.e. as $T\to\infty$;

(II) there exists an integrable function $f:[0,1]^2\to[0,\infty)$ such that

$$\mathbb{E}|I_T(\tau,\sigma)|\leq f(\tau,\sigma)$$

for each $T\geq 1$ and $(\tau,\sigma)\in[0,1]^2$;

(III) for each $(\tau,\sigma)\in[0,1]^2$,

$$\mathbb{E}|I_T(\tau,\sigma)|\to\mathbb{E}|\mathscr{G}(\tau,\sigma)| \quad \text{as} \quad T\to\infty.$$

We establish sufficient conditions for the convergence of the fdds in Subsection 5.1. A sufficient condition for the existence of a non-negative integrable function that satisfies condition (II) is established in Subsection 5.2. Finally, the required convergence of moments is established using the fact that if $I_T(\tau,\sigma)$ converges in distribution to $\mathscr{G}(\tau,\sigma)$ as $T\to\infty$ and $\sup_{T\geq 1}\mathbb{E}|I_T(\tau,\sigma)|^2<\infty$, then $\mathbb{E}|I_T(\tau,\sigma)|<\infty$ and $\mathbb{E}|I_T(\tau,\sigma)|\to\mathbb{E}|\mathscr{G}(\tau,\sigma)|$ as $T\to\infty$ for each $(\tau,\sigma)\in[0,1]^2$ (see Theorem 25.12 and its Corollary on p. 338 of Billingsley (1995)). We have that $\sup_{T\geq 1}\mathbb{E}|I_T(\tau,\sigma)|^2<\infty$ since $\mathbb{E}|I_T(\tau,\sigma)|^2=\mathrm{cum}_2(I_T(\tau,\sigma))=O(1)$ as $T\to\infty$ (by (I)).

We write

$$\sqrt{T}(\widehat{M^2}-M_0^2)=2\pi\int_0^1\int_0^1 I_T(\tau,\sigma)d\tau d\sigma.$$

As the map $I:L^2([0,1]^2,\mathbb{C})\to\mathbb{C}$ defined by $I(f)=\int_0^1\int_0^1 f(\tau,\sigma)d\tau d\sigma$ is continuous, an application of the continuous mapping Theorem gives

$$\sqrt{T}(\widehat{M^2}-M_0^2)\xrightarrow{d}2\pi\int_0^1\int_0^1\mathscr{G}(\tau,\sigma)d\tau d\sigma.$$

Therefore we finally have

$$\sqrt{T}(\widehat{M^2}-M_0^2)\xrightarrow{d}N\Big(0,4\pi^2\int_0^1\int_0^1\int_0^1\int_0^1\nu^2((\tau_1,\sigma_1)(\tau_2,\sigma_2))d\tau_1 d\sigma_1 d\tau_2 d\sigma_2\Big),$$

where the kernel v^2 is defined in Proposition 6.2 of the following section. The asymptotic variance in Theorem 3.2 is now obtained by a straightforward calculation of the integral observing the representation (6.2). Under H_0 the spectral densities f_ω and $f_{\omega_1,-\omega_1,\omega_2}$ are free of ω and ω_1,ω_2 respectively. Therefore under H_0 the limiting variance simplifies to:

$$
\begin{aligned}
v_{H_0}^2 =& 8\pi^2 \int_0^1 \int_0^1 \int_0^1 \int_0^1 |f_0(\tau_1,\sigma_2) f_0(\tau_2,\sigma_1)|^2 d\tau_1 d\tau_2 d\sigma_1 d\sigma_2 \\
=& 8\pi^2 \left(\int_{[0,1]^2} \left| f_0(\tau,\sigma) \right|^2 d\tau d\sigma \right)^2 . \qquad (5.1)
\end{aligned}
$$

5.1 The convergence of the finite-dimensional distributions

To establish the convergence of the fdds, we need to show that

$$
\begin{pmatrix} I_T(\tau_1,\sigma_1), & \dots & , I_T(\tau_d,\sigma_d) \end{pmatrix}^{\mathrm{T}} \xrightarrow{d} \begin{pmatrix} \mathscr{G}(\tau_1,\sigma_1), & \dots & , \mathscr{G}(\tau_d,\sigma_d) \end{pmatrix}^{\mathrm{T}}
$$

as $T \to \infty$ for each $d \geq 1$ and $(\tau_1,\sigma_1),\dots,(\tau_d,\sigma_d) \in S$ where $S \subset [0,1]^{2d}$ and S has Lebesgue measure 1. In order to do that we restrict our attention to the vector

$$
\widetilde{I}_T(\tau_1,\dots,\tau_d,\sigma_1,\dots,\sigma_d) := \sqrt{T} \begin{pmatrix} S_{T,1}(\tau_1,\sigma_1) - \mathbb{E}\left(S_{T,1}(\tau_1,\sigma_1)\right) \\ \vdots \\ S_{T,1}(\tau_d,\sigma_d) - \mathbb{E}\left(S_{T,1}(\tau_d,\sigma_d)\right) \\ S_{T,2}(\tau_1,\sigma_1) - \mathbb{E}\left(S_{T,2}(\tau_1,\sigma_1)\right) \\ \vdots \\ S_{T,2}(\tau_d,\sigma_d) - \mathbb{E}\left(S_{T,2}(\tau_d,\sigma_d)\right) \end{pmatrix} . \qquad (5.2)
$$

First we show that the aforementioned vector converges in distribution to some $N(0,\Sigma)$ random vector and then use the delta method to obtain the desired result. Here we only deal with the case $d = 1$, as the general case can be established similarly with an additional amount of notation. In order to show that the limit distribution of $\tilde{I}_T$ converges to multivariate normal, we use the Cramér-Wold device and show for any vector $c \in \mathbb{R}^2$, the random variable $c'\widetilde{I}_T(\tau,\sigma)$ converges in distribution to $N(0, c'\Sigma c)$ variable.
For this purpose we prove that the cumulants of $c'\widetilde{I}_T(\tau,\sigma)$ converge to the cumulants of a normal distribution. The first cumulant is trivially zero. Using the fact that cumulants of order l are invariant under centering for $l \geq 2$, Proposition 5.1 shows the convergence of higher order cumulants of $c'\widetilde{I}_T(\tau,\sigma)$ to the cumulants of a normal distribution.

Proposition 5.1. *Under assumption* (ii) *of Theorem 3.2 we have,*

$$
\begin{aligned}
\mathrm{cum}_l\left(c_1\sqrt{T} S_{T,2}(\tau,\sigma) + c_2\sqrt{T} S_{T,1}(\tau,\sigma)\right) &= O(1) \quad \text{for} \quad \mathrm{l} = 2, \\
&= o(1) \quad \text{for} \quad \mathrm{l} > 2,
\end{aligned}
$$

for any $c_1, c_2 \in \mathbb{R}$ *and a.e.* $(\tau,\sigma) \in [0,1]^2$.

Proof. Introduce the notation $\mathrm{cum}_{n_1,n_2}(X,Y) = \mathrm{cum}(\underbrace{X,\ldots,X}_{n_1},\underbrace{Y,\ldots,Y}_{n_2})$. Using the linearity of cumulants as in Theorem 2.3.1 from Brillinger (2001) we write

$$\mathrm{cum}_l\left(c_1\sqrt{T}S_{T,1}(\tau,\sigma)+c_2\sqrt{T}S_{T,2}(\tau,\sigma)\right)=\sum_{n=0}^{l}c_1^n c_2^{l-n}\,\mathrm{cum}_{n,l-n}\left(\sqrt{T}S_{T,1}(\tau,\sigma),\sqrt{T}S_{T,2}(\tau,\sigma)\right).$$

We will show that $\mathrm{cum}_{n,l-n}\left(\sqrt{T}S_{T,1}(\tau,\sigma),\sqrt{T}S_{T,2}(\tau,\sigma)\right)$ is bounded for $l=2$ and converges to 0 for $l>2$ for $n=0,\ldots,l$. First we will show it for $n=0$ and then use induction on n, i.e., assuming the result is true for $n=t-1$, We will show the order of $\mathrm{cum}_{n,l-n}\left(\sqrt{T}S_{T,1}(\tau,\sigma),\sqrt{T}S_{T,2}(\tau,\sigma)\right)$ remains same for $n=t$.

Using linearity again, we obtain

$$\begin{aligned}
&\mathrm{cum}_{n,l-n}\left(\sqrt{T}S_{T,1}(\tau,\sigma),\sqrt{T}S_{T,2}(\tau,\sigma)\right)\\
=&\,\mathrm{cum}_{n,l-n}\left(\frac{1}{\sqrt{T}}\sum_{k=1}^{\lfloor T/2\rfloor}(p_{\omega_k}(\tau,\sigma)+p_{\omega_k}(\sigma,\tau)),\frac{2}{\sqrt{T}}\sum_{k=1}^{\lfloor T/2\rfloor}p^{(T)}_{\omega_k}(\tau,\sigma)p^{(T)}_{\omega_{k-1}}(\sigma,\tau)\right)\\
=&\,\mathrm{cum}_{n,l-n}\left(\frac{1}{\sqrt{T}}\sum_{k=1}^{\lfloor T/2\rfloor}p_{\omega_k}(\tau,\sigma),\frac{2}{\sqrt{T}}\sum_{k=1}^{\lfloor T/2\rfloor}p^{(T)}_{\omega_k}(\tau,\sigma)p^{(T)}_{\omega_{k-1}}(\sigma,\tau)\right)\\
&+\mathrm{cum}_{n,l-n}\left(\frac{1}{\sqrt{T}}\sum_{k=1}^{\lfloor T/2\rfloor}p_{\omega_k}(\sigma,\tau),\frac{2}{\sqrt{T}}\sum_{k=1}^{\lfloor T/2\rfloor}p^{(T)}_{\omega_k}(\tau,\sigma)p^{(T)}_{\omega_{k-1}}(\sigma,\tau)\right)\\
=&\,C_{n1}+C_{n2}
\end{aligned}$$

We will argue only for the first term, the second term can be handled similarly.

$$\begin{aligned}
C_{n1}=&\,\mathrm{cum}_{n,l-n}\left(\frac{1}{\sqrt{T}}\sum_{k=1}^{\lfloor T/2\rfloor}p_{\omega_k}(\tau,\sigma),\frac{2}{\sqrt{T}}\sum_{k=1}^{\lfloor T/2\rfloor}p^{(T)}_{\omega_k}(\tau,\sigma)p^{(T)}_{\omega_{k-1}}(\sigma,\tau)\right)\\
=&\frac{2^{l-n}}{T^{l/2}}\sum_{k_1,\ldots,k_l=1}^{\lfloor T/2\rfloor}\mathrm{cum}\left(p^{(T)}_{\omega_{k_1}}(\tau,\sigma),\ldots,p^{(T)}_{\omega_{k_n}}(\tau,\sigma),p^{(T)}_{\omega_{k_{n+1}}}(\tau,\sigma)p^{(T)}_{\omega_{k_{n+1}-1}}(\sigma,\tau),\ldots,p^{(T)}_{\omega_{k_l}}(\tau,\sigma)p^{(T)}_{\omega_{k_l-1}}(\sigma,\tau)\right)\\
=&\frac{2^{l-n}}{T^{l/2}}\sum_{k_1,\ldots,k_l=1}^{\lfloor T/2\rfloor}\mathrm{cum}(Z_{k_1,1}Z_{k_1,2},\ldots,Z_{k_n,1}Z_{k_n,2},Z_{k_{n+1},1}Z_{k_{n+1},2}Z_{k_{n+1},3}Z_{k_{n+1},4}\ldots Z_{k_l,1}Z_{k_l,2}Z_{k_l,3}Z_{k_l,4})
\end{aligned}$$

where $Z_{i1}:=\widetilde{X}^{(T)}_{\omega_i}(\tau)$, $Z_{i2}:=\widetilde{X}^{(T)}_{-\omega_i}(\sigma)$, $Z_{i3}:=\widetilde{X}^{(T)}_{\omega_{i-1}}(\sigma)$ and $Z_{i4}:=\widetilde{X}^{(T)}_{-\omega_{i-1}}(\tau)$. Now using Theorem 2.3.2 from Brillinger (2001) we write

$$C_{n1}=\frac{2^{l-n}}{T^{l/2}}\sum_{k_1=1}^{\lfloor T/2\rfloor}\cdots\sum_{k_l=1}^{\lfloor T/2\rfloor}\sum_{\nu}\mathrm{cum}\left(Z_{ij}:ij\in\nu_1\right)\ldots\mathrm{cum}\left(Z_{ij}:ij\in\nu_p\right)=\sum_{\nu}C_n(\nu)\tag{5.3}$$

with

$$C_n(\nu)=\frac{2^{l-n}}{T^{l/2}}\sum_{k_1=1}^{\lfloor T/2\rfloor}\cdots\sum_{k_l=1}^{\lfloor T/2\rfloor}\mathrm{cum}\left(Z_{ij}:ij\in\nu_1\right)\ldots\mathrm{cum}\left(Z_{ij}:ij\in\nu_p\right)$$

for all indecomposable partitions $\nu = \nu_1 \cup \nu_2 \cup \cdots \cup \nu_p$ of the table

$$T_n = \begin{cases} (k_1,1) & (k_1,2) \\ \vdots & \vdots \\ (k_n,1) & (k_n,2) \\ (k_{n+1},1) & (k_{n+1},2) & (k_{n+1},3) & (k_{n+1},4) \\ \vdots & \vdots & \vdots & \vdots \\ (k_l,1) & (k_l,2) & (k_l,3) & (k_l,4) \end{cases} \tag{5.4}$$

As there are finitely many partitions, we will show that $C_n(\nu)$ is of right order for all indecomposable partition of T_n and $n = 0,1,\ldots,l$. To this end, we claim that $C_0(\nu) = O(1)$ for $l = 2$ and $o(1)$ for $l > 2$ for all indecomposable partitions ν of T_0. The proof of this claim is presented in Section 6.2.

Now suppose that we have proved $C_n(\nu)$ has the right order ($O(1)$ for $l = 2$ and $o(1)$ for $l > 2$) for $n = 0,\ldots,t-1$. Let $\nu = \nu_1 \cup \cdots \cup \nu_p$ be an indecomposable partition of table T_t and $\nu_{p+1} = \{(k_t,3),(k_t,4)\}$. Then $\nu' = \nu_1 \cup \cdots \cup \nu_p \cup \nu_{p+1}$ is an indecomposable partition of T_{t-1}. Using Theorem B.2 from Panaretos and Tavakoli (2013b) we get $\mathrm{cum}(Z_{ij} : ij \in \nu_{p+1}) = O(1)$ and

$$\begin{aligned} C_{t-1}(\nu') =& \frac{2^{l-t+1}}{T^{l/2}} \sum_{k_1=1}^{\lfloor T/2 \rfloor} \cdots \sum_{k_l=1}^{\lfloor T/2 \rfloor} \mathrm{cum}\big(Z_{ij} : ij \in \nu_1\big) \ldots \mathrm{cum}\big(Z_{ij} : ij \in \nu_p\big)\, O(1) \\ =& 2C_t(\nu) \times O(1). \end{aligned}$$

Therefore $C_t(\nu)$ is of the right order and hence the result is true. □

Note that $(I_T(\tau_1,\sigma_1),\ldots,I_T(\tau_d,\sigma_d)) = g(\tilde{I}_t(\tau_1,\ldots,\sigma_d))$ where $g : \mathbb{R}^{2d} \to \mathbb{R}^d$, defined as $g(x_1,x_2,\ldots,x_{2d}) = (x_{d+1} - x_1^2,\ldots,x_{2d} - x_d^2)$. Therefore an application of the delta method (Theorem 8.22 from Lehmann and Casella (1998)) along with Lemma 6.1 establishes the convergence of the fdds of I_T to the fdds of $\mathcal{G}$ almost everywhere as $T \to \infty$.

5.2 The dominating function

We establish a sufficient condition for the existence of a non-negative integrable function that satisfies condition (II).

Theorem 5.1. *There exists an integrable function $f : [0,1]^2 \to [0,\infty)$ such that*

$$\mathbb{E}|I_T(\tau,\sigma)| \le f(\tau,\sigma)$$

for each $T \ge 1$ and $(\tau,\sigma) \in [0,1]^2$ if

$$\sum_{t_1,t_2,t_3=-\infty}^{\infty} \mathbb{E}[\|X_{t_1}X_{t_2}\|_1 \|X_{t_3}X_0\|_1] < \infty,$$

where fg denotes a pointwise product of two functions $f : [0,1] \to \mathbb{R}$ and $g : [0,1] \to \mathbb{R}$ defined by $fg(\tau) = f(\tau)g(\tau)$ for each $\tau \in [0,1]$.

We need an auxiliary lemma to prove Theorem 5.1.

Lemma 5.1. *For each* $(\tau,\sigma)\in[0,1]^2$, $\omega,\lambda\in\mathbb{R}$ *and* $T\geq 1$,

$$\mathbb{E}|p_\omega^{(T)}(\tau,\sigma)p_\lambda^{(T)}(\tau,\sigma)| \leq \frac{7}{(2\pi)^2 T}\sum_{t_1,t_2,t_3=-\infty}^{\infty}|\mathbb{E}[X_{t_1}(\tau)X_{t_2}(\sigma)X_{t_3}(\tau)X_0(\sigma)]|.$$

Proof. By the Cauchy-Schwarz inequality,

$$\mathbb{E}|p_\omega^{(T)}(\tau,\sigma)p_\lambda^{(T)}(\tau,\sigma)| \leq (\mathbb{E}|p_\omega^{(T)}(\tau,\sigma)|^2)^{1/2}(\mathbb{E}|p_\lambda^{(T)}(\tau,\sigma)|^2)^{1/2}.$$

We have that

$$\mathbb{E}|p_\omega^{(T)}(\tau,\sigma)|^2 = \mathbb{E}[\widetilde{X}_\omega(\tau)\widetilde{X}_{-\omega}(\sigma)\widetilde{X}_{-\omega}(\tau)\widetilde{X}_\omega(\sigma)].$$

The definition of the fDFT, the linearity of the expectation and the stationarity of the sequence $\{X_t\}_{t\in\mathbb{Z}}$ yield

$$\begin{aligned}
&E[\widetilde{X}_\omega(\tau)\widetilde{X}_{-\omega}(\sigma)\widetilde{X}_{-\omega}(\tau)\widetilde{X}_\omega(\sigma)]\\
&=\frac{1}{(2\pi T)^2}\sum_{u_1,u_2,u_3,u_4=0}^{T-1}\exp(-i\omega(u_1-u_2-u_3+u_4))\mathbb{E}[X_{u_1}(\tau)X_{u_2}(\sigma)X_{u_3}(\tau)X_{u_4}(\sigma)]\\
&=\frac{1}{(2\pi T)^2}\sum_{u_1,u_2,u_3,u_4=0}^{T-1}\exp(-i\omega(u_1-u_2-u_3+u_4))\mathbb{E}[X_{u_1-u_4}(\tau)X_{u_2-u_4}(\sigma)X_{u_3-u_4}(\tau)X_0(\sigma)].
\end{aligned}$$

Let $h^T(t)=1$ for $0\leq t\leq T-1$ and 0 otherwise. By setting $t_i=u_i-u_4$ for $1\leq i\leq 3$ and $t=t_4$, we obtain

$$\begin{aligned}
&\mathbb{E}[\widetilde{X}_\omega(\tau)\widetilde{X}_{-\omega}(\sigma)\widetilde{X}_{-\omega}(\tau)\widetilde{X}_\omega(\sigma)]\\
&=\frac{1}{(2\pi T)^2}\sum_{t_1,t_2,t_3=-(T-1)}^{T-1}\exp(-i\omega(t_1-t_2-t_3))\mathbb{E}[X_{t_1}(\tau)X_{t_2}(\sigma)X_{t_3}(\tau)X_0(\sigma)]\\
&\quad\times\sum_{t\in\mathbb{Z}}h^{(T)}(t_1+t)h^{(T)}(t_2+t)h^{(T)}(t_3+t)h^{(T)}(t)\\
&=\frac{1}{(2\pi)^2 T}\sum_{t_1,t_2,t_3=-(T-1)}^{T-1}\exp(-i\omega(t_1-t_2-t_3))\mathbb{E}[X_{t_1}(\tau)X_{t_2}(\sigma)X_{t_3}(\tau)X_0(\sigma)]\\
&\quad+\frac{1}{(2\pi T)^2}\sum_{t_1,t_2,t_3=-(T-1)}^{T-1}\exp(-i\omega(t_1-t_2-t_3))\mathbb{E}[X_{t_1}(\tau)X_{t_2}(\sigma)X_{t_3}(\tau)X_0(\sigma)]\\
&\quad\times\sum_{t\in\mathbb{Z}}[h^{(T)}(u_1+t)h^{(T)}(u_2+t)h^{(T)}(u_3+t)h^{(T)}(t)-[h^{(T)}(t)]^4]
\end{aligned}$$

since

$$\sum_{t\in\mathbb{Z}}[h^{(T)}(t)]^4=T.$$

Using Lemma F.7 from Panaretos and Tavakoli (2013b) it follows that

$$\mathbb{E}[\widetilde{X}_\omega(\tau)\widetilde{X}_{-\omega}(\sigma)\widetilde{X}_{-\omega}(\tau)\widetilde{X}_\omega(\sigma)]$$

$$
\begin{aligned}
&\leq \frac{1}{(2\pi)^2 T} \sum_{t_1,t_2,t_3=-(T-1)}^{T-1} |\mathbb{E}[X_{t_1}(\tau)X_{t_2}(\sigma)X_{t_3}(\tau)X_0(\sigma)]| \\
&\quad + \frac{2}{(2\pi T)^2} \sum_{t_1,t_2,t_3=-(T-1)}^{T-1} (|t_1|+|t_2|+|t_3|)|\mathbb{E}[X_{t_1}(\tau)X_{t_2}(\sigma)X_{t_3}(\tau)X_0(\sigma)]| \\
&\leq \frac{7}{(2\pi)^2 T} \sum_{t_1,t_2,t_3=-(T-1)}^{T-1} |\mathbb{E}[X_{t_1}(\tau)X_{t_2}(\sigma)X_{t_3}(\tau)X_0(\sigma)]| \\
&\leq \frac{7}{(2\pi)^2 T} \sum_{t_1,t_2,t_3=-\infty}^{\infty} |\mathbb{E}[X_{t_1}(\tau)X_{t_2}(\sigma)X_{t_3}(\tau)X_0(\sigma)]|.
\end{aligned}
$$

Analogously, we obtain

$$
\mathbb{E}[\widetilde{X}_\lambda(\tau)\widetilde{X}_{-\lambda}(\sigma)\widetilde{X}_{-\lambda}(\tau)\widetilde{X}_\lambda(\sigma)] \leq \frac{7}{(2\pi)^2 T} \sum_{t_1,t_2,t_3=-\infty}^{\infty} |\mathbb{E}[X_{t_1}(\tau)X_{t_2}(\sigma)X_{t_3}(\tau)X_0(\sigma)]|.
$$

which completes the proof of Lemma 5.1. □

Proof of Theorem 5.1. We have that

$$
\mathbb{E}|I_T(\tau,\sigma)| \leq 2\sqrt{T}\mathbb{E}|S_{T,2}(\tau,\sigma)| + 2\sqrt{T}\mathbb{E}|S_{T,1}(\tau,\sigma)\overline{S}_{T,1}(\tau,\sigma)|
$$

using the inequality $\mathbb{E}|\xi - \mathbb{E}\xi| \leq 2\mathbb{E}|\xi|$ for any random variable ξ such that $\mathbb{E}|\xi| < \infty$. Using the definition of $S_{T,2}$ and the triangle inequality,

$$
\sqrt{T}\mathbb{E}|S_{T,2}(\tau,\sigma)| \leq \frac{2}{\sqrt{T}} \sum_{k=1}^{\lfloor T/2 \rfloor} \mathbb{E}|p_{\omega_k}^{(T)}(\tau,\sigma)\overline{p}_{\omega_{k-1}}^{(T)}(\tau,\sigma)|.
$$

The fact that $\overline{p}_{\omega_{k-1}}^{(T)}(\tau,\sigma) = p_{-\omega_{k-1}}^{(T)}(\tau,\sigma)$ and the bound of Lemma 5.1 now yield

$$
\sqrt{T}\mathbb{E}|S_{T,2}(\tau,\sigma)| \leq \frac{7}{(2\pi)^2} \sum_{t_1,t_2,t_3=-\infty}^{\infty} |\mathbb{E}[X_{t_1}(\tau)X_{t_2}(\sigma)X_{t_3}(\tau)X_0(\sigma)]|
$$

for each $T \geq 1$ and $(\tau,\sigma) \in [0,1]^2$. Similarly, using the triangle inequality,

$$
\begin{aligned}
\mathbb{E}|S_{T,1}(\tau,\sigma)\overline{S}_{T,1}(\tau,\sigma)| \leq \frac{1}{T^2}\Big[&\sum_{k=1}^{\lfloor T/2 \rfloor}\sum_{l=1}^{\lfloor T/2 \rfloor} \mathbb{E}|p_{\omega_k}^{(T)}(\tau,\sigma)p_{\omega_l}^{(T)}(\tau,\sigma)| \\
&+ \sum_{k=1}^{\lfloor T/2 \rfloor}\sum_{l=1}^{\lfloor T/2 \rfloor} \mathbb{E}|\overline{p}_{\omega_k}^{(T)}(\tau,\sigma)\overline{p}_{\omega_l}^{(T)}(\tau,\sigma)| + 2\sum_{k=1}^{\lfloor T/2 \rfloor}\sum_{l=1}^{\lfloor T/2 \rfloor} \mathbb{E}|\overline{p}_{\omega_k}^{(T)}(\tau,\sigma)p_{\omega_l}^{(T)}(\tau,\sigma)|\Big]
\end{aligned}
$$

and

$$
\sqrt{T}\mathbb{E}|S_{T,1}(\tau,\sigma)\overline{S}_{T,1}(\tau,\sigma)| \leq \frac{7}{(2\pi)^2} \sum_{t_1,t_2,t_3=-\infty}^{\infty} |\mathbb{E}[X_{t_1}(\tau)X_{t_2}(\sigma)X_{t_3}(\tau)X_0(\sigma)]|
$$

using the bound of Lemma 5.1. The proof is complete. □

6 More technical details

6.1 Limiting Mean and Variance Calculation

In this Section we calculate the limiting covariance kernel of the process I_T. The main result of this section is stated in Proposition 6.2.

Lemma 6.1. *Under the assumption* (ii) *of Theorem 3.2, we have*

$$\sqrt{T}\left(\mathbb{E}(S_{T,2}(\tau,\sigma)) - \frac{1}{2\pi}\int_{-\pi}^{\pi} |f_\omega(\tau,\sigma)|^2 d\omega\right) \to 0$$

$$\sqrt{T}\left(\mathbb{E}(S_{T,1}(\tau,\sigma)) - \frac{1}{2\pi}\int_{-\pi}^{\pi} f_\omega(\tau,\sigma) d\omega\right) \to 0$$

as $T \to \infty$ *for almost every* $(\tau,\sigma) \in [0,1]^2$.

Proof. Using Proposition 2.5 and 2.6 from Panaretos and Tavakoli (2013a) we obtain for almost every $(\tau,\sigma) \in [0,1]^2$,

$$\begin{aligned}
\mathbb{E}(S_{T,2}(\tau,\sigma)) =& \frac{2}{T}\sum_{k=1}^{\lfloor T/2\rfloor} \mathbb{E}\left(p^{(T)}_{\omega_k}(\tau,\sigma)p^{(T)}_{\omega_{k-1}}(\sigma,\tau)\right) \\
=& \frac{2}{T}\sum_{k=1}^{\lfloor T/2\rfloor} \mathrm{Cov}\left(p^{(T)}_{\omega_k}(\tau,\sigma), p^{(T)}_{\omega_{k-1}}(\sigma,\tau)\right) + \frac{2}{T}\sum_{k=1}^{\lfloor T/2\rfloor} \mathbb{E}\left(p^{(T)}_{\omega_k}(\tau,\sigma)\right)\mathbb{E}\left(p^{(T)}_{\omega_{k-1}}(\sigma,\tau)\right) \\
=& \frac{2}{T}\sum_{k=1}^{\lfloor T/2\rfloor} O(T^{-1}) + \frac{2}{T}\sum_{k=1}^{\lfloor T/2\rfloor} \left(f_{\omega_k}(\tau,\sigma)f_{\omega_{k-1}}(\sigma,\tau) + O(T^{-1})\right) \\
=& \frac{2}{T}\sum_{k=1}^{\lfloor T/2\rfloor} f_{\omega_k}(\tau,\sigma)f_{\omega_{k-1}}(\sigma,\tau) + O(T^{-1}) \\
\mathbb{E}(S_{T,1}(\tau,\sigma)) =& \frac{1}{T}\sum_{k=1}^{\lfloor T/2\rfloor} \mathbb{E}\left(p^{(T)}_{\omega_k}(\tau,\sigma) + p^{(T)}_{\omega_k}(\sigma,\tau)\right) = \frac{1}{T}\sum_{k=1}^{\lfloor T/2\rfloor} \left(f_{\omega_k}(\tau,\sigma) + f_{\omega_k}(\sigma,\tau)\right) + O(T^{-1}).
\end{aligned}$$

An upper bound on the approximation error for the integral by sum is given by

$$\begin{aligned}
&\left|\frac{1}{T}\sum_{k=1}^{\lfloor T/2\rfloor} f_{\omega_k}(\tau,\sigma)f_{\omega_{k-1}}(\sigma,\tau) - \frac{1}{2\pi}\int_0^{\pi} |f_\omega(\tau,\sigma)|^2 d\omega\right| \\
\le& \frac{1}{2\pi}\sum_{k=1}^{\lfloor T/2\rfloor}\sum_{t_1,t_2}\int_{\omega_{k-1}}^{\omega_k} \Big|\exp(-i\omega_k t_1 - i\omega_{k-1}t_2) - \exp(-i\omega(t_1+t_2))\Big| d\omega |r_{t_1}(\tau,\sigma)r_{t_2}(\sigma,\tau)| \\
\le& \frac{2\pi}{T^2}\sum_{k=1}^{\lfloor T/2\rfloor}\sum_{t_1,t_2\le N} \frac{1}{|t_1+t_2|}|r_{t_1}(\tau,\sigma)r_{t_2}(\sigma,\tau)| + \sum_{t_1,t_2>N} |r_{t_1}(\tau,\sigma)r_{t_2}(\sigma,\tau)| \qquad (6.1)
\end{aligned}$$

As $\sum_{t_1,t_2}\int_0^1\int_0^1 |r_{t_1}(\tau,\sigma)r_{t_2}(\sigma,\tau)|d\tau d\sigma \le (\sum_t \|r_t\|_2)^2 < \infty$, therefore $\sum_{t_1,t_2}|r_{t_1}(\tau,\sigma)r_{t_2}(\sigma,\tau)| < \infty$ for almost every (τ,σ), and hence we can choose N appropriately so that the upper bound given in (6.1) is of order $O(T^{-1})$. The term $\mathbb{E}(S_{T,1}(\tau,\sigma)$ can be dealt with similarly. □

Lemma 6.2. *Under assumption (ii) of Theorem 3.2, for almost all* $(\tau_1,\sigma_1,\tau_2,\sigma_2)\in[0,1]^4$*, the limit of the covariance matrix* Σ *of the vector* $\widetilde{I}_T(\tau_1,\tau_2,\sigma_1,\sigma_2)$ *defiend in* (5.2) *is given by*

$$\begin{aligned}
\Sigma_{12}=\Sigma_{21}\to&\frac{1}{2\pi}\int_{-\pi}^{\pi}\big(f_\omega(\tau_1,\sigma_2)f_\omega(\tau_2,\sigma_1)+f_\omega(\tau_1,\tau_2)f_\omega(\sigma_2,\sigma_1)\big)\,d\omega\\
&+\frac{1}{2\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}f_{\omega_1,-\omega_1,\omega_2}(\tau_1,\sigma_1,\tau_2,\sigma_2)\omega_1d\omega_2\\
\Sigma_{34}=\Sigma_{43}\to&\frac{2}{\pi}\int_{-\pi}^{\pi}f_\omega(\tau_1,\sigma_1)f_\omega(\sigma_1,\tau_2)f_\omega(\sigma_2,\tau_1)f_\omega(\tau_2,\sigma_2)d\omega\\
&+\frac{2}{\pi}\int_{-\pi}^{\pi}f_\omega(\tau_1,\sigma_1)f_\omega(\tau_2,\tau_1)f_\omega(\sigma_1,\sigma_2)f_\omega(\sigma_2,\tau_2)d\omega\\
&+\frac{1}{\pi}\int_{-\pi}^{\pi}f_\omega(\tau_1,\sigma_2)f_\omega(\tau_2,\sigma_1)f_\omega(\sigma_1,\tau_2)f_\omega(\sigma_2,\tau_1)d\omega\\
&+\frac{2}{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}f_{\omega_1}(\tau_1,\sigma_1)f_{\omega_2}(\tau_2,\sigma_2)f_{\omega_1,-\omega_1,\omega_2}(\sigma_1,\tau_1,\sigma_2,\tau_2)d\omega_1d\omega_2\\
\Sigma_{23}=\Sigma_{32}\to&\frac{1}{\pi}\int_{-\pi}^{\pi}f_\omega(\tau_1,\sigma_1)f_\omega(\sigma_1,\sigma_2)f_\omega(\tau_2,\tau_1)d\omega+\frac{1}{\pi}\int_{-\pi}^{\pi}f_\omega(\tau_1,\sigma_1)f_\omega(\sigma_1,\tau_2)f_\omega(\sigma_2,\tau_1)d\omega\\
&+\frac{1}{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}f_{\omega_1}(\tau_1,\sigma_1)f_{\omega_1,-\omega_1,\omega_2}(\sigma_1,\tau_1,\sigma_2,\tau_2)d\omega_1d\omega_2\\
\Sigma_{14}=\Sigma_{41}\to&\frac{1}{\pi}\int_{-\pi}^{\pi}f_\omega(\tau_2,\sigma_2)f_\omega(\sigma_2,\sigma_1)f_\omega(\tau_1,\tau_2)d\omega+\frac{1}{\pi}\int_{-\pi}^{\pi}f_\omega(\tau_2,\sigma_2)f_\omega(\sigma_2,\tau_1)f_\omega(\sigma_1,\tau_2)d\omega\\
&+\frac{1}{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}f_{\omega_1}(\tau_2,\sigma_2)f_{\omega_1,-\omega_1,\omega_2}(\sigma_2,\tau_2,\sigma_1,\tau_1)d\omega_1d\omega_2\\
\Sigma_{13}=\Sigma_{31}\to&\frac{1}{\pi}\int_{-\pi}^{\pi}f_\omega(\tau_1,\sigma_1)f_\omega(\sigma_1,\sigma_1)f_\omega(\tau_1,\tau_1)d\omega+\frac{1}{\pi}\int_{-\pi}^{\pi}f_\omega(\tau_1,\sigma_1)f_\omega^2(\sigma_1,\tau_1)d\omega\\
&+\frac{1}{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}f_{\omega_1}(\tau_1,\sigma_1)f_{\omega_1,-\omega_1,\omega_2}(\sigma_1,\tau_1,\sigma_1,\tau_1)d\omega_1d\omega_2\\
\Sigma_{24}=\Sigma_{42}\to&\frac{1}{\pi}\int_{-\pi}^{\pi}f_\omega(\tau_2,\sigma_2)f_\omega(\sigma_2,\sigma_2)f_\omega(\tau_2,\tau_2)d\omega+\frac{1}{\pi}\int_{-\pi}^{\pi}f_\omega(\tau_2,\sigma_2)f_\omega^2(\sigma_2,\tau_2)d\omega\\
&+\frac{1}{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}f_{\omega_1}(\tau_2,\sigma_2)f_{\omega_1,-\omega_1,\omega_2}(\sigma_2,\tau_2,\sigma_2,\tau_2)d\omega_1d\omega_2.
\end{aligned}$$

$$\begin{aligned}
\Sigma_{ii}\to&\frac{1}{2\pi}\int_{-\pi}^{\pi}\big(f_\omega^2(\tau_i,\sigma_i)+f_\omega(\tau_i,\tau_i)f_\omega(\sigma_i,\sigma_i)\big)\,d\omega\\
&+\frac{1}{2\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}f_{\omega_1,-\omega_1,\omega_2}(\tau_i,\sigma_i,\tau_i,\sigma_i)\omega_1d\omega_2\quad\text{for}\quad \mathrm{i}=1,2.\\
\Sigma_{ii}\to&\frac{2}{\pi}\int_{-\pi}^{\pi}\big|f_\omega(\tau_{i-2},\sigma_{i-2})\big|^4d\omega+\frac{2}{\pi}\int_{-\pi}^{\pi}\big|f_\omega(\tau_{i-2},\sigma_{i-2})\big|^2f_\omega(\tau_{i-2},\tau_{i-2})f_\omega(\sigma_{i-2},\sigma_{i-2})d\omega\\
&+\frac{2}{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}f_{\omega_1}(\tau_{i-2},\sigma_{i-2})f_{\omega_2}(\tau_{i-2},\sigma_{i-2})f_{\omega_1,-\omega_1,\omega_2}(\sigma_{i-2},\tau_{i-2},\sigma_{i-2},\tau_{i-2})d\omega_1d\omega_2,\quad\text{for}\quad\mathrm{i}=3,4.
\end{aligned}$$

Proof. We start with

$$\Sigma_{12}=\Sigma_{21}=T\mathrm{Cov}\big(S_{T,1}(\tau_1,\sigma_1),S_{T,1}(\tau_2,\sigma_2)\big)$$

$$
\begin{aligned}
&= \frac{1}{T}\mathrm{Cov}\left(\sum_{k=1}^{\lfloor T/2\rfloor}\left(p_{\omega_k}^{(T)}(\tau_1,\sigma_1)+p_{\omega_k}^{(T)}(\sigma_1,\tau_1)\right),\sum_{k=1}^{\lfloor T/2\rfloor}\left(p_{\omega_k}^{(T)}(\tau_2,\sigma_2)+p_{\omega_k}^{(T)}(\sigma_2,\tau_2)\right)\right)\\
&= \frac{1}{T}\sum_{k=1}^{\lfloor T/2\rfloor}\sum_{l=1}^{\lfloor T/2\rfloor}\mathrm{Cov}\left(p_{\omega_k}^{(T)}(\tau_1,\sigma_1),p_{\omega_l}^{(T)}(\tau_2,\sigma_2)\right)+\frac{1}{T}\sum_{k=1}^{\lfloor T/2\rfloor}\sum_{l=1}^{\lfloor T/2\rfloor}\mathrm{Cov}\left(p_{\omega_k}^{(T)}(\sigma_1,\tau_1),p_{\omega_l}^{(T)}(\sigma_2,\tau_2)\right)\\
&\quad+\frac{1}{T}\sum_{k=1}^{\lfloor T/2\rfloor}\sum_{l=1}^{\lfloor T/2\rfloor}\mathrm{Cov}\left(p_{\omega_k}^{(T)}(\tau_1,\sigma_1),p_{\omega_l}^{(T)}(\sigma_2,\tau_2)\right)+\frac{1}{T}\sum_{k=1}^{\lfloor T/2\rfloor}\sum_{l=1}^{\lfloor T/2\rfloor}\mathrm{Cov}\left(p_{\omega_k}^{(T)}(\sigma_1,\tau_1),p_{\omega_l}^{(T)}(\tau_2,\sigma_2)\right)
\end{aligned}
$$

First we calculate

$$
\begin{aligned}
\mathrm{Cov}\left(p_{\omega_k}^{(T)}(\tau_1,\sigma_1),p_{\omega_l}^{(T)}(\tau_2,\sigma_2)\right) &= \mathrm{cum}(AB,CD)\\
&=(A,B,C,D)+(A,C)(B,D)+(A,D)(B,C)
\end{aligned}
$$

where $A=\widetilde{X}_{\omega_k}^T(\tau_1)$, $B=\widetilde{X}_{-\omega_k}^T(\sigma_1)$, $C=\widetilde{X}_{\omega_l}^T(\tau_2)$ and $D=\widetilde{X}_{-\omega_l}^T(\sigma_2)$ and the last equality holds by Theorem 2.3.2 from Brillinger (2001) and the fact that $(A)=(B)=(C)=(D)=0$.

$$
\begin{aligned}
(A,B,C,D) &= 2\pi T^{-2}(Tf_{\omega_k,-\omega_k,\omega_l}(\tau_1,\sigma_1,\tau_2,\sigma_2)+O(1)\\
&= \frac{2\pi}{T}f_{\omega_k,-\omega_k,\omega_l}(\tau_1,\sigma_1,\tau_2,\sigma_2)+O(T^{-2})\\
(A,C)(B,D) &= T^{-2}O(1)O(1)=O(T^{-2})\\
(A,D)(B,C) &= T^{-2}(T\delta_{k,l}f_{\omega_k}(\tau_1,\sigma_2)+O(1))(T\delta_{k,l}f_{\omega_l}(\tau_2,\sigma_1)+O(1))\\
&= \delta_{k,l}f_{\omega_k}(\tau_1,\sigma_2)f_{\omega_l}(\tau_2,\sigma_1)+\delta_{k,l}O(T^{-1})+O(T^{-2})
\end{aligned}
$$

Therefore if follows

$$
\begin{aligned}
\Sigma_{12}=\Sigma_{21} &= \frac{1}{T}\sum_{k=1}^{\lfloor T/2\rfloor}f_{\omega_k}(\tau_1,\sigma_2)f_{\omega_k}(\tau_2,\sigma_1)+\frac{1}{T}\sum_{k=1}^{\lfloor T/2\rfloor}f_{\omega_k}(\sigma_1,\tau_2)f_{\omega_k}(\sigma_2,\tau_1)\\
&\quad+\frac{1}{T}\sum_{k=1}^{\lfloor T/2\rfloor}f_{\omega_k}(\tau_1,\tau_2)f_{\omega_k}(\sigma_2,\sigma_1)+\frac{1}{T}\sum_{k=1}^{\lfloor T/2\rfloor}f_{\omega_k}(\sigma_1,\sigma_2)f_{\omega_k}(\tau_2,\tau_1)\\
&\quad+\frac{2\pi}{T^2}\sum_{k=1}^{\lfloor T/2\rfloor}\sum_{l=1}^{\lfloor T/2\rfloor}f_{\omega_k,-\omega_k,\omega_l}(\tau_1,\sigma_1,\tau_2,\sigma_2)+\frac{2\pi}{T^2}\sum_{k=1}^{\lfloor T/2\rfloor}\sum_{l=1}^{\lfloor T/2\rfloor}f_{\omega_k,-\omega_k,\omega_l}(\sigma_1,\tau_1,\sigma_2,\tau_2)\\
&\quad+\frac{2\pi}{T^2}\sum_{k=1}^{\lfloor T/2\rfloor}\sum_{l=1}^{\lfloor T/2\rfloor}f_{\omega_k,-\omega_k,\omega_l}(\tau_1,\sigma_1,\sigma_2,\tau_2)+\frac{2\pi}{T^2}\sum_{k=1}^{\lfloor T/2\rfloor}\sum_{l=1}^{\lfloor T/2\rfloor}f_{\omega_k,-\omega_k,\omega_l}(\sigma_1,\tau_1,\tau_2,\sigma_2)\\
&\to \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(f_\omega(\tau_1,\sigma_2)f_\omega(\tau_2,\sigma_1)+f_\omega(\tau_1,\tau_2)f_\omega(\sigma_2,\sigma_1)\right)d\omega\\
&\quad+\frac{1}{2\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}f_{\omega_1,-\omega_1,\omega_2}(\tau_1,\sigma_1,\tau_2,\sigma_2)\omega_1 d\omega_2 \quad \text{as } T\to\infty.
\end{aligned}
$$

Similarly for $i=1,2$, we have

$$
\begin{aligned}
\Sigma_{ii} \to& \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(f_\omega^2(\tau_i,\sigma_i)+f_\omega(\tau_i,\tau_i)f_\omega(\sigma_i,\sigma_i)\right)d\omega\\
&+\frac{1}{2\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}f_{\omega_1,-\omega_1,\omega_2}(\tau_i,\sigma_i,\tau_i,\sigma_i)\omega_1 d\omega_2 \quad \text{as } T\to\infty.
\end{aligned}
$$

Next consider:

$$\begin{aligned}
\Sigma_{34} = \Sigma_{43} =& T\mathrm{Cov}\big(S_{T,2}(\tau_1,\sigma_1), S_{T,2}(\tau_2,\sigma_2)\big)\\
=& \frac{4}{T}\mathrm{Cov}\left(\sum_{k=1}^{\lfloor T/2\rfloor}\big(p_{\omega_k}^{(T)}(\tau_1,\sigma_1)p_{\omega_{k-1}}^{(T)}(\sigma_1,\tau_1)\big), \sum_{k=1}^{\lfloor T/2\rfloor}\big(p_{\omega_k}^{(T)}(\tau_2,\sigma_2)p_{\omega_{k-1}}^{(T)}(\sigma_2,\tau_2)\big)\right)\\
=& \frac{4}{T}\sum_{k=1}^{\lfloor T/2\rfloor}\sum_{l=1}^{\lfloor T/2\rfloor}\mathrm{Cov}\big(p_{\omega_k}^{(T)}(\tau_1,\sigma_1)p_{\omega_{k-1}}^{(T)}(\sigma_1,\tau_1), p_{\omega_l}^{(T)}(\tau_2,\sigma_2)p_{\omega_{l-1}}^{(T)}(\sigma_2,\tau_2)\big)
\end{aligned}$$

We calculate

$$\begin{aligned}
C_{kl} :=& \mathrm{Cov}\big(p_{\omega_k}^{(T)}(\tau_1,\sigma_2)p_{\omega_{k-1}}^{(T)}(\sigma_1,\tau_1), p_{\omega_l}^{(T)}(\tau_2,\sigma_2)p_{\omega_{l-1}}^{(T)}(\sigma_2,\tau_2)\big)\\
=& \mathrm{cum}\big(\tilde{X}_{\omega_k}^{(T)}(\tau_1)\tilde{X}_{-\omega_k}^{(T)}(\sigma_1)\tilde{X}_{\omega_{k-1}}^{(T)}(\sigma_1)\tilde{X}_{-\omega_{k-1}}^{(T)}(\tau_1), \tilde{X}_{\omega_l}^{(T)}(\tau_2)\tilde{X}_{-\omega_l}^{(T)}(\sigma_2)\tilde{X}_{\omega_{l-1}}^{(T)}(\sigma_2)\tilde{X}_{-\omega_{l-1}}^{(T)}(\tau_2)\big)\\
=& \mathrm{cum}(ABCD, EFGH)
\end{aligned}$$

We use Theorem 2.3.2 from Brillinger (2001) to calculate the cumulant. As argued in the proof of Proposition 5.1 we only need to look at the partitions ν with $p = 1,2,3,4$.
p=1:

$$(A,B,C,D,E,F,G,H) = \frac{(2\pi)^3}{T^3} f_{\omega_k,-\omega_k,\omega_{k-1},-\omega_{k-1},\omega_l,-\omega_l,\omega_{l-1}}(\tau_1,\sigma_1,\sigma_1,\tau_1,\tau_2,\sigma_2,\sigma_2,\tau_2)+O(T^{-4})$$

p=2:

$$\begin{aligned}
(A,B)(C,D,E,F,G,H) =& \frac{(2\pi)^2}{T^4}(Tf_{\omega_k}(\tau_1,\sigma_1)+O(1))(Tf_{\omega_{k-1},-\omega_{k-1},\omega_l,-\omega_l,\omega_{l-1}}(\sigma_1,\tau_1,\tau_2,\sigma_2,\sigma_2,\tau_2)+O(1))\\
=& \frac{(2\pi)^2}{T^2} f_{\omega_k}(\tau_1,\sigma_1) f_{\omega_{k-1},-\omega_{k-1},\omega_l,-\omega_l,\omega_{l-1}}(\sigma_1,\tau_1,\tau_2,\sigma_2,\sigma_2,\tau_2)+O(T^{-3})
\end{aligned}$$

Similarly all $2+6$ partitions are of order $O(T^{-2})$

$$(A,B,C)(D,E,F,G,H) = \frac{(2\pi)^2}{T^4} O(1)O(1) = O(T^{-4}).$$

Similarly all $3+5$ partitions are of order$O(T^{-4})$.

$$(A,B,E,F)(C,D,G,H) = \frac{(2\pi)^2}{T^4}(Tf_{\omega_k,-\omega_k,\omega_l}(\tau_1,\sigma_1,\tau_2,\sigma_2)+O(1))(Tf_{\omega_{k-1},-\omega_{k-1},\omega_{l-1}}(\sigma_1,\tau_1,\sigma_2,\tau_2)+O(1))$$

Similarly all $4+4$ partitions are of order $O(T^{-2})$.
p=3: The partitions with significant contributions are:

$$\begin{aligned}
(A,B)(C,F)(D,E,G,H) =& T^{-4}(Tf_{\omega_k}(\tau_1,\sigma_1)+O(1))(T\delta_{k-1,l}f_{\omega_{k-1}}(\sigma_1,\sigma_2)+O(1))\\
& (T\delta_{k-1,l}(2\pi)f_{-\omega_{k-1},\omega_l,\omega_{l-1}}(\tau_1,\tau_2,\sigma_2,\tau_2)+O(1))\\
=& \delta_{k-1,l}O(T^{-1})+O(T^{-3}).
\end{aligned}$$

$$\begin{aligned}
(A,B)(C,H)(D,E,F,G) =& \delta_{k,l}O(T^{-1}) + O(T^{-3}).\\
(A,B)(D,E)(C,F,G,H) =& \delta_{k-1,l}O(T^{-1}) + O(T^{-3}).\\
(A,B)(D,G)(C,E,F,H) =& \delta_{k,l}O(T^{-1}) + O(T^{-3}).\\
(A,B)(E,F)(C,D,G,H) =& \frac{1}{T^4}(Tf_{\omega_k}(\tau_1,\sigma_1) + O(1))(Tf_{\omega_l}(\tau_2,\sigma_2) + O(1))(2\pi T f_{\omega_{k-1},-\omega_{k-1},\omega_{l-1}}(\sigma_1,\tau_1,\sigma_2,\tau_2) + O\\
=& \frac{2\pi}{T} f_{\omega_k}(\tau_1,\sigma_1) f_{\omega_l}(\tau_2,\sigma_2) f_{\omega_{k-1},-\omega_{k-1},\omega_{l-1}}(\sigma_1,\tau_1,\sigma_2,\tau_2) + O(T^{-2})\\
(A,B)(G,H)(C,D,E,F) =& \frac{2\pi}{T} f_{\omega_k}(\tau_1,\sigma_1) f_{\omega_{l-1}}(\sigma_2,\tau_2) f_{\omega_{k-1},-\omega_{k-1},\omega_l}(\sigma_1,\tau_1,\tau_2,\sigma_2) + O(T^{-2})\\
(C,D)(E,F)(A,B,G,H) =& \frac{2\pi}{T} f_{\omega_{k-1}}(\sigma_1,\tau_1) f_{\omega_l}(\tau_2,\sigma_2) f_{\omega_k,-\omega_k,\omega_{l-1}}(\tau_1,\sigma_1,\sigma_2,\tau_2) + O(T^{-2})\\
(C,D)(G,H)(A,B,E,F) =& \frac{2\pi}{T} f_{\omega_{k-1}}(\sigma_1,\tau_1) f_{\omega_{l-1}}(\sigma_2,\tau_2) f_{\omega_k,-\omega_k,\omega_l}(\tau_1,\sigma_1,\tau_2,\sigma_2) + O(T^{-2}).
\end{aligned}$$

All other terms are of order $O(T^{-3})$.
p=4: The partitions with significant contributions are

$$\begin{aligned}
(A,B)(C,F)(D,E)(G,H) =& T^{-4}(Tf_{\omega_k}(\tau_1,\sigma_1) + O(1))(T\delta_{k-1,l} f_{\omega_{k-1}}(\sigma_1,\sigma_2) + O(1))\\
& (T\delta_{k-1,l} f_{\omega_l}(\tau_2,\tau_1) + O(1))(Tf_{\omega_{l-1}}(\sigma_2,\tau_2) + O(1))\\
=& \delta_{k-1,l} f_{\omega_k}(\tau_1,\sigma_1) f_{\omega_{k-1}}(\sigma_1,\sigma_2) f_{\omega_l}(\tau_2,\tau_1) f_{\omega_{l-1}}(\sigma_2,\tau_2) + O(T^{-2})\\
(A,B)(C,H)(D,G)(E,F) =& T^{-4}(Tf_{\omega_k}(\tau_1,\sigma_1) + O(1))(T\delta_{k,l} f_{\omega_{k-1}}(\sigma_1,\tau_2) + O(1))\\
& (T\delta_{k,l} f_{\omega_{l-1}}(\sigma_2,\tau_1) + O(1))(Tf_{\omega_l}(\tau_2,\sigma_2) + O(1))\\
=& \delta_{k-1,l} f_{\omega_k}(\tau_1,\sigma_1) f_{\omega_{k-1}}(\sigma_1,\tau_2) f_{\omega_{l-1}}(\sigma_2,\tau_1) f_{\omega_l}(\tau_2,\sigma_2) + O(T^{-2})\\
(A,F)(B,E)(C,D)(G,H) =& \delta_{k,l} f_{\omega_k}(\tau_1,\sigma_2) f_{\omega_l}(\tau_2,\sigma_1) f_{\omega_{k-1}}(\sigma_1,\tau_1) f_{\omega_{l-1}}(\sigma_2,\tau_2) + O(T^{-2})\\
(A,H)(B,G)(C,D)(E,F) =& \delta_{k,l-1} f_{\omega_k}(\tau_1,\tau_2) f_{\omega_{l-1}}(\sigma_2,\sigma_1) f_{\omega_{k-1}}(\sigma_1,\tau_1) f_{\omega_l}(\tau_2,\sigma_2) + O(T^{-2})\\
(A,F)(B,E)(C,H)(D,G) =& \delta_{k,l} f_{\omega_k}(\tau_1,\sigma_2) f_{\omega_l}(\tau_2,\sigma_1) f_{\omega_{k-1}}(\sigma_1,\tau_2) f_{\omega_{l-1}}(\sigma_2,\tau_1) + O(T^{-2})
\end{aligned}$$

Contributions of all the other partitions with $p = 4$ are $\leq O(T^{-2})$. Summing up all these terms we get

$$\begin{aligned}
\Sigma_{34} = \Sigma_{43} =& \frac{4}{T} \sum_{k=1}^{\lfloor T/2 \rfloor} f_{\omega_k}(\tau_1,\sigma_1) f_{\omega_{k-1}}(\sigma_1,\sigma_2) f_{\omega_{k-1}}(\tau_2,\tau_1) f_{\omega_{k-2}}(\sigma_2,\tau_2)\\
&+ \frac{4}{T} \sum_{k=1}^{\lfloor T/2 \rfloor} f_{\omega_k}(\tau_1,\sigma_1) f_{\omega_{k-1}}(\sigma_1,\tau_2) f_{\omega_{k-2}}(\sigma_2,\tau_1) f_{\omega_{k-1}}(\tau_2,\sigma_2)\\
&+ \frac{4}{T} \sum_{k=1}^{\lfloor T/2 \rfloor} f_{\omega_k}(\tau_1,\sigma_2) f_{\omega_k}(\tau_2,\sigma_1) f_{\omega_{k-1}}(\sigma_1,\tau_1) f_{\omega_{k-1}}(\sigma_2,\tau_2)\\
&+ \frac{4}{T} \sum_{k=1}^{\lfloor T/2 \rfloor} f_{\omega_k}(\tau_1,\tau_2) f_{\omega_k}(\sigma_2,\sigma_1) f_{\omega_{k-1}}(\sigma_1,\tau_1) f_{\omega_{k+1}}(\tau_2,\sigma_2)\\
&+ \frac{4}{T} \sum_{k=1}^{\lfloor T/2 \rfloor} f_{\omega_k}(\tau_1,\sigma_2) f_{\omega_k}(\tau_2,\sigma_1) f_{\omega_{k-1}}(\sigma_1,\tau_2) f_{\omega_{k-1}}(\sigma_2,\tau_1)
\end{aligned}$$

$$
\begin{aligned}
&+\frac{4(2\pi)}{T^2}\sum_{k=1}^{\lfloor T/2\rfloor}\sum_{l=1}^{\lfloor T/2\rfloor} f_{\omega_k}(\tau_1,\sigma_1)f_{\omega_l}(\tau_2,\sigma_2)f_{\omega_{k-1},-\omega_{k-1},\omega_{l-1}}(\sigma_1,\tau_1,\sigma_2,\tau_2)\\
&+\frac{4(2\pi)}{T^2}\sum_{k=1}^{\lfloor T/2\rfloor}\sum_{l=1}^{\lfloor T/2\rfloor} f_{\omega_k}(\tau_1,\sigma_1)f_{\omega_{l-1}}(\sigma_2,\tau_2)f_{\omega_{k-1},-\omega_{k-1},\omega_l}(\sigma_1,\tau_1,\tau_2,\sigma_2)\\
&+\frac{4(2\pi)}{T^2}\sum_{k=1}^{\lfloor T/2\rfloor}\sum_{l=1}^{\lfloor T/2\rfloor} f_{\omega_{k-1}}(\sigma_1,\tau_1)f_{\omega_l}(\tau_2,\sigma_2)f_{\omega_k,-\omega_k,\omega_{l-1}}(\tau_1,\sigma_1,\sigma_2,\tau_2)\\
&+\frac{4(2\pi)}{T^2}\sum_{k=1}^{\lfloor T/2\rfloor}\sum_{l=1}^{\lfloor T/2\rfloor} f_{\omega_{k-1}}(\sigma_1,\tau_1)f_{\omega_{l-1}}(\sigma_2,\tau_2)f_{\omega_k,-\omega_k,\omega_l}(\tau_1,\sigma_1,\tau_2,\sigma_2)+o(1)\\
\to&\frac{2}{\pi}\int_{-\pi}^{\pi} f_\omega(\tau_1,\sigma_1)f_\omega(\sigma_1,\tau_2)f_\omega(\sigma_2,\tau_1)f_\omega(\tau_2,\sigma_2)d\omega\\
&+\frac{2}{\pi}\int_{-\pi}^{\pi} f_\omega(\tau_1,\sigma_1)f_\omega(\tau_2,\tau_1)f_\omega(\sigma_1,\sigma_2)f_\omega(\sigma_2,\tau_2)d\omega\\
&+\frac{1}{\pi}\int_{-\pi}^{\pi} f_\omega(\tau_1,\sigma_2)f_\omega(\tau_2,\sigma_1)f_\omega(\sigma_1,\tau_2)f_\omega(\sigma_2,\tau_1)d\omega\\
&+\frac{2}{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} f_{\omega_1}(\tau_1,\sigma_1)f_{\omega_2}(\tau_2,\sigma_2)f_{\omega_1,-\omega_1,\omega_2}(\sigma_1,\tau_1,\sigma_2,\tau_2)d\omega_1 d\omega_2
\end{aligned}
$$

as $T\to\infty$. Similarly for $i=3,4$,

$$
\begin{aligned}
\Sigma_{ii}\to&\frac{2}{\pi}\int_{-\pi}^{\pi}\left|f_\omega(\tau_{i-2},\sigma_{i-2})\right|^4 d\omega+\frac{2}{\pi}\int_{-\pi}^{\pi}\left|f_\omega(\tau_{i-2},\sigma_{i-2})\right|^2 f_\omega(\tau_{i-2},\tau_{i-2})f_\omega(\sigma_{i-2},\sigma_{i-2})d\omega\\
&+\frac{2}{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} f_{\omega_1}(\tau_{i-2},\sigma_{i-2})f_{\omega_2}(\tau_{i-2},\sigma_{i-2})f_{\omega_1,-\omega_1,\omega_2}(\sigma_{i-2},\tau_{i-2},\sigma_{i-2},\tau_{i-2})d\omega_1 d\omega_2,
\end{aligned}
$$

as $T\to\infty$. Finally we calculate

$$
\begin{aligned}
\Sigma_{23}=\Sigma_{32}=&T\mathrm{Cov}\big(S_{T,1}(\tau_2,\sigma_2),S_{T2}(\tau_1,\sigma_1)\big)\\
=&\frac{2}{T}\sum_{k=1}^{\lfloor T/2\rfloor}\sum_{l=1}^{\lfloor T/2\rfloor}\mathrm{Cov}\Big(p_{\omega_k}^{(T)}(\tau_1,\sigma_1)p_{\omega_{k-1}}^{(T)}(\sigma_1,\tau_1),p_{\omega_l}^{(T)}(\tau_2,\sigma_2)+p_{\omega_l}^{(T)}(\sigma_2,\tau_2)\Big).
\end{aligned}
$$

As earlier we consider each of the terms in the summation separately and calculate

$$
\mathrm{Cov}\Big(p_{\omega_k}^{(T)}(\tau_1,\sigma_1)p_{\omega_{k-1}}^{(T)}(\sigma_1,\tau_1),p_{\omega_l}^{(T)}(\tau_2,\sigma_2)\Big)=\mathrm{cum}(ABCD,EF)
$$

We employ Theorem 2.3.2 from Brillinger (2001) and only calculate cumulants for partitions with size $p=1,2,3$.
For $p=1$: $(A,B,C,D,E,F)=O(T^{-2})$.
For $p=2$: all 3+3 partitions $=O(T^{-3})$ and significant 2+4 partitions are as follows:

$$
\begin{aligned}
(A,B)(C,D,E,F)=&(2\pi)T^{-3}(Tf_{\omega_k}(\tau_1,\sigma_1)+O(1))(Tf_{\omega_{k-1},-\omega_{k-1},\omega_l}(\sigma_1,\tau_1,\tau_2,\sigma_2)+O(1))\\
=&\frac{2\pi}{T}f_{\omega_k}(\tau_1,\sigma_1)f_{\omega_{k-1},-\omega_{k-1},\omega_l}(\sigma_1,\tau_1,\tau_2,\sigma_2)+O(T^{-2})\\
(A,F)(B,C,D,E)=&(2\pi)T^{-3}(T\delta_{k,l}f_{\omega_k}(\tau_1,\sigma_2)+O(1))(T\delta_{k,l}f_{-\omega_k,\omega_{k-1},-\omega_{k-l}}(\sigma_1,\sigma_1,\tau_1,\tau_2)+O(1))
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{2\pi}{T}\delta_{k,l} f_{\omega_k}(\tau_1,\sigma_2) f_{-\omega_k,\omega_{k-1},-\omega_{k-l}}(\sigma_1,\sigma_1,\tau_1,\tau_2) + O(T^{-2}) \\
(B,E)(A,C,D,F) &= (2\pi)T^{-3}(T\delta_{k,l} f_{-\omega_k}(\sigma_1,\tau_2) + O(1))(T\delta_{k,l} f_{\omega_k,\omega_{k-1},-\omega_{k-l}}(\tau_1,\sigma_1,\tau_1,\sigma_2) + O(1)) \\
&= \frac{2\pi}{T}\delta_{k,l} f_{-\omega_k}(\sigma_1,\tau_2) f_{\omega_k,\omega_{k-1},-\omega_{k-l}}(\tau_1,\sigma_1,\tau_1,\sigma_2) + O(T^{-2}) \\
(C,D)(A,B,E,F) &= \frac{2\pi}{T} f_{\omega_{k-1}}(\sigma_1,\tau_1) f_{\omega_k,-\omega_k,\omega_l}(\tau_1,\sigma_1,\tau_2,\sigma_2) + O(T^{-2}) \\
(C,F)(A,B,D,F) &= \frac{2\pi}{T}\delta_{k-1,l} f_{\omega_{k-1}}(\sigma_1,\sigma_2) f_{\omega_k,-\omega_k,\omega_l}(\tau_1,\sigma_1,\tau_2,\tau_1) + O(T^{-2}) \\
(D,E)(A,B,C,F) &= \delta_{k-1,l} O(T^{-1}) + O(T^{-2}).
\end{aligned}
$$

Other 2+4 partitions are $O(T^{-3})$.
For p=3:

$$
\begin{aligned}
(A,B)(C,F)(D,E) &= T^{-3}(T f_{\omega_k}(\tau_1,\sigma_1) + O(1))(T\delta_{k-1,l} f_{\omega_{k-1}}(\sigma_1,\sigma_2) + O(1))(T\delta_{k-1,l} f_{\omega_l}(\tau_2,\tau_1) + O(1)) \\
&= \delta_{k-1,l} f_{\omega_k}(\tau_1,\sigma_1) f_{\omega_{k-1}}(\sigma_1,\sigma_2) f_{\omega_l}(\tau_2,\tau_1) + \delta_{k-1,l} O(T^{-1}) + O(T^{-2}) \\
(A,F)(B,E)(C,D) &= \delta_{k,l} f_{\omega_k}(\tau_1,\sigma_2) f_{\omega_l}(\tau_2,\sigma_1) f_{\omega_{k-1}}(\sigma_1,\tau_1) + \delta_{k,l} O(T^{-1}) + O(T^{-2})
\end{aligned}
$$

and other partitions are $O(T^{-2})$. Summing up all the cumulants we obtain

$$
\begin{aligned}
\Sigma_{23} = \Sigma_{32} &= \frac{2}{T}\sum_{k=1}^{\lfloor T/2\rfloor}\big(f_{\omega_k}(\tau_1,\sigma_1) f_{\omega_{k-1}}(\sigma_1,\sigma_2) f_{\omega_{k-1}}(\tau_2,\tau_1) + f_{\omega_k}(\tau_1,\sigma_1) f_{\omega_{k-1}}(\sigma_1,\tau_2) f_{\omega_{k-1}}(\sigma_2,\tau_1)\big) \\
&\quad + \frac{2}{T}\sum_{k=1}^{\lfloor T/2\rfloor}\big(f_{\omega_k}(\tau_1,\sigma_2) f_{\omega_k}(\tau_2,\sigma_1) f_{\omega_{k-1}}(\sigma_1,\tau_1) + f_{\omega_k}(\tau_1,\tau_2) f_{\omega_k}(\sigma_2,\sigma_1) f_{\omega_{k-1}}(\sigma_1,\tau_1)\big) \\
&\quad + \frac{4\pi}{T^2}\sum_{k=1}^{\lfloor T/2\rfloor}\sum_{l=1}^{\lfloor T/2\rfloor} f_{\omega_k}(\tau_1,\sigma_1) f_{\omega_{k-1},-\omega_{k-1}\omega_l}(\sigma_1,\tau_1,\tau_2,\sigma_2) \\
&\quad + \frac{4\pi}{T^2}\sum_{k=1}^{\lfloor T/2\rfloor}\sum_{l=1}^{\lfloor T/2\rfloor} f_{\omega_k}(\tau_1,\sigma_1) f_{\omega_{k-1},-\omega_{k-1}\omega_l}(\sigma_1,\tau_1,\sigma_2,\tau_2) \\
&\quad + \frac{4\pi}{T^2}\sum_{k=1}^{\lfloor T/2\rfloor}\sum_{l=1}^{\lfloor T/2\rfloor} f_{\omega_{k-1}}(\sigma_1,\tau_1) f_{\omega_k,-\omega_k\omega_l}(\tau_1,\sigma_1,\tau_2,\sigma_2) \\
&\quad + \frac{4\pi}{T^2}\sum_{k=1}^{\lfloor T/2\rfloor}\sum_{l=1}^{\lfloor T/2\rfloor} f_{\omega_{k-1}}(\sigma_1,\tau_1) f_{\omega_k,-\omega_k\omega_l}(\tau_1,\sigma_1,\sigma_2,\tau_2) + o(1) \\
&\to \frac{1}{\pi}\int_{-\pi}^{\pi} f_\omega(\tau_1,\sigma_1) f_\omega(\sigma_1,\sigma_2) f_\omega(\tau_2,\tau_1)d\omega + \frac{1}{\pi}\int_{-\pi}^{\pi} f_\omega(\tau_1,\sigma_1) f_\omega(\sigma_1,\tau_2) f_\omega(\sigma_2,\tau_1)d\omega \\
&\quad + \frac{1}{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} f_{\omega_1}(\tau_1,\sigma_1) f_{\omega_1,-\omega_1,\omega_2}(\sigma_1,\tau_1,\sigma_2,\tau_2)d\omega_1 d\omega_2 \quad \text{as } T\to\infty.
\end{aligned}
$$

The other terms are obtained similarly. □

Proposition 6.1. *Under assumption* (ii) *of Theorem 3.2, for almost every* $(\tau,\sigma)\in[0,1]^2$,

$$
\mathbb{E}(S_{T2}(\tau,\sigma) - S_{T1}(\tau,\sigma)\overline{S_{T1}(\tau,\sigma)}) = \frac{1}{2\pi}\int_{-\pi}^{\pi}|f_\omega(\tau,\sigma)|^2 d\omega - \frac{1}{4\pi^2}\left|\int_{-\pi}^{\pi} f_\omega(\tau,\sigma)d\omega\right|^2 + O(T^{-1}).
$$

Proof. The result follows from Lemma 6.1, Lemma 6.2 and the fact $\mathbb{E}\left(S_{T1}(\tau,\sigma)\overline{S_{T1}(\tau,\sigma)}\right) = \mathrm{Var}(S_{T1}(\tau,\sigma)) + \mathbb{E}^2(S_{T,1}(\tau,\sigma))$. □

Proposition 6.2. *Under assumption (ii) of Theorem 3.2, for almost every* $(\tau_1,\sigma_1,\tau_2,\sigma_2) \in [0,1]^4$,

$$\begin{aligned}
v^2(\tau_1,\sigma_1,\tau_2,\sigma_2) &= \lim_{T\to\infty} \mathrm{Cov}(I_T(\tau_1,\sigma_1), I_T(\tau_2,\sigma_2)) \qquad (6.2)\\
&= \frac{2}{\pi}\int_{-\pi}^{\pi} f_\omega(\tau_1,\sigma_1) f_\omega(\sigma_1,\tau_2) f_\omega(\sigma_2,\tau_1) f_\omega(\tau_2,\sigma_2) d\omega\\
&+ \frac{2}{\pi}\int_{-\pi}^{\pi} f_\omega(\tau_1,\sigma_1) f_\omega(\tau_2,\tau_1) f_\omega(\sigma_1,\sigma_2) f_\omega(\sigma_2,\tau_2) d\omega\\
&+ \frac{1}{\pi}\int_{-\pi}^{\pi} f_\omega(\tau_1,\sigma_2) f_\omega(\tau_2,\sigma_1) f_\omega(\sigma_1,\tau_2) f_\omega(\sigma_2,\tau_1) d\omega\\
&+ \frac{2}{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} f_{\omega_1}(\tau_1,\sigma_1) f_{\omega_2}(\tau_2,\sigma_2) f_{\omega_1,-\omega_1,\omega_2}(\sigma_1,\tau_1,\sigma_2,\tau_2) d\omega_1 d\omega_2\\
&- \frac{1}{\pi^2}\int_{-\pi}^{\pi} f_\omega(\tau_1,\sigma_1) d\omega \int_{-\pi}^{\pi} f_\omega(\tau_2,\sigma_2) f_\omega(\sigma_2,\sigma_1) f_\omega(\tau_1,\tau_2) d\omega\\
&- \frac{1}{\pi^2}\int_{-\pi}^{\pi} f_\omega(\tau_1,\sigma_1) d\omega \int_{-\pi}^{\pi} f_\omega(\tau_2,\sigma_2) f_\omega(\sigma_2,\tau_1) f_\omega(\sigma_1,\tau_2) d\omega\\
&- \frac{1}{\pi^2}\int_{-\pi}^{\pi} f_\omega(\tau_1,\sigma_1) d\omega \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} f_{\omega_1}(\tau_2,\sigma_2) f_{\omega_1,-\omega_1,\omega_2}(\sigma_2,\tau_2,\sigma_1,\tau_1) d\omega_1 d\omega_2\\
&- \frac{1}{\pi^2}\int_{-\pi}^{\pi} f_\omega(\tau_2,\sigma_2) d\omega \int_{-\pi}^{\pi} f_\omega(\tau_1,\sigma_1) f_\omega(\sigma_1,\sigma_2) f_\omega(\tau_2,\tau_1) d\omega\\
&- \frac{1}{\pi^2}\int_{-\pi}^{\pi} f_\omega(\tau_2,\sigma_2) d\omega \int_{-\pi}^{\pi} f_\omega(\tau_1,\sigma_1) f_\omega(\sigma_1,\tau_2) f_\omega(\sigma_2,\tau_1) d\omega\\
&- \frac{1}{\pi^2}\int_{-\pi}^{\pi} f_\omega(\tau_2,\sigma_2) d\omega \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} f_{\omega_1}(\tau_1,\sigma_1) f_{\omega_1,-\omega_1,\omega_2}(\sigma_1,\tau_1,\sigma_2,\tau_2) d\omega_1 d\omega_2\\
&+ \frac{1}{2\pi^3}\int_{-\pi}^{\pi} f_\omega(\tau_1,\sigma_1) d\omega \int_{-\pi}^{\pi} f_\omega(\tau_2,\sigma_2) d\omega \int_{-\pi}^{\pi} f_\omega(\tau_1,\sigma_2) f_\omega(\tau_2,\sigma_1) d\omega\\
&+ \frac{1}{2\pi^3}\int_{-\pi}^{\pi} f_\omega(\tau_1,\sigma_1) d\omega \int_{-\pi}^{\pi} f_\omega(\tau_2,\sigma_2) d\omega \int_{-\pi}^{\pi} f_\omega(\tau_1,\tau_2) f_\omega(\sigma_2,\sigma_1) d\omega\\
&+ \frac{1}{2\pi^3}\int_{-\pi}^{\pi} f_\omega(\tau_1,\sigma_1) d\omega \int_{-\pi}^{\pi} f_\omega(\tau_2,\sigma_2) d\omega \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} f_{\omega_1,-\omega_1,\omega_2}(\tau_1,\sigma_1,\tau_2,\sigma_2)\omega_1 d\omega_2
\end{aligned}$$

Proof. We have proved in Section 5.1 the vector

$$\sqrt{T}\begin{pmatrix} S_{T,1}(\tau_1,\sigma_1) - \mathbb{E}(S_{T,1}(\tau_1,\sigma_1))\\ S_{T,1}(\tau_2,\sigma_2) - \mathbb{E}(S_{T,1}(\tau_2,\sigma_2))\\ S_{T,2}(\tau_1,\sigma_1) - \mathbb{E}(S_{T,2}(\tau_1,\sigma_1))\\ S_{T,2}(\tau_2,\sigma_2) - \mathbb{E}(S_{T,2}(\tau_2,\sigma_2)) \end{pmatrix}$$

converges in distribution to a normal distribution. To obtain the covariance kernel of $I_T(\tau,\sigma)$, we use delta-method on this vector with $g(x_1,x_2,x_3,x_4) := (x_3 - x_1^2, x_4 - x_2^2)$. Using

Lemma 6.1 it follows that

$$\sqrt{T}\begin{pmatrix} S_{T2}(\tau_1,\sigma_1) - S_{T1}(\tau_1,\sigma_1)\overline{S_{T1}(\tau_1,\sigma_1)} - \frac{1}{2\pi}\int_{-\pi}^{\pi}|f_\omega(\tau_1,\sigma_1)|^2 d\omega + \frac{1}{4\pi^2}\left|\int_{-\pi}^{\pi} f_\omega(\tau_1,\sigma_1)d\omega\right|^2 \\ S_{T2}(\tau_2,\sigma_2) - S_{T1}(\tau_2,\sigma_2)\overline{S_{T1}(\tau_2,\sigma_2)} - \frac{1}{2\pi}\int_{-\pi}^{\pi}|f_\omega(\tau_2,\sigma_2)|^2 d\omega + \frac{1}{4\pi^2}\left|\int_{-\pi}^{\pi} f_\omega(\tau_2,\sigma_2)d\omega\right|^2 \end{pmatrix} \xrightarrow{d} N(0,\widetilde{\Sigma}).$$

where

$$\begin{aligned}\widetilde{\Sigma}_{ij} =& \Sigma_{(i+2)(j+2)} - 2\frac{1}{2\pi}\int_{-\pi}^{\pi} f_\omega(\tau_i,\sigma_i)d\omega\Sigma_{i(j+2)} - 2\frac{1}{2\pi}\int_{-\pi}^{\pi} f_\omega(\tau_j,\sigma_j)d\omega\Sigma_{(i+2)j} \\ &+ 4\frac{1}{4\pi^2}\int_{-\pi}^{\pi} f_\omega(\tau_i,\sigma_i)d\omega\int_{-\pi}^{\pi} f_\omega(\tau_j,\sigma_j)d\omega\Sigma_{ij}\end{aligned}$$

for $i=1,2$; $j=1,2$ and Σ_{ij} are as in Lemma 6.2. Substituting the values of Σ_{ij} we obtain (6.2). □

6.2 The order of $C_0(\nu)$ in Proposition 5.1

Using the notations of Proposition 5.1 we write

$$C_0(\nu) = \frac{2^l}{T^{l/2}}\sum_{k_1=1}^{\lfloor T/2\rfloor}\cdots\sum_{k_l=1}^{\lfloor T/2\rfloor}\text{cum}\left(Z_{ij} : ij\in\nu_1\right)\dots\text{cum}\left(Z_{ij} : ij\in\nu_p\right), \tag{6.3}$$

where $Z_{i1} := \widetilde{X}^{(T)}_{\omega_i}(\tau)$, $Z_{i2} := \widetilde{X}^{(T)}_{-\omega_i}(\sigma)$, $Z_{i3} := \widetilde{X}^{(T)}_{\omega_{i-1}}(\sigma)$ and $Z_{i4} := \widetilde{X}^{(T)}_{-\omega_{i-1}}(\tau)$ and $\nu = \nu_1\cup\nu_2\cup\cdots\cup\nu_p$ is any indecomposable partition of the table

$$\begin{matrix} (k_1,1) & (k_1,2) & (k_1,3) & (k_1,4) \\ (k_2,1) & (k_2,2) & (k_2,3) & (k_2,4) \\ \vdots & \vdots & \vdots & \vdots \\ (k_l,1) & (k_l,2) & (k_l,3) & (k_l,4) \end{matrix}. \tag{6.4}$$

To calculate these cumulants we will use Theorem B.2 from Panaretos and Tavakoli (2013b), which says

$$\text{cum}\left(\widetilde{X}^{(T)}_{\omega_1}(\tau_1),\dots,\widetilde{X}^{(T)}_{\omega_k}(\tau_k)\right) = \frac{(2\pi)^{k/2-1}}{T^{k/2}}\Delta^{(T)}(\omega_1+\dots,\omega_k)f_{\omega_1,\dots,\omega_{k-1}}(\tau_1,\dots,\tau_k) + \epsilon_T. \tag{6.5}$$

In the above equation $f_{\omega_1,\dots,\omega_{k-1}} = O(1)$, $\epsilon_T = O(1)$ uniformly over ω and the equality is interpreted in $L^2([0,1]^k)$.
For $\omega = 2\pi k/T$, $k\in\mathbb{Z}$, the function $\Delta^{(T)}(\omega) = T$ if $k = 0$ $(mod\ T)$ and 0 otherwise.
Note that if for some ν, we have $p > 2l$ then there is at least one $|\nu_m| = 1$ for some $1\le m\le p$. In that case $\text{cum}(Z_{ij} : ij\in\nu_m) = 0$ and therefore $C(\nu) = 0$.
Therefore let us look at the indecomposable partitions ν with $p\le 2l$, such that each ν_i has at least 2 elements.
Let $\mathscr{P}$ be the set of all partitions of set $\{k_1,k_2,\dots,k_l\}$ and define a function $s:\mathbb{N}^l\times\mathscr{P}\mapsto$

$\{0,1\}^p$, such that $[s(\{k_1,\ldots,k_l\},v)]_i := 0$ if $\sum \omega_k = 0$ in v_i and it is 1 otherwise. For a fixed partition $v = v_1 \cup \cdots \cup v_p$, using (6.5),the sum in (6.3) can be written as

$$C(v) = \frac{2^l}{T^{l/2}} \sum_{s(v)\in\{0,1\}^p} \sum_{\substack{\{k_1,\ldots,k_l\}: \\ s(\{k_1,\ldots,k_l\},v)=s(v)}} T^{-2l} T^{p-\|s(v)\|} O(1). \tag{6.6}$$

For every possible value of $s(v)$, we will find r, possible order for the set $\{k_1,\ldots,k_l\}$ and an upper bound for $r + p - \|s(v)\|$. Note that for some values of $s(v)$ there are no feasible solutions for $\{k_1,\ldots,k_l\}$ and hence the contribution of such $s(v)$ in the sum will be 0. So we focus on consistent values of $s(v)$.

To this end let $s(v_j) = [s(v)]_j$ and partition table (6.4) in blocks in the following way. First we look at the rows for which there is no v_j such that $v_j \cap \{k_i, -k_i, k_i - 1, -(k_i - 1)\} \neq \emptyset$ and $s(v_j) = 0$. In other words if any set v_j in the partition has an elements from i-th row then $s(v_j) = 1$. Each of these rows are one of the blocks, call them $B_{11},\ldots,B_{1r_1}$. On the rest of the rows define the following equivalence relationship. We say $i \sim j$ if there is a chain of sets $v_{m_1},\ldots,v_{m_t}$ connecting i-th and j-th row, such that $s(v_{m_k}) = 0$ for all k. It is easy to see it is in fact an equivalence relation. Therefore consider all the equivalence classes and that will give us a partition of the rows of the table. Each of these partitions are considered as separate blocks. Note that by construction each row in one block has a linear relationship with all the other rows in the same block. Reorder and label the blocks as $B_{21},\ldots,B_{2r_2},\ldots,B_{l1},\ldots,B_{lr_l}$ such that B_{ij} has $(i-1)$ independent solutions for the rows of B_{ij}, in the sense that, if we fix any $i-1$ rows of the block the rest will be fixed. Also by construction, if a set v_i has an element from both $B_{i_1 j_1}$ and $B_{i_2 j_2}$, with $(i_1, j_1) \neq (i_2, j_2)$, then $s(v_i) = 1$.

Claim 1. *For all blocks B_{ij} with $i > 2$, there exists sets $v_{m_1},\ldots,v_{m_t}$ with the property that $s(v_{m_k}) = 0$ and $|v_{m_k}| > 2$ for $k = 1,\ldots t$; $\cup v_{m_j} \subset B_{ij}$ and $|\cup v_{m_k}| \geq i + 2(t-1)$.*

Proof. By construction we can always find $v_{m_1},\ldots,v_{m_t}$ such that $s(v_{m_k}) = 0$ and $\cup v_{m_j} \subset B_{ij}$. First suppose $|v_{m_k}| = 2$ for $k = 1,\ldots t$. Note that by construction all the rows in the block must have one element in the v_{m_k}'s. Therefore all the rows are linearly related and if we fix one row, the rest of the rows is also fixed. This is a contradiction to the property that B_{ij} has $(i-1)$ independent solutions for $\{k_1,\ldots,k_l\}$. Therefore there must be at least one set with cardinality > 2 among the v_{m_k}'s.

Now look at the set, v_{m_1} with maximal cardinality. If $|v_{m_1}| \geq i$, then the claim holds with $t = 1$. If not,consider the rows that do not occur in v_{m_1}. By construction, all these rows must hook with all the rows appearing in v_{m_1} with sets v_k such that $s(v_k) = 0$. Find v_{m_2} such that $s(v_{m_2}) = 0$ and v_{m_2} has at least one elements from rows appearing in v_{m_1} and at least two rows from the rows that do not appear in v_{m_1}. There must be one such set, because if not once the rows appearing in v_{m_1} are fixed, the rest of the rows in the block is also fixed and in this situation number of independent rows from B_{ij} is $\leq |v_{m_1}| - 1 < i - 1$. Continue in this way to find $v_{m_1},\ldots v_{m_t}$. Each of these sets adds at most $|v_{m_k}| - 2$ independent variables, therefore number of independent rows in the block

$$i - 1 \leq |v_{m_1}| - 1 + \sum_{k=2}^{t} (|v_{m_k}| - 2) = \sum_{k=1}^{t} |v_{m_k}| - 1 - 2(t-1).$$

Consequently, we have $i+2(t-1) \leq \sum_{k=1}^{t} |v_{m_k}| = |\cup v_{m_k}|$. □

Note that each of the first r_1 rows can be chosen independently. Therefore from the construction $r = r_1 + r_2 + 2r_3 + \dots (m-1)r_m \leq l$.

Define the set $I := \{i \in \{1, \dots, p\} : s(v_i) = 0\}$. Note that $|I| = p - \|s(v)\|$. First we find t_{ij} sets from B_{ij} obtained by Claim 1. Let n be the number of elements left in the table, then $n \leq 4l - \sum_{i=3}^{m} ir_i - 2\sum_{i=3}^{m}\sum_{j=1}^{r_i}(t_{ij}-1)$.

Next we find a lower bound on number of elements in the set $\cup_{k\in I^c} v_k$. Because all the sets of the partition must have at least two elements, we have

$$|I| \leq \sum_{i=3}^{m}\sum_{j=1}^{r_i} t_{ij} + \frac{n - |\cup_{k\in I^c} v_k|}{2} \leq \sum_{i=3}^{m} r_i + 2l - \frac{\sum_{i=3}^{m} ir_i + |\cup_{k\in I^c} v_k|}{2} \tag{6.7}$$

To this end we consider the following cases separately:
Case I: There are more than one blocks, i.e., $\sum r_i > 1$.

Claim 2. *Each of the blocks must have at least 2 elements from the set* $\cup_{i\in I^c} v_i$.

Proof. All the blocks must communicate, therefore there must be at least one element in each block which belongs to some v_j that connects two different blocks and hence $s(v_j) = 1$. If there are more than one such elements we are done. If not, then there must be at least one other set v_j consisting of only elements in that block with $s(v_j) = 1$, owing to the fact that sum of all elements in one block is 0 and it must be nonzero if we take just one element out from the block. □

Therefore in this case each of the blocks must have at least 2 elements in the set $\cup_{i\in I^c} v_i$ and hence $|\cup_{i\in I^c} v_i| \geq 2(r_1 + r_2 + \cdots + r_m)$ and

$$|I| \leq 2l - r_1 - r_2 - \cdots - \frac{mr_m}{2}.$$

Consequently, it follows that

$$\begin{aligned} r + p - \|s(v)\| = r + |I| &\leq 2l + r_3/2 + \cdots + (m/2-1)r_m \\ &< 2l + r_1 + r_2 + \dots [(m-1)/2]r_m \leq 2l + l/2. \end{aligned}$$

The second inequality is strict because of the fact that at least one of the r_i's is positive. Therefore in this case $C(v) = O(T^{\delta})$ for some $\delta < 0$.
Case II: There is only one block, i.e., $r_k = 1$ for some $2 \leq k \leq m$ and $r_j = 0$ for all $j \neq k$.
Case II.1: $l = 2$: In this situation $k = 2$, as $k \leq l$. Therefore substituting $r = 1$ and $p \leq 2l$, we obtain

$$C(v) \leq T^{-l/2} O(T) T^{-2l} T^{2l} O(1) = O(T^{1-l/2}) = O(1).$$

Case II.2: $l > 2$: Here $k \leq l$ and $r = k-1$. By Claim 1, the total number of sets in the partition $p \leq t + (4l - k - 2(t-1))/2 = 2l - k/2 + 1/2$. Finally, as $\|s(v)\| \geq 0$, we get

$$C(v) \leq T^{-l/2} O(T^{k-1}) T^{-2l} T^{2l-k/2+1/2} O(1) = O(T^{k/2-l/2-1/2}) \leq O(T^{-1/2}).$$

This implies that the 2nd order cumulant is finite and cumulants of order > 2 converges to 0 as $T \to \infty$.