SFB 823

A note on quadratic forms of stationary functional time series under mild conditions

Anne van Delft

Discussion Paper

Nr. 18/2019

A note on quadratic forms of stationary functional time series under mild conditions

Anne van Delft*

Ruhr-Universität Bochum
Fakultät für Mathematik
44780 Bochum, Germany

May 31, 2019

Abstract

We study the distributional properties of a quadratic form of a stationary functional time series under mild moment conditions. As an important application, we obtain consistency rates of estimators of spectral density operators and prove joint weak convergence to a vector of complex Gaussian random operators. Weak convergence is established based on an approximation of the form via transforms of Hilbert-valued martingale difference sequences. As a side-result, the distributional properties of the long-run covariance operator are established.

keywords: functional data, time series, spectral analysis, martingales

AMS subject classification: Primary: 62HG99, 60G10, 62M15, Secondary 62M10.

1 Introduction

The subject of this paper are quadratic forms of a stationary time series $(X_t : t \in \mathbb{Z})$ with paths in some function space H. From a technical perspective, we shall adhere to existing literature and assume H is a separable Hilbert space. Each realization is therefore a function. Such *Functional time series* are of growing interest due to the fact that many processes are almost continuously measured on their domain of definition. Consequently, the number of realizations can be substantially smaller than the intrinsic variation of the process and inference methods must take this into account. While quadratic forms of Euclidean-valued random variables have received considerable attention and have been studied under various dependence conditions [see i.a. 18, 24, 20, 1, 30, 21, and references therein], this is not so much the case for quadratic forms of functional-valued random variables. Yet, they do arise naturally in a variety of inference problems. A quadratic form statistic of a functional time series can be given by

$$\hat{\mathscr{Q}}_T = \sum_{s,t=1}^{T} \Phi_{T,s,t}(X_s \otimes X_t) \tag{1}$$

where $\{\Phi_{T,t,s}\}_{t,s \in \{1,\ldots,T\}}$ defines a sequence of bounded linear operators which will vary depending on the application. Important applications in which statistics of the form (1) arise, are those

*anne.vandelft@rub.de

that concern the consistent estimation of the second-order characteristics of the process (or related operators). This is especially relevant in functional data because the smoothness properties of the random functions are encoded in the second-order structure and are key in obtaining optimal finite-dimensional representations. For example, if we denote $I_{H\otimes H}$ the identity operator on the tensor product space $H \otimes H$, then the specification $\Phi_{T,t,s} = \frac{1}{T} 1_{s=t} I_{H\otimes H}$ trivially yields the sample covariance operator. In the case of $i.i.d.$ functional data, this object captures the full second-order structure and its eigen decomposition plays a central role in the extraction to finite dimension of the process's properties, e.g., via the Karhunen-Loève representation if $H = L^2$. Not surprisingly, the sample covariance operator received considerable attention in the corresponding line of literature [e.g., 11, 7, 28] but also in case of linear processes [see among others 3, 8, 15, and references therein]. However, when there is serial correlation between observations the covariance operator clearly does not capture the full dynamics. For dependent functional data, a more meaningful object is therefore the spectral density operator

$$\mathscr{F}^{(\lambda)} = \frac{1}{2\pi} \sum_{h\in\mathbb{Z}} C_h e^{-\mathrm{i}\lambda h} \quad \lambda \in (-\pi, \pi] \tag{2}$$

where C_h is the h-lag covariance operator of the process X. As an estimator of $\mathscr{F}^{(\lambda)}$ for a process X with mean function μ, one can consider

$$\hat{\mathscr{F}}^{(\lambda)} = \frac{1}{2\pi T} \sum_{s,t=1}^{T} \underbrace{w(b_T(t-s))e^{\mathrm{i}\lambda(t-s)}}_{\phi^{(\lambda)}_{T,t-s}} \big((X_s - \mu) \otimes (X_t - \mu)\big), \quad \lambda \in (-\pi, \pi]$$

which simply corresponds to the quadratic form in (1) with $\Phi^{(\lambda)}_{T,t,s} = 2\pi T \phi^{(\lambda)}_{T,t-s} I_{H\otimes H}$. Here $w(\cdot)$ is an even, bounded function on $\mathbb{R}$ that is continuous at zero and b_T is a bandwidth parameter converging to zero at a rate such that $b_T T \to \infty$ as the sample size T tends to infinity. The properties of this estimator and its relation to the smoothed periodogram operator are discussed in detail in Section 4. For $\lambda = 0$, $2\pi\hat{\mathscr{F}}^{(0)}$ is an estimator of the long-run covariance operator. Because it arises as the limiting covariance operator of the sample mean function, properties of the long-run covariance operator have been studied in several contexts within the framework of L^p_m-approximability [see e.g., 16, 13].

Frequency domain analysis of functional time series, i.e., the case $\lambda \neq 0$, has received considerably less attention than time domain analysis. Yet, not only does frequency domain analysis and hence the spectral density operator arise in various applications in a natural manner such as in high-resolution medical data or biology, it captures moreover the full second order dynamics of dependent functional data. It can therefore be seen to take on a similar role for dependent functional data as the covariance operator takes on in the case of $i.i.d.$ functional data. In fact, it allows to extract the uncountably infinite variation to a countably infinite space in an optimal manner via a *dynamic* Karhunen-Loève representation provided the function space is sufficiently smooth. Moreover, frequency domain based inference methods enable powerful nonparametric tools for hypothesis testing. Because of its relevance for dependent functional data, estimators of $\mathscr{F}^{(\lambda)}$ in the context of L^p_m-dependence as well as under functional cumulant-mixing conditions were introduced earlier this decade. Under L^p_m approximability, [12] considered dynamic principal components for stationary functional time series and obtained a consistency result for a lag window estimator. Under cumulant-mixing conditions, [25] derived consistency and asymptotic normality of a smoothed periodogram operator estimator. Estimation and distributional properties of an estimator for a time-varying spectral density operator were derived in [9], who introduced a framework for locally stationary functional time series. Note that all of the aforementioned estimators can be written in the form (1). It is worth mentioning that these works have paved the way for frequency domain-based inference of functional time series, leading to an upsurge in the available literature in the past few years [see e.g. 14, 22, 27, 10, for some recent works and references therein].

Cumulant tensors and spectral cumulant tensors can be shown to form Fourier pairs, provided appropriate summability conditions are satisfied. The consideration of functional cumulant mixing conditions as in [25] can therefore to some extent be seen to provide a natural framework for the derivation of sampling properties. Yet, the central limit theorem and consistency result as derived in [25] rely on existence of all moments and summability conditions of the cumulant tensors. In certain applications such required summability conditions might be too strong and worthwhile to be relaxed. To the author's knowledge, there is currently no CLT available under L^p_m-dependence and the consistency rate available in this setting [12] is sub-optimal compared to the one derived under cumulant mixing conditions in [25].

Broadly speaking, the aim of this paper is therefore twofold. We wish to derive a general central limit theorem for quadratic forms of stationary functional time series under sharp moment conditions. At the same time, we aim to obtain the best possible convergence and consistency rates for the aforementioned applications. It is worth mentioning that our conditions on the dependence structure are also weaker than those considered within the L^p_m-dependence framework. Underlying our approach is an approximation of the quadratic form with a Hilbertian-valued martingale process. To construct this process, we shall use a martingale approximation of the quadratic form. The idea to approximate a normalized partial sum process via a related martingale process was first put forward by [17]. [31] introduced this approach to derive distributional properties of the Discrete Fourier transform (DFT) of a Euclidean-valued ergodic time series. The latter has since then been applied in a variety of problems [see e.g., 26, 30, 23]. In [6], the result of [31] and [26] was generalized to a CLT of the discrete Fourier transform of a Hilbertian-valued time series.

The structure of this note is as follows. In Section 2, we introduce necessary notation and conditions. In Section 3, we explain the approach in more detail and provide a joint central limit theorem for a set of quadratic forms as in (1). In Section 4 we focus on the estimation of the spectral density operator and long-run covariance as particular applications. More specifically, a consistency rate and distributional properties are established. Various technical results and proofs are relegated to the Appendix.

2 Framework

Let $(\Omega, \mathscr{A}, \mathbb{P})$ be a probability space and $(T, \mathscr{B}, \mu)$ a σ-finite measure space. For a separable Banach space $(B, \|\cdot\|_B)$ with dual B' we denote the space $L^p_B(T, \mathscr{B}, \mu)$ the space of p-th integrable B-valued functions equiped with norm $\|\cdot\|_{B,p} = (\int |\cdot|^p d\mu(t))^{1/p}$. We call a function $f : \Omega \to B$ $\mathscr{A}$-measurable if, for all $E \in \mathscr{B}(T)$, $f^{-1}(E) \in \mathscr{A}$. A function is strongly $\mathscr{A}$-measurable if and only if for all $x' \in B'$, the function $\langle f, x'\rangle$ is $\mathscr{A}$-measurable, i.e., if it is the pointwise limit of $\mathscr{A}$-measurable simple functions. A B-valued random element X over $(\Omega, \mathscr{A}, \mathbb{P})$ is then a strongly measurable function $X : (\Omega, \mathscr{A}, \mathbb{P}) \to B$. We denote $X \in \mathscr{L}^p_B$ if $\|X\|_{\mathbb{B},p} := (\mathbb{E}\|X\|^p_B)^{1/p} < \infty$. For $X \in \mathscr{L}^1_B(\Omega, \mathscr{A}, \mathbb{P})$ and $\mathscr{A}_o$ a sub-algebra of $\mathscr{A}$, we define the conditional expectation $\mathbb{E}[\cdot|\mathscr{A}_o]$: $\mathscr{L}^1_B(\Omega, \mathscr{A}, \mathbb{P}) \to \mathscr{L}^1_B(\Omega, \mathscr{A}_o, \mathbb{P})$ to be the mapping such that

$$\int_A \mathbb{E}[X|\mathscr{A}_o] d\mathbb{P} = \int_A X d\mathbb{P} \quad A \in \mathscr{A}_o$$

where the expectations should all be understood in the sense of a Bochner integral. Note that classical properties of conditional expectation remain valid in the context of separable Banach spaces. Observe that $\mathcal{L}_B^p$ is a Banach space w.r.t. the norm $\|X\|_{\mathbb{B},p}$ and for $p = 2$ it is a Hilbert space consisting of B-valued random variables X with finite second moment. Troughout this paper, we will focus on Hilbert-valued random variables and denote the corresponding space by $\mathcal{L}_H^p$ and its norm by $\|X\|_{\mathbb{H},p}$. For elements of H, we shall denote the inner product by $\langle\cdot,\cdot\rangle$ and the induced norm by $\|\cdot\|_H$. We let $H_1 \otimes H_2$ denote the Hilbert tensor product of the Hilbert spaces $(H_i, \langle\cdot,\cdot\rangle_j)_{j=1,2}$. This Hilbert space can be constructed from the algebraic tensor product $H_1 \otimes_{alg} H_2$ together with a bilinear map $\psi : H_1 \times H_2 \to H_1 \otimes_{alg} H_2$ that satisfies $\langle\psi(x_1, x_2), \psi(y_1, y_2)\rangle = \langle x_1, x_2\rangle_1\langle y_1, y_2\rangle_2$ for $x_1, y_1 \in H_1$ and $x_2, y_2 \in H_2$ and then taking the completion with respect to the induced norm [see e.g., 19, for details].

Furthermore, we require some terminology for operators acting on H. Firstly, denote by $S_\infty(H)$ the Banach space of bounded linear operators $A\colon H \to H$ equipped with the operator norm $|||A|||_\infty = \sup_{\|g\|\leq 1} \|Ag\|_H, g \in H$. Let $\{\chi_i\}_{i\geq 1}$ be an arbitrary orthonormal basis of H and $A^\dagger$ denote the adjoint of A. An operator A is called self-adjoint if $\langle Af, g\rangle = \langle f, A^\dagger g\rangle = \langle f, Ag\rangle$ for all $f, g \in H$ and non-negative definite if $\langle Ag, g\rangle \geq 0$ for all $g \in H$. For $A, B, C \in S_\infty(H)$ we define the kronecker product as $(A\widetilde{\otimes}B)C = ACB^\dagger$, while the transpose Kronecker product is given by $(A\widetilde{\otimes}_\top B)C = (A\widetilde{\otimes}\overline{B})\overline{C}^\dagger$.

A compact operator $A : H \to H$ belongs to the class of Hilbert-Schmidt operators, denoted by $S_2(H)$, if it has finite Hilbert-schmidt norm which is defined as $|||A|||_2 := (\sum_{i=1}^\infty \|A(\chi_i)\|_H^2)^{1/2}$. If A has finite-trass class norm, where the norm is given by $|||A|||_1 = \sum_{i=1}^\infty \langle (AA^\dagger)^{1/2})(\chi_i), \chi_i\rangle$, we say A is trace-class and denote $A \in S_1(H)$. The space $(S_2(H), |||\cdot|||_2)$ is a Hilbert space with inner product $\langle A, B\rangle_S = \mathrm{Tr}(AB^\dagger) = \sum_{i=1}^\infty \langle A(\chi_i), B(\chi_i)\rangle, A, B \in S_2(H)$. For $f, g, v \in H$, define the tensor product $f \otimes g\colon H \otimes H \to H$ as the bounded linear operator $(f \otimes g)v = \langle v, g\rangle f$. The mapping $\mathcal{T}\colon H \otimes H \to S_2(H)$ defined by the linear extension of $\mathcal{T}(f \otimes g) = f \otimes \overline{g}$ is an isometric isomorphism. For two zero-mean elements $X, Y \in \mathcal{L}_H^2$, the cross- covariance operator is given by $\mathrm{Cov}(X, Y) = \mathbb{E}(X \otimes Y)$ and belongs to $S_1(H)$. We note in particular that $\|X\|_{\mathbb{H},2}^2 = \mathrm{Tr}(\mathrm{var}(X \otimes X))$. For a filtration $\{\mathcal{G}_j\}$ of sub σ-algebras of $\mathcal{A}$, we shall make extensive use of projection operators defined by

$$P_k = \mathbb{E}[\cdot|\mathcal{G}_k] - \mathbb{E}[\cdot|\mathcal{G}_{k-1}], \quad k \in \mathbb{Z}$$

which are linear operators on $\mathcal{L}_H^1$ and are strongly orthogonal elements in $\mathcal{L}_H^2$, i.e.,

$$\mathrm{Cov}(P_k(X_1), P_j(X_2)) = O_H \quad \forall X_1, X_2 \in \mathcal{L}_H^2 \text{ and } k \neq j \in \mathbb{Z}.$$

Finally, we let $\Rightarrow_N$ indicate convergence in distribution as $N \to \infty$, where $N \in \mathbb{N}$.

3 Main result

Throughout this article, we are interested in weakly stationary functional time series $\{X_t\colon t \in \mathbb{Z}\}$ taking values in $\mathcal{L}_H^2$. In particular this means that the mean $\mathbb{E}X_t = \mu$ and the h-lag covariance operator C_h are invariant under translations in time, i.e, $C_h = \mathbb{E}(X_h - \mu) \otimes (X_0 - \mu)$. Without loss of generality, we shall assume throughout this article that the data are centered. When the mean is unknown one can consider centering the data by subtracting the sample mean function (see Remark 4.1). Furthermore, we assume the process admits a representation of the form

$$X_t = g(\epsilon_t, \epsilon_{t-1}, \ldots,)$$

where $\{\epsilon_t : t \in \mathbb{Z}\}$ is an i.i.d. sequence of elements in some normed vector space S and $g : S^\infty \to H$ is measurable. Functional processes with such representation are widely applicable and allow for example for nonlinear dynamics[e.g., 13]. It is clear from this representation that X is stationary and ergodic and we can consider the filtration $\mathscr{G}_t = \sigma(\epsilon_t, \epsilon_{t-1}, \dots)$. Moreover, it is straightforward to show that a stationary ergodic process can be written as

$$X_h = \sum_{j=-\infty}^{h} P_j(X_h) \tag{3}$$

where the equality holds in $\mathscr{L}^2_H$. Note that $\{P_j(X)\}$ form a martingale difference sequence with respect to the backward filtration of σ-algebras $\{\mathscr{G}_{-j} : j > 0\}$. In order to formulate conditions on the dependence structure we consider a generalized version of the physical dependence measure of [32]. More specifically, let $\{\epsilon'_t : t \in \mathbb{Z}\}$ be an independent copy of $\{\epsilon_t : t \in \mathbb{Z}\}$ defined on $(\Omega, \mathscr{A}, \mathbb{P})$. For a set $I \subset \mathbb{Z}$, let $\mathscr{G}_{t,I} = \sigma(\epsilon_{t,I}, \epsilon_{t-1,I}, \dots)$ where $\epsilon_{t,I} = \epsilon'_t$ if $t \in I$ and $\epsilon_{t,I} = \epsilon_t$ if $t \not\in I$ and as a measure of dependence define

$$\nu_{\mathbb{H},p}(X_t) = \| X_t - \mathbb{E}[X_t | \mathscr{G}_{t,\{0\}}] \|_{\mathbb{H},p}$$

Additionally we define the m-dependent process

$$X_t^{(m)} = P_{t,t-m} X_t = \mathbb{E}[X_t | \sigma(\epsilon_t, \epsilon_{t-1}, \dots, \epsilon_{t-m})].$$

The following summarizes the assumption on the dependence structure made throughout this paper.

Assumption 3.1. *Let $\{X_t\colon\ t \in \mathbb{Z}\}$ be a centered stationary functional time series in $\mathscr{L}^p_H$ such that*

$$\sum_{j=0}^{\infty} \nu_{\mathbb{H},p}(X_j) < \infty \tag{4}$$

with $p = 4$.

Observe that Assumption 3.1 is weaker than L^p_m-dependence. More specifically, L^p_m-dependence would correspond to the condition $\sum_{m=0}^{\infty} \| X_t - X_t^{(m)} \|_{\mathbb{H},p} < \infty$. Additionally, note that by Jensen's inequality and the contraction property of the conditional expectation that

$$\| P_0(X_t) \|_{\mathbb{H},p} = \| \mathbb{E}[X_t - \mathbb{E}[X_t | \mathscr{G}_{t,\{0\}}] | \mathscr{G}_0] \|_{\mathbb{H},p} \le \nu_{\mathbb{H},p}(X_t).$$

Hence, under condition (4) we have $\sum_{j=0}^{\infty} \| P_0(X_j) \|_{\mathbb{H},p} < \infty$.

The assumption $\{X_t\colon\ t \in \mathbb{Z}\} \in \mathscr{L}^4_H$ ensures a finite second-order structure of a random element of the form $X_t \otimes X_s$. Note that the latter can be viewed as a random element of $S_2(H)$, i.e., it is a measurable mapping from $(\Omega, \mathscr{A})$ into $(S_2(H), \mathscr{B})$ and thus $X_t \otimes X_s \in \mathscr{L}^2_{S_2(H)}$. Existence of a limit of the quadratic form in (1) requires conditions on both the weight sequence as well as on the dependence structure. To elaborate on the latter, the condition in (4) has two implications (Proposition 3.1 and Proposition 3.2, resp.), which we shall require in order to derive distributional properties of the quadratic form. Denote the functional Discrete Fourier Transform (fDFT) of the stationary process X by

$$\mathscr{D}_T^{(\lambda)} = \frac{1}{\sqrt{2\pi T}} \sum_{t=1}^{T} X_t e^{-\mathrm{i}\lambda t}. \tag{5}$$

The second-order structure of this object –if well-defined– provides information on how the variation that is contained in the process is distributed over frequencies and can be viewed as an estimator of a Hilbertian-valued orthogonal increment process[see 10, for necessary conditions]. Provided the memory of the process decays fast enough, its limiting variance is given by the spectral density operator in (2). Assumption 3.1 with $p = 2$ provides sufficiently fast decay in memory for this to be the case.

Proposition 3.1. *Under Assumption 3.1 with $p = 2$, $\sum_{h\in\mathbb{Z}} |||C_h|||_2 < \infty$ and*

$$\lim_{T\to\infty} \mathrm{Var}(\mathscr{D}_T^{(\lambda)}) = \mathscr{F}^{(\lambda)}$$

exists as a non-negative definite Hermitian element of $S_2(H)$.

Proof of Proposition 3.1. We obtain by orthogonality of the projections, stationarity and Jensen's inequality

$$\begin{aligned}\sum_{h=0}^{\infty} |||C_h|||_2 &\le \sum_{h=0}^{\infty}\sum_{j=-\infty}^{h} \mathbb{E}\Big|\Big|\Big| P_0(X_{h-j}) \otimes P_0(X_{-j})\Big|\Big|\Big|_2 \le \sum_{h=0}^{\infty}\sum_{j=-\infty}^{h} \sqrt{\mathbb{E}\|P_0(X_{h-j})\|_H^2 \mathbb{E}\|P_0(X_{-j})\|_H^2} \\ &\le \Big(\sum_{j=0}^{\infty} \|P_0(X_j)\|_{\mathbb{H},2}\Big)^2 < \big(\sum_{j=0}^{\infty} \nu_{\mathbb{H},2}(X_t)\big)^2 < \infty\end{aligned}$$

and similarly for $h < 0$. Hence, $\mathscr{F}_\lambda = \frac{1}{2\pi}\sum_{h\in\mathbb{Z}} C_h e^{-\mathrm{i}h\lambda}$ converges in norm $|||\cdot|||_2$ for all $\lambda \in (-\pi,\pi]$. It follows that $\mathscr{F}^{(\lambda)}$ is a non-negative definite, Hermitian $S_2(H)$-valued density function over frequencies that satisfies $C_h = \int_{-\pi}^{\pi} \mathscr{F}^{(\lambda)} e^{\mathrm{i}h\lambda} d\lambda$ [e.g., 10, Thm 3.7]. Moreover, from the dominated convergence theorem one obtains

$$\lim_{T\to\infty} \mathrm{Var}(\mathscr{D}_T^{(\lambda)}) = \lim_{T\to\infty} \sum_{h\le T} (1 - \frac{|h|}{T}) \mathbb{E}(X_h \otimes X_0) e^{-i\lambda h} = \mathscr{F}^{(\lambda)} \quad \lambda \in (-\pi,\pi]. \tag{6}$$

□

Hence, (2) exists as a limit of Césaro averages of $\{C_h e^{-\mathrm{i}h\lambda} : h \in \mathbb{Z}\}$ in $S_2(H)$. Without stronger assumptions, such as summability conditions, the derivation of several distributional properties of the quadratic form in (1) do not appear obvious. Yet Assumption 3.1 allows to proceed via an approximating $S_2(H)$-valued random process. Underlying this approximation is the following process

$$D_{m,k,T}^{(\lambda)} := \frac{1}{\sqrt{2\pi}} \sum_{t=0}^{T} P_k(X_{t+k}^{(m)}) e^{-\mathrm{i}t\lambda}. \tag{7}$$

The second-order structure of (7) is closely related to that of $\mathscr{D}_T^{(\lambda)}$, but moreover has several useful properties that we shall make extensive use of.

Proposition 3.2. *Under the conditions of Assumption 3.1 with $p = 4$, we have for all $\lambda \in (-\pi,\pi]$,*

(i) The process $D_{m,k}^{(\lambda)} := D_{m,k,\infty}^{(\lambda)}$ forms a m-dependent stationary martingale difference sequence with respect to the filtration $\{\mathscr{G}_k\}$ in $\mathscr{L}_H^4$.

(ii) The process $\{D_{\infty,0,T}^{(\lambda)}\}_{T\ge 1}$ is Cauchy in $\mathscr{L}_H^4$ with limit

$$D_0^{(\lambda)} := \sum_{t=0}^{\infty} P_0(X_t) e^{-\mathrm{i}t\lambda}$$

and the process $\{D_{\infty,0,T}^{(\lambda)} \otimes D_{\infty,0,T}^{(\lambda)}\}_{\{T\ge 1\}}$ is Cauchy in $\mathscr{L}_{S_2(H)}^2$ with limit $D_0^{(\lambda)} \otimes D_0^{(\lambda)}$.

(iii) $\lim_{m\to\infty} \lim_{T\to\infty} \mathrm{Tr}\Big(\Pi_{ijkl}\big(\mathrm{Var}(D_{m,0,T}^{(\lambda)}) \widetilde{\otimes} \mathrm{Var}(D_{m,0,T}^{(\lambda)})\big)\Big) = \lim_{m\to\infty} \mathrm{Tr}\Big(\Pi_{ijkl}\big(\mathscr{F}_m^{(\lambda)} \widetilde{\otimes} \mathscr{F}_m^{(\lambda)}\big)\Big) = \mathrm{Tr}\Big(\Pi_{ijkl}\big(\mathscr{F}^{(\lambda)} \widetilde{\otimes} \mathscr{F}^{(\lambda)}\big)\Big) < \infty.$

Here, Π_{ijkl} denote the permutation operator on $\otimes_{i=1}^{4} H$ that permutes the components of a tensor product of simple tensors according to the permutation $(1,2,3,4) \mapsto (i,j,k,l)$, that is, $\Pi_{ijkl}(x_1 \otimes \cdots \otimes x_4) = (x_i \otimes \cdots \otimes x_l)$. The details of the proof can be found in Appendix A. For fixed m and $p = 2$, the first statement is almost immediate from the properties of the projection operators which form martingale difference sequences with respect to $\{\mathscr{G}_k\}$ and the fact that the process $\{X_t^{(m)}\}$ is m-dependent. For $p = 4$, the proof of the above statements require extensions of inequalities such as Burkholder's inequality for linear transforms of Hilbert-valued martinagles; see Appendix A. The Cauchy property will be necessary to verify several aspects of the distributional properties, including verification of tightness on the function space of the quadratic form. Proposition 3.2(iii) shows in particular that the iterated limit in T and m, resp, of a certain functional of the variance operators of the family of martingale processes $\{D_{m,0,T}^{(\lambda)}\}_{T\geq 1}, m \geq 1$ converge to that of the corresponding functional of $\mathscr{F}^{(\lambda)}$, i.e. of the limiting variance operators of the fDFT, and that this functional is finite.

Next, we require the following conditions on the sequence of weight operators. We assume that we have a representation $\Phi_{T,s,t} = (\phi_{T,s,t} \widetilde{\otimes} I_H)$, where $\phi_{T,s,t} \in S_\infty(H)$ such that $\Phi_{T,s,t} = \Phi^\dagger_{T,s,t}$. Observe that this is an operator in $S_\infty(S_2(H))$ with the property

$$(\phi_{T,s,t} \widetilde{\otimes} I_H)(X_s \otimes X_t) = \phi_{T,s,t}(X_s) \otimes I_H(X_t) = (\phi_{T,s,t} \widetilde{\otimes} I_H)^\dagger (X_s \otimes X_t) = I_H(X_s) \otimes \phi^\dagger_{T,s,t}(X_t) \qquad (8)$$

Note that the identity operator can be replaced with any arbitrary bounded linear operator $B_T \in S_\infty(H)$. Additionally, we require a few technical conditions ensuring that the the weights are "well-behaved", i.e., the quadratic form exists as a well-defined random element of $S_2(H)$ for which no degenerate (non-Gaussian) limiting distributions can arise.

Assumption 3.2 (Conditions on $\phi : \mathbb{Z} \times \mathbb{Z} \to S_\infty(H)$)**.** *Let $T \in \mathbb{N}$ and $\lambda \in (-\pi, \pi]$. Let $A_{T,(\cdot)} : \mathbb{Z} \to S_\infty(H)$ be a continuous mapping such that $A_{T,t} \equiv A_{T,-t}, \forall t \in \mathbb{Z}$ and set $\phi_{T,t}^{(\lambda)} = A_{T,t} e^{\mathrm{i}\lambda t}$. Denote*

$$|||\Phi_T|||_F^2 := \sum_{t=1}^{T} \sum_{s=1}^{T} |||\phi_{T,t-s}|||_\infty^2 \text{ and } \varrho_T^2 := \sum_{t=1}^{T} |||\phi_{T,t}|||_\infty^2.$$

We assume,

(i) $T\varrho_T^2 = O(|||\Phi_T|||_F^2)$;

(ii) $\max_{1\leq t\leq T} |||\phi_{T,t}|||_\infty^2 = \max_{1\leq t\leq T} |||A_{T,t}|||_\infty^2 = o(\varrho_T^2)$;

(iii) $\sum_{t=1}^{T} |||A_{T,t} - A_{T,t-1}|||_\infty^2 = o(\varrho_T^2)$;

(iv) $\sum_{j=1}^{T-1} \sum_{s=1}^{j-1} |||\sum_{t=j+1}^{T} \phi_{T,s-t} \widetilde{\otimes} \phi_{T,j-t}|||_\infty^2 = o(\|\Phi_T\|_F^4)$.

Note that the first condition simply ensures a balance in order, i.e., the left-hand side of the same order as the total sum of weights operator when the latter is a functional-valued operator on $\mathbb{Z} \times \mathbb{Z}$. Together with the second, this means the norm of none of the individual weight contributions dominates the order of the variance. The third condition ensures a "smooth" contribution of each component $\Phi_{T,s,t}(X_s \otimes X_T)$ to the total mass of the quadratic form. The fourth condition is required to ensure that, as the overlap of the two bivariate operator-valued functions over $\mathbb{Z} \times \mathbb{Z}$ gets smaller, the contribution to the total mass must become negligible. Observe that for the examples mentioned in the introduction where $\phi_{T,t}^{(\lambda)}$ are scalar-valued, the norms $|||\cdot|||_\infty$ can be replaced by $|\cdot|$. Condition (iv) on the kernel then simply means a bandwidth parameter $b_T << 1$ must ensure a local smoothing occurs. As will become clear in the next section, it predictably excludes that the periodogram operator without smoothing can provide an asymptotically Gaussian consistent estimator of the spectral density operator. Many different weight functions used for the consistent estimation of $\mathscr{F}^{\lambda}$ will satisfy the above conditions, including the common choice of a bounded piecewise continous lag window function with compact support, provided the bandwidth parameter ensures condition (iv) holds true (see Section 4).

In order to derive the properties of the quadratic form, a natural and common approach is to decompose $\hat{\mathscr{Q}}_T$ into off-diagonal elements and diagonal elements as follows

$$\hat{\mathscr{Q}}_T = \sum_{t=2}^{T}\sum_{s=1}^{t-1} \Phi^{(\lambda)}_{T,s,t}(X_s \otimes X_t) + \Big(\sum_{t=2}^{T}\sum_{s=1}^{t-1} \Phi^{(\lambda)}_{T,s,t}(X_s \otimes X_t)\Big)^{\dagger} + \sum_{1\le t\le T} \Phi_{T,t,t}(X_t \otimes X_t). \tag{9}$$

The main ingredient to the proof is to use that the off-diagonal elements, after centering around their mean, can be approximated by the process

$$\mathscr{M}^{(\lambda)}_{T,m} = \sum_{t=2}^{T}\sum_{s=1}^{t-1} \Phi^{(\lambda)}_{T,t,s}\Big(D^{(\lambda)}_{m,t} \otimes D^{(\lambda)}_{m,s}\Big), \tag{10}$$

where the functionals $D^{(\lambda)}_{m,t}$ are defined via (7) in Proposition 3.2(i). The intuition is therefore similar in spirit to the strategy applied in the Euclidean setting [see e.g., 23, 30]. We emphasize that the aim of this paper is not the same nor can the weak convergence result in our paper be seen as a trivial extension of these works. We aim to derive consistency rates and joint distributional convergence of a set of operators where the quadratic form is very general, consisting of operator-valued weight operators of a Hilbertian-valued stochastic process. The derivation of the operator approximations and of the distributional properties, including the verification of tightness on the function space, are therefore far more involved. The convenient properties of (10) are given in the next statement.

Proposition 3.3. *Let $\mathscr{M}^{(\lambda)}_{T,m}$ as defined in* (10). *Under Assumption 3.1 with $p = 4$ and fixed m, the process*

$$\Big\{|||\Phi_T|||_F^{-1}\mathscr{M}^{(\lambda)}_{T,m}\Big\}_{T\ge 1}$$

is a martingale process in $\mathscr{L}^2_{S_2(H)}$ with respect to the filtration $\{\mathscr{G}_T\}$ for all fixed $\lambda \in (-\pi,\pi]$.

Proof of Proposition 3.3. It is immediate that $\mathscr{M}^{(\lambda)}_{T,m}$ is adapted to the filtation $\mathscr{G}_T$. Secondly, from the properties of the operators $\{\Phi_{T,s,t}\}$, we can write

$$\mathscr{M}^{(\lambda)}_{T,m} = \sum_{t=2}^{T}\sum_{s=1}^{t-1} \Phi^{(\lambda)}_{T,s,t}\Big(D^{(\lambda)}_{m,t} \otimes D^{(\lambda)}_{m,s}\Big) = \sum_{t=2}^{T} D^{(\lambda)}_{m,t} \otimes \Big(\sum_{s=1}^{t-1} \phi^{(\lambda)}_{T,s,t} D^{(\lambda)}_{m,s}\Big)$$

From Proposition 3.2(i), $D^{(\lambda)}_{m,t}$ forms a stationary martingale difference sequence in $\mathscr{L}^4_H$ with respect to $\mathscr{G}_t$. Hence, using orthogonality of the increments and by Lemma A.1

$$\mathbb{E}|||\mathscr{M}^{(\lambda)}_{T,m}|||_2^2 \le \sum_{t=2}^{T} \mathbb{E}\Big|\Big|\Big|D^{(\lambda)}_{m,t} \otimes \Big(\sum_{s=1}^{t-1} \phi^{(\lambda)}_{T,s,t} D^{(\lambda)}_{m,s}\Big)\Big|\Big|\Big|_2^2 \le \|D^{(\lambda)}_{m,0}\|^2_{\mathbb{H},2}\|D^{(\lambda)}_{m,0}\|^2_{\mathbb{H},2} \sum_{t=2}^{T}\sum_{s=1}^{t-1} |||\phi^{(\lambda)}_{T,s,t}|||^2_\infty$$

Noting that $\sum_{t=2}^{T}\sum_{s=1}^{t-1}|||\phi^{(\lambda)}_{T,s,t}|||^2_\infty \approx 1/2|||\Phi_T|||_F^{-1}$, we obtain $|||\Phi_T|||_F^{-1}\mathbb{E}|||\mathscr{M}^{(\lambda)}_{T,m}|||_2^2 < \infty$. Finally, observe that

$$\mathbb{E}\Big[D^{(\lambda)}_{m,t} \otimes \Big(\sum_{s=1}^{t-1} \phi^{(\lambda)}_{T,s,t} D^{(\lambda)}_{m,s}\Big)|\mathscr{G}_{t-1}\Big] = \mathbb{E}\Big[D^{(\lambda)}_{m,t}|\mathscr{G}_{t-1}\Big] \otimes \Big(\sum_{s=1}^{t-1} \phi^{(\lambda)}_{T,s,t} D^{(\lambda)}_{m,s}\Big) = O_H$$

where we used that $\sum_{s=1}^{t-1}\phi^{(\lambda)}_{T,s,t}D^{(\lambda)}_{m,s}$ is $\mathscr{G}_{t-1}$-measurable and that $D^{(\lambda)}_{m,t}$ is a H-valued martingale with respect to $\mathscr{G}_t$. The result now follows. □

The following theorem states the distributional properties of the quadratic form.

Theorem 3.1 (asymptotic normality of $\hat{\mathscr{Q}}_T^{\lambda}$)**.** *Let $\{X_t\}$ be a random sequence with paths in a separable Hilbert space H for which assumption* (3.1) *holds with $p = 4$ and suppose that the sequence $\{\Phi_T^{\lambda}\}$ satisfies Assumption 3.2. Then the quadratic form in* (1) *satisfies*

$$(|||\Phi_T|||_F^2)^{-1/2}\big(\hat{\mathscr{Q}}_T^{\lambda_j} - \mathbb{E}(\hat{\mathscr{Q}}_T^{\lambda_j}\big)_{j=1,\ldots,d} \Rightarrow \big(\breve{\mathscr{Q}}^{\lambda_j}\big)_{j=1,\ldots,d}$$

where, $\breve{\mathscr{Q}}^{\lambda_j}$, $j = 1,\ldots,d$ are jointly complex Gaussian elements of $S_2(H)$

$$\begin{pmatrix} \Re(\breve{\mathscr{Q}}^{\lambda_j}) \\ \Im(\breve{\mathscr{Q}}^{\lambda_j}) \end{pmatrix}_{j=1,\ldots,d} \sim \mathscr{N}_{(S_2(H))^d \times (S_2(H))^d}\left(\begin{pmatrix} O_H \\ O_H \end{pmatrix}, \frac{1}{2}\begin{pmatrix} \Re(\Gamma+\Sigma) & \Im(-\Gamma+\Sigma) \\ \Im(\Gamma+\Sigma) & \Re(\Gamma-\Sigma) \end{pmatrix}\right).$$

The (i,j)-th element of the covariance operator is given by

$$\Gamma_{i,j} = \eta(\lambda_i \pm \lambda_j)4\pi^2\Big(\mathscr{F}^{(\lambda_j)}\widetilde{\otimes}\mathscr{F}^{(\lambda_j)} + 1_{\{0,\pi\}}\mathscr{F}^{(\lambda_j)}\widetilde{\otimes}_{\top}\mathscr{F}^{(\lambda_j)}\Big)$$

and of the pseudocovariance operator by

$$\Sigma_{i,j} = \eta(\lambda_i \pm \lambda_j)4\pi^2\Big(1_{\{0,\pi\}}\mathscr{F}^{(\lambda_j)}\widetilde{\otimes}\mathscr{F}^{(\lambda_j)} + \mathscr{F}^{(\lambda_j)}\widetilde{\otimes}_{\top}\mathscr{F}^{(\lambda_j)}\Big),$$

and where $\eta(x) = 1$ for $x = 2\pi z, z \in \mathbb{Z}$ and zero otherwise.

In particular,for distinct frequencies $\lambda_1,\ldots,\lambda_d \in [0,\pi]$, Γ and Σ are $d \times d$ diagonal matrices with $S_1(H \otimes H)$-valued components and hence for such choice of frequencies, the components of $\big(\breve{\mathscr{Q}}^{\lambda_j}\big)_{j=1,\ldots,d}$ are asympticallly independent.

Proof of Theorem 3.1. We consider the sequence of processes $\{\xi_T^{\lambda} : T \in \mathbb{N}\}$ where

$$\xi_T^{\lambda} := (|||\Phi_T^{\lambda}|||_F^2)^{-1/2}(\hat{\mathscr{Q}}_T^{\lambda} - \mathbb{E}\hat{\mathscr{Q}}_T^{\lambda}).$$

Observe that $\{\xi_T^{\lambda} : T \in \mathbb{N}\}$ is a measurable stochastic processes with sample paths in the Hilbert space $S_2(H)$. We shall verify the following two conditions to show weak convergence on $S_2(H)$ [see e.g., 2]

Lemma 3.1 (weak convergence)**.** *Let $\{\xi_T : T \in \mathbb{N}\}$ be a stochastic process with sample paths in a separable Hilbert space. If the following two conditions are satisfied*

i) *The finite dimensional distributions of ξ_T converge to those of ξ a.e.;*
ii) *The family of laws $\mathscr{P} := (\mathbb{P}_T)_{T\in\mathbb{N}}$ of $\{\xi_T : T \in \mathbb{N}\}$ is tight.*

then, $\xi_T \Rightarrow_T \xi$.

First we derive that, for all $m \geq 1$, $\xi_{T,m} \Rightarrow_T \xi_m$, where ξ_m defines a zero-mean Gaussian element of $S_2(H)$ and where the double indexed process is given by

$$\xi_{T,m}^{\lambda} := (|||\Phi_T|||_F^2)^{-1/2}(\mathscr{M}_{T,m}^{(\lambda)} + \mathscr{M}_{T,m}^{\dagger(\lambda)})$$

with $\mathscr{M}_{T,m}^{(\lambda)}$ as in (10). By Proposition 3.3, for every fixed m, $\mathscr{M}_{T,m}^{(\lambda)}$ is a martingale process in $\mathscr{L}^2_{S_2(H)}(\Omega,\mathscr{A},\mathbb{P})$ with respect to the filtration $\{\mathscr{G}_T\}$. Note the same holds for $\mathscr{M}_{T,m}^{\dagger(\lambda)}$. Let $(\chi_l)_{l\geq1}$ be an orthonormal basis of H. Then $(\chi_{l,l'})_{l,l'} := (\chi_l \otimes \chi_{l'})_{l,l'}$ defines an orthonormal basis of $H \otimes H$ and denote

$$\xi_{T,m}^{\lambda}(\chi) = \langle \xi_{T,m}^{\lambda}, \chi \rangle.$$

for any $H \otimes H$. The result below shows that, for a finite set of frequencies, the finite-dimensional distributions of $\xi_{T,m}^{\lambda}$ for fixed m converge jointly to those of ξ_m^{λ} as $T \to \infty$, where these are asymptotically independent at distinct frequencies.

Theorem 3.2. *Suppose the conditions of Theorem 3.1 hold. Then, for a finite set of distinct frequencies $\lambda_1, \dots, \lambda_d \in [0,\pi]$, for all $\forall m \geq 1$ and any $\chi_{l_j l'_j} \in H \otimes H$, we have*

$$\{\xi^{\lambda_j}_{T,m}(\chi_{l_j l'_j})\}_{j=1,\dots,d} \Rightarrow_T \{\xi^{\lambda_j}_{m}(\chi_{l_j l'_j})\}_{j=1,\dots,d} \sim \mathcal{N}_{\mathbb{C}^d}\Big(0, \mathrm{diag}\big(\Gamma^{\lambda_j}_m(\chi_{l_j l'_j})\big), \mathrm{diag}\big(\Sigma^{\lambda_j}_m(\chi_{l_j l'_j})\big)\Big)$$

where

$$\Gamma^{\lambda_j}_m(\chi_{l_j l'_j}) = 4\pi^2\Big(\mathscr{F}^{(\lambda_j)}_m(\chi_{l_j l_j})\overline{\mathscr{F}^{(\lambda_j)}_m(\chi_{l'_j l'_j})} + 1_{\{0,\pi\}}\big(\mathscr{F}^{(\lambda_j)}_m(\chi_{l_j l'_j})\mathscr{F}^{(\lambda_j)}_m(\chi_{l_j l'_j})\big)\Big) \tag{11}$$

and

$$\Sigma^{\lambda_j}_m(\chi_{l_j l'_j}) = 4\pi^2\Big(1_{\{0,\pi\}}\big(\mathscr{F}^{(\lambda_j)}_m(\chi_{l'_j l'_j})\mathscr{F}^{(\lambda_j)}_m(\chi_{l_j l_j})\big) + \mathscr{F}^{(\lambda_j)}_m(\chi_{l_j l'_j})\mathscr{F}^{(\lambda_j)}_m(\chi_{l_j l'_j})\Big)$$

where $\mathscr{F}^{(\lambda_j)}_m(\chi_{l_j l'_j}) = \big(\mathbb{E}(D^{(\lambda_j)}_{m,0} \otimes D^{(\lambda_j)}_{m,0})\big)(\chi_{l_j l'_j})$.

The proof is tedious and relegated to Appendix B. Next, we show that $\forall m \geq 1$, $\{\xi^{\lambda}_{T,m}, T \geq 1\}$ is tight. In order to verify tightness we shall use the following result, which is a particular case of [29, Theorem 3] who considers tightness criteria for more general Schauder decomposable Banach spaces.

Lemma 3.2 (tightness on a separable Hilbert space)**.** *Let $(\chi_{ll'})$ be an orthonormal basis of $H \otimes H$. A family of probability measures $\mathscr{P} := (\mathbb{P}_T)_{T\in\mathbb{N}}$ on $S_2(H)$ is tight if and only if*

i) $\forall k \geq 1: \quad \lim_{h\to\infty} \sup_T \mathbb{P}_T\Big\{x \in S_2(H) : \sum_{l,l'<k} |\langle x, \chi_{ll'}\rangle|^2 > h\Big\}\Big) = 0$*;*

ii) $\forall \epsilon > 0: \quad \lim_{k\to\infty} \sup_T \mathbb{P}_T\Big(\Big\{x \in S_2(H) : \sum_{l,l':l+l'>k} |\langle x, \chi_{ll'}\rangle|^2 > \epsilon\Big\}\Big) = 0.$

In order to verify the first condition, note that, since k is fixed,

$$\lim_{h\to\infty} \sup_T \mathbb{P}\Big(\sum_{l,l'<k} |\langle \xi^{\lambda}_{T,m}, \chi_{ll'}\rangle|^2 > h\Big) \leq \sum_{l,l'<k} \lim_{h\to\infty} \sup_T \mathbb{P}\Big(|\langle \xi^{\lambda}_{T,m}, \chi_{ll'}\rangle|^2 > h\Big)$$

and hence the first condition is implied by

$$\forall l, l' \geq 1: \quad \lim_{h\to\infty} \sup_T \mathbb{P}\Big(|\langle \xi^{\lambda}_{T,m}, \chi_{ll'}\rangle|^2 > h\Big) = 0 \tag{12}$$

for which we moreover have

$$\mathbb{P}\Big(|\langle \xi^{\lambda}_{T,m}, \chi_{ll'}\rangle|^2 > h\Big) \leq \mathbb{P}\Big(\Re(\langle \xi^{\lambda}_{T,m}, \chi_{ll'}\rangle)^2 > h/2\Big) + \mathbb{P}\Big(\Im(\langle \xi^{\lambda}_{T,m}, \chi_{ll'}\rangle)^2 > h/2\Big).$$

Since the real and imaginary part of the random variables $\langle \xi^{\lambda}_{T,m}, \chi_{ll'}\rangle$ converge to real-valued random variables by Theorem 3.2, the corresponding sequence of probability measures is tight on $(\mathbb{R}, \mathscr{B})$. (18) therefore follows from the continuous mapping theorem. In order to verify the second condition of Lemma 3.2, note that by Markov's inequality it suffices to prove that

$$\lim_{k\to\infty} \sup_T \sum_{l,l':l+l'\geq k} \mathbb{E}|\xi^{\lambda}_{m,T}(\psi_{ll'})|^2 = 0. \tag{13}$$

Firstly, observe that $\mathbb{E}|\xi^{\lambda}_{m,T}(\psi_{ll'})|^2 \geq 0$ and $\langle \mathrm{Var}(\xi^{\lambda}_m)(\psi_{ll'}), (\psi_{ll'})\rangle \geq 0$. Note then from (11), that, for any $\chi_{ll'} \in H \otimes H$, as $T \to \infty$

$$\lim_{T\to\infty} \mathbb{E}|\xi^{\lambda}_{m,T}(\chi_{ll'})|^2 = \Gamma^{\lambda}_m(\chi_{ll'}) = \mathrm{Var}(\xi^{\lambda}_m(\chi_{ll'})) < \infty. \tag{14}$$

Together with Parseval's identity the monotone convergence and by definition of the (transpose) Kronecker tensor product, Theorem 3.2 implies

$$\begin{aligned}
\lim_{T\to\infty} \mathbb{E}\big[|||\xi^{\lambda}_{T,m}|||_2^2\big] &= \sum_{l,l'=1}^{\infty} \lim_{T\to\infty} \mathbb{E}\big[|\xi^{\lambda}_{T,m}(\psi_{ll'})|^2\big] \\
&= \sum_{l,l'=1}^{\infty} 4\pi^2 \Big(\mathscr{F}^{(\lambda)}_m(\psi_{ll})\overline{\mathscr{F}^{(\lambda)}_m(\psi_{l'l'})} + 1_{\{0,\pi\}}\big(\mathscr{F}^{(\lambda)}_m(\psi_{ll'})\mathscr{F}^{(\lambda)}_m(\psi_{ll'})\big)\Big) \\
&= 4\pi^2 \operatorname{Tr}\Big(\mathscr{F}^{(\lambda)}_m \widetilde{\otimes} \mathscr{F}^{(\lambda)}_m)\Big) + \operatorname{Tr}\Big(1_{\{0,\pi\}}\big(\mathscr{F}^{(\lambda)}_m \widetilde{\otimes}_{\top} \mathscr{F}^{(\lambda)}_m)\Big) = \mathbb{E}|||\xi^{\lambda}_m|||_2^2
\end{aligned}$$

From Proposition 3.2(iii.) we obtain immediately that

$$\mathbb{E}|||\xi^{\lambda}_m|||_2^2 = \operatorname{Tr}(\operatorname{Var}(\xi^{\lambda}_m)) < \infty. \tag{15}$$

Consequently, we can choose an $\epsilon > 0$ such that for all $k \geq k_0$

$$|\operatorname{Tr}(\operatorname{Var}(\xi^{\lambda}_m)) - \sum_{l+l' \leq k_0} \langle (\operatorname{Var}(\xi^{\lambda}_m))(\psi_{ll'}), (\psi_{ll'})\rangle| < \epsilon.$$

From the pointwise convergence (14) and from the sequence convergence in (15), we obtain

$$\begin{aligned}
\lim_{T\to\infty} \sum_{l,l':l+l'=k_0}^{\infty} \mathbb{E}|\xi^{\lambda}_{m,T}(\psi_{ll'})|^2 &= \lim_{T\to\infty} \Big(\sum_{l,l'=1}^{\infty} \mathbb{E}|\xi^{\lambda}_{m,T}(\psi_{ll'})|^2 - \sum_{l+l'<k_0} \mathbb{E}|\xi^{\lambda}_{m,T}(\psi_{ll'})|^2\Big) \\
&= \operatorname{Tr}(\operatorname{Var}(\xi^{\lambda}_m)) - \sum_{l+l'<k_0} \langle (\operatorname{Var}(\xi^{\lambda}_m))(\psi_{ll'}), (\psi_{ll'})\rangle.
\end{aligned}$$

In other words, there must exist a T_0 such that for all $T \geq T_0$ and $k \geq k_0$

$$|\mathbb{E}|||\xi^{\lambda}_{m,T}|||_2^2 - \sum_{l+l'<k} \mathbb{E}|\xi^{\lambda}_{m,T}(\psi_{ll'})|^2| < \epsilon.$$

Moreover, we can choose a $\tilde{k} \geq k_0$ such that for all $1 \leq T < T_0$,

$$\mathbb{E}|||\xi^{\lambda}_{m,T}|||_2^2 - \sum_{l+l'<\tilde{k}} \mathbb{E}|\xi^{\lambda}_{m,T}(\psi_{ll'})|_2^2 < \epsilon.$$

(13) now follows by taking $\max(k, \tilde{k})$. Therefore, we have established both conditions of Lemma 3.1, and thus $\xi^{\lambda}_{T,m} \Rightarrow_T \xi^{\lambda}_m$. Next, we will show that $\xi^{\lambda}_m \Rightarrow_m \breve{\mathscr{Q}}^{\lambda}$ where $\breve{\mathscr{Q}}^{\lambda}$ denotes the limiting process given in Theorem 3.1. We again verify the conditions of Lemma 3.1. From Theorem 3.2 and Proposition 3.2(iii.), we find

$$\begin{aligned}
\lim_{m\to\infty} \mathbb{E}|||\xi^{\lambda}_m|||_2^2 &= \lim_{m\to\infty} 4\pi^2 \operatorname{Tr}\Big(\mathscr{F}^{(\lambda)}_m \widetilde{\otimes} \mathscr{F}^{(\lambda)}_m + 1_{\{0,\pi\}}\big(\mathscr{F}^{(\lambda)}_m \widetilde{\otimes}_{\top} \mathscr{F}^{(\lambda)}_m)\Big) \\
&= 4\pi^2 \operatorname{Tr}\Big(\mathscr{F}^{(\lambda)} \widetilde{\otimes} \mathscr{F}^{(\lambda)} + 1_{\{0,\pi\}}\big(\mathscr{F}^{(\lambda)} \widetilde{\otimes}_{\top} \mathscr{F}^{(\lambda)})\Big) = \mathbb{E}|||\breve{\mathscr{Q}}^{\lambda}|||_2^2 < \infty.
\end{aligned} \tag{16}$$

Recall then that Theorem 3.2 shows that for fixed m, $\xi^{\lambda}_m(\chi)$, for any $\chi \in H \otimes H$ is a zero mean complex-valued Gaussian random variable. Hence $\xi^{\lambda}_m(\chi) \Rightarrow_m \breve{\mathscr{Q}}^{\lambda}(\chi)$ if we can show that the covariance structure satisfies

$$\begin{aligned}
&\lim_{m\to\infty} \Gamma_m(\chi_{ll'}) = \Gamma(\chi_{ll'}) \\
&\lim_{m\to\infty} \Sigma_m(\chi_{ll'}) = \Sigma(\chi_{ll'})
\end{aligned} \tag{17}$$

where Γ and Σ are the covariance and pseudocovariance operator given in Theorem 3.1. This however follows immediately from (16). Hence, $\xi_m^\lambda(\chi_{ll'}) \Rightarrow_m \breve{\mathscr{Q}}^\lambda(\chi_{ll'})$ showing the finite-dimensional distributions converge. Similar to (18) this implies that

$$\forall\, l, l' \geq 1: \quad \lim_{h\to\infty} \sup_m \mathbb{P}\Big(|\langle \xi_m^\lambda, \chi_{ll'}\rangle|^2 > h\Big) = 0. \tag{18}$$

Hence, condition Lemma 3.2(i) is satisfied. The tightness condition Lemma 3.2(ii) is satisfied if

$$\lim_{k\to\infty} \sup_m \sum_{l+l'\geq k} \mathbb{E}|\xi_m^\lambda(\psi_{ll'})|^2 = 0. \tag{19}$$

From the pointwise convergence (17) and the convergence of (16) as $m \to \infty$, this now however follows similarly to the proof of (19). Altogether, this establishes $\xi_{T,m}^\lambda \Rightarrow_T \xi_m^\lambda \Rightarrow_m \breve{\mathscr{Q}}^\lambda$. Finally, it remains to show $\xi_T^\lambda \Rightarrow_T \breve{\mathscr{Q}}^\lambda$, for which we make use of the following lemma.

Lemma 3.3. *Under the conditions of Theorem 3.1*

$$\lim_{m\to\infty} \limsup_{T\to\infty} \frac{1}{|||\Phi_T|||_F} \left\| \hat{\mathscr{Q}}_T^\lambda - \mathbb{E}\hat{\mathscr{Q}}_T^\lambda - \mathscr{M}_{T,m}^{(\lambda)} - \mathscr{M}_{T,m}^{\dagger(\lambda)} \right\|_{S_2,2} = 0. \tag{20}$$

The proof can be found in Appendix C. Since $S_2(H)$ is a complete metric space, let F be a closed set of $S_2(H)$ and fix $\epsilon > 0$. Then,

$$\mathbb{P}(\xi_T^\lambda \in F) \leq \mathbb{P}(|||\xi_{T,m}^\lambda - \xi_T^\lambda|||_2 \geq \epsilon) + \mathbb{P}\big(\xi_{T,m}^\lambda \in \{x : |||x-y|||_2 \leq \epsilon, y \in F\}\big)$$

and since by the weak convergence of $\xi_{T,m}^\lambda \Rightarrow_T \xi_m^\lambda \Rightarrow_m \breve{\mathscr{Q}}^\lambda$, we have

$$\lim_{m\to\infty} \limsup_{T\to\infty} \mathbb{P}\big(\xi_{T,m}^\lambda \in \{x : |||x-y|||_2 \leq \epsilon, y \in F\}\big) \leq \mathbb{P}\big(\breve{\mathscr{Q}}^\lambda \in \{x : |||x-y|||_2 \leq \epsilon, y \in F\}\big).$$

Using then Lemma 3.3, Markov's inequality yields

$$\limsup_{T\to\infty} \mathbb{P}(\xi_T^\lambda \in F) \leq \mathbb{P}\big(\breve{\mathscr{Q}}^\lambda \in \{x : |||x-y|||_2 \leq \epsilon, y \in F\}\big),$$

so that taking $\epsilon \to 0$, completes the proof. □

4 Estimation of the spectral density operator

In this section we focus on the application of the above theorem to estimate the spectral density operator

$$\mathscr{F}^{(\lambda)} = \frac{1}{2\pi} \sum_{h\in\mathbb{Z}} C_h e^{-\mathrm{i}\lambda h}.$$

Proofs of the statements in this section are postponed to Appendix D. It is well-known that under various conditions [see e.g., 12, 25] an asymptotically unbiased estimator is given by the periodogram operator

$$\mathscr{I}_T^\lambda := \mathscr{D}_T^\lambda \otimes \mathscr{D}_T^\lambda$$

where $\mathscr{D}_T^\lambda$ are the fDFT of X given in Section 2. Note that, by construction, this operator is hermitian, non-negative definite and $\lambda \mapsto \mathscr{I}_T^\lambda$ is 2π-periodic. From (6), we can immediately conclude that, under the stated conditions, the periodogram operator is indeed an asymptotically unbiased estimator of $\mathscr{F}^{(\lambda)}$. It can however never be consistent because it is based upon one frequency observation. A consistent estimator of the spectral density operator can be obtained via smoothing the operator-valued function $\lambda \mapsto \mathscr{I}_T^{\lambda}$ over neighboring frequency ordinates, i.e., via convolving the periodogram operator with a window function K. For example, it is very common to consider an estimator of the form

$$\hat{\mathscr{F}}^{\omega} = \frac{1}{b_T}\int_{-\infty}^{\infty} K\big(\frac{\omega-\lambda}{b_T}\big)\mathscr{D}_T^{\lambda}\otimes\mathscr{D}_T^{\lambda}d\lambda, \tag{21}$$

where $K:\mathbb{R}\to\mathbb{R}_+$ is assumed to be an even, non-negative weight function that is integrable. Under Assumption 3.1 with $p=4$, it is immediate from an application of Cauchy-Schwarz inequality and Lemma A.2(i) that $\sup_{\lambda}\|\mathscr{I}_T^{\lambda}\|_{S_2,2} = O(1)$ uniformly in T. By Holder's inequality, (21)therefore exists as en element of $\|\cdot\|_{S_2,2}$. In order to exploit the results from the previous section, we however require the estimator can be formulated in terms of a quadratic form. As remarked in the introduction, we consider

$$\hat{\mathscr{F}}^{\omega} = \frac{1}{2\pi T}\sum_{s,t=1}^{T}\big((X_s-\mu)\otimes(X_t-\mu)\big)w(b_T(t-s))e^{\mathrm{i}\omega(t-s)} \tag{22}$$

Note that $\hat{\mathscr{F}}^{\omega} = \frac{1}{2\pi T}\hat{\mathscr{Q}}_T^{\omega}$ with $\Phi_{T,t,s} = \phi^{\omega}_{T,(t-s)}I_{H\otimes H} = w(b_T(t-s))e^{\mathrm{i}\omega(t-s)}I_{H\otimes H}$ thus yields the representation in terms of the quadratic form introduced in the previous section. Provided $w(\cdot)$ and $K(\cdot)$ form Fourier pairs, there is a clear connection between (22) and (21). Namely, a change of variables gives

$$\begin{aligned}
\hat{\mathscr{F}}^{\omega} &= \frac{1}{b_T}\int_{-\infty}^{\infty} K\big(\frac{\omega-\lambda}{b_T}\big)\mathscr{D}_T^{\lambda}\otimes\mathscr{D}_T^{\lambda}d\lambda = \int_{-\infty}^{\infty}K(x)\mathscr{I}_T^{\omega+xb_T}dx\\
&= \int_{-\infty}^{\infty}K(x)\frac{1}{2\pi T}\sum_{s,t=1}^{T}e^{-\mathrm{i}(\omega+xb_T)(s-t)}(X_s\otimes X_t)dx\\
&= \frac{1}{2\pi T}\sum_{s,t=1}^{T}(X_s\otimes X_t)e^{-\mathrm{i}\omega(s-t)}\int_{-\infty}^{\infty}K(x)e^{\mathrm{i}xb_T(t-s)}dx\\
&= \frac{1}{2\pi T}\sum_{s,t=1}^{T}(X_s\otimes X_t)w(b_T(t-s))e^{\mathrm{i}\omega(t-s)}
\end{aligned}$$

where the equality is with respect to $\|\cdot\|_{S_2,2}$. In order to verify consistency and asymptotic normality, we shall require the following assumptions on the weight function $w(\cdot)$ in (22).

Assumption 4.1. *Let w be an even, bounded function on $\mathbb{R}$ with $\lim_{x\to 0} w(x) = 1$ that is continuous except at a finite number of points. Furthermore, suppose that $\lim_{b\to 0} b\sum_{h\in\mathbb{Z}} w^2(bh) = \kappa$ where $\kappa := \int_{\infty}^{\infty} w^2(x)dx < \infty$ such that $\sup_{0\le b\le 1} b\sum_{h\ge M/b} w^2(bh) \to 0$ as $M\to\infty$.*

Observe that these are rather mild conditions for window functions and includes a wide range of common choices [see e.g. 5]. Under these conditions we can obtain consistency in mean square of the spectral density operator.

Theorem 4.1. *Suppose Assumption 3.1 with $p=4$ and Assumption 4.1 are satisfied. Then,*

(i) $\sup_{\lambda\in[0,\pi]}\mathbb{E}|||\hat{\mathscr{F}}_T^{\lambda}-\mathscr{F}^{\lambda}|||_2^2 \to 0$ *if $b_T\to 0$ as $T\to\infty$ such that $b_T T\to\infty$.*

(ii) If, in addition, $\sum_{h\in\mathbb{Z}} h\|P_0(X_h)\|_{\mathbb{H},2} < \infty$ and $\lim_{x\to 0}|w(x)-1| = O(x)$, then

$$\mathbb{E}|||\hat{\mathscr{F}}_T^{\lambda}-\mathscr{F}^{\lambda}|||_2^2 = O(\frac{1}{b_T T}+b_T^2)$$

uniformly in $\lambda\in[0,\pi]$.

Note that Theorem 4.1 does not rely on a martingale approximation to exist but relies on the ergodicity properties of the underlying process. Without loss of generality, we can restrict to the interval $[0,\pi]$ since the mappings $\lambda \mapsto \hat{\mathscr{F}}_T^{\lambda}$ and $\lambda \mapsto \mathscr{F}_T^{\lambda}$ are even and 2π-periodic. Under Assumption 3.1 with $p=4$ and Assumption 4.1 we in fact obtain that $\sup_{\lambda\in[0,\pi]} \mathbb{E}|||\hat{\mathscr{F}}_T^{\lambda} - \mathbb{E}\mathscr{F}^{\lambda}|||_2^2 = O(\frac{1}{b_T T})$. It is however often of importance to obtain a specific rate of consistency and hence additionally to be able to control the order of the bias in norm. As given in the second part of the statement, this requires mild additional conditions on the smoothness of the process as well as a smoothness condition of the weight function around 0.

Remark 4.1 (**If the function μ is unknown**)**.** In case the mean function μ is unknown, we can instead consider the estimator

$$\hat{\hat{\mathscr{F}}}^{(\lambda)} = \frac{1}{2\pi T}\sum_{s,t=1}^{T}\big((X_s-\hat{\mu})\otimes(X_t-\hat{\mu})\big)w(b_T(t-s))e^{\mathrm{i}\lambda(t-s)} \tag{23}$$

where $\hat{\mu} = \frac{1}{T}\sum_{j=1}^{T} X_T$ denotes the sample mean function and which defines a random element of H. We obtain the following error bound with the estimator in (22), which shows the results in this section are not affected by centering the data using the sample mean.

Lemma 4.1. *Suppose Assumption 3.1 with $p=4$ is satisfied and Assumption 4.1 holds. Then*

$$\sup_{\lambda\in[0,\pi]} \mathbb{E}|||\hat{\hat{\mathscr{F}}}^{(\lambda)} - \hat{\mathscr{F}}^{(\lambda)}|||_2^2 = O((b_T T)^{-2}).$$

More generally, if $X \in \mathscr{L}_H^{2p}$, then $\|\hat{\hat{\mathscr{F}}}^{(\lambda)} - \hat{\mathscr{F}}^{(\lambda)}\|_{S_2,p} = O((b_T T)^{-1})$, $p \geq 1$.

The next result is the joint distributional convergence of a set of estimators at distinct frequencies to uncorrelated Gaussian elements of $S_2(H)$.

Theorem 4.2. *Suppose Assumption 3.1 with $p=4$ and Assumption 4.1 are satisfied. Let $\lambda_1,\ldots,\lambda_d \in [0,\pi]$ be distinct. Then, for $b_T \to 0$ such that $b_T T \to \infty$ as $T \to \infty$*

$$\sqrt{b_T T}\big(\hat{\mathscr{F}}^{\lambda_j} - \mathbb{E}\hat{\mathscr{F}}^{\lambda_j}\big)_{j=1,\ldots,d} \Rightarrow_T \big(\mathfrak{F}^{\lambda_j}\big)_{j=1,\ldots,d}$$

where $\mathfrak{F}^{\lambda_j}$, $j=1,\ldots,d$ are zero-mean jointly independent complex Gaussian elements of $S_2(H)$, with covariance operator

$$\mathrm{Cov}\big(\mathfrak{F}^{\lambda_j},\mathfrak{F}^{\lambda_j}\big) = 2\pi\kappa^2\Big(\mathscr{F}^{(\lambda_j)}\widetilde{\otimes}\mathscr{F}^{(\lambda_j)} + 1_{\{0,\pi\}}\mathscr{F}^{(\lambda_j)}\widetilde{\otimes}_{\top}\mathscr{F}^{(\lambda_j)}\Big)$$

and with pseudocovariance operator

$$\mathrm{Cov}\big(\mathfrak{F}^{\lambda_j},\overline{\mathfrak{F}^{\lambda_j}}\big) = 2\pi\kappa^2\Big(1_{\{0,\pi\}}\mathscr{F}^{(\lambda_j)}\widetilde{\otimes}\mathscr{F}^{(\lambda_j)} + \mathscr{F}^{(\lambda_j)}\widetilde{\otimes}_{\top}\mathscr{F}^{(\lambda_j)}\Big).$$

If the conditions of Theorem 4.1(ii) *are also satisfied, then*

$$\sqrt{b_T T}\big(\hat{\mathscr{F}}^{\lambda_j} - \mathscr{F}^{\lambda_j}\big)_{j=1,\ldots,d} \Rightarrow_T \big(\mathfrak{F}^{\lambda_j}\big)_{j=1,\ldots,d}$$

Observe that if $\lambda_j \in \{0,\pi\}$, then $\mathfrak{F}^{\lambda_j}$ is real Gaussian. Finally, we obtain the following corollary on the distributional properties of the estimator of the long run covariance operator.

Corollary 4.1. *Under the conditions of Theorem 4.2,*

$$\sqrt{b_T T}2\pi\big(\hat{\mathscr{F}}^{(0)} - \mathscr{F}^{(0)}\big) \Rightarrow_T \mathscr{N}_{S_2(H)}(0,4\pi^2\Gamma^{(0)})$$

where $\Gamma^{(0)} = 2\pi\kappa^2\Big(\mathscr{F}^{(0)}\widetilde{\otimes}\mathscr{F}^{(0)} + \mathscr{F}^{(0)}\widetilde{\otimes}_{\top}\mathscr{F}^{(0)}\Big)$.

Acknowledgements This work has been supported by the Collaborative Research Center "Statistical modeling of nonlinear dynamic processes" (SFB 823, Teilprojekt A1, A7, C1) of the German Research Foundation (DFG). The author kindly thanks Holger Dette and Davide Giraudo for some useful discussions.

References

[1] Bhansali, R., Giraitis, L., & Kokoszka, P. Approximations and limit theory for quadratic forms of linear variables stochastic processes and their applications. *Stochastic processes and their applications*, 117:71–95 (2007).

[2] Billingsley, P. *Convergence of Probability Measures.* Wiley, New York (1968).

[3] Bosq, D. Estimation of mean and covariance operator of autoregressive processes in Banach spaces *Statistical Inference for Stochastic Processes*, 5:287–306 (2002).

[4] Burkholder, D. L. Sharp inequalities for martingales and stochastic integrals. *Astérisque*, tome 157-158, p. 75–94 (1988).

[5] Brockwell, P. & Davis, R. *Time Series: Theory and Methods.* Springer, New York (1991).

[6] Cerovecki, C. & Hörmann, S. On the CLT for discrete Fourier transforms of functional time series. *Journal of multivariate analysis*, 154:282–295 (2017).

[7] Dauxois J., Pousse, A. & Romain, Y. Asymptotic theory for the principal component analysis of a vector random function: some applications to statistical inference. *Journal of multivariate analysis*, 12:136–154 (1982).

[8] Dehling H., Sharipov, O.S. Estimation of mean and covariance operator for Banach space valued autoregressive processes with dependent innovations *Statistical Inference for Stochastic Processes*, 2:137–149 (2005).

[9] van Delft, A., & Eichler, M. Locally stationary functional time series. *Electronic Journal of Statistics*, **12**(1), 107–170 (2018).

[10] van Delft, A. & Eichler, M. A note on Herglotz's Theorem for time series on the function space. *ArXiv Preprint*, arXiv:1801.04262 (2018).

[11] Grenander, U. *Abstract inference.* Wiley, New York (1981).

[12] Hörmann, S., Kidziński, L. & Hallin, M. Dynamic functional principal components. *The Journal of the Royal Statistical Society: Series B*, 77:319–348 (2015).

[13] Hörmann, S., & Kokoszka, P. Weakly dependent functional data. *The Annals of Statistics*, 38:1845–1884 (2010).

[14] Hörmann, S., Kokoszka, P. & Nisol, G. Detection of periodicity in functional time series. *The Annals of Statistics, forthcoming.* (2017).

[15] Hörvath, L., & Kokoszka, P. *Inference for Functional Data with Applications.* Springer, New York (2012).

[16] Hörvath, L., Kokoszka, P. & Reeder, R. Estimation of the mean of functional time series and a two-sample problem. *The Journal of the Royal Statistical Society: Series B*, 75:103–122 (2013).

[17] Gordin M. I. The central limit theorem for stationary processes. *Soviet. Math. Dokl.*, 10:1174–1176 (1969).

[18] de Jong, P. A central limit theorem for generalized quadratic forms. *Probability Theory and Related Fields*, 75:261–277 (1987).

[19] Kadison, R.V. & Ringorse, J. R. *Fundamentals of the Theory of Operator Algebras. Graduate Studies in Mathematics.* Amer. Math. Soc., Providence, RI (1997).

[20] Kokoszka, P., & Taqqu, M. S. The asymptotic behaviour of quadratic forms in heavy-tailed strongly dependent random variables. *Stochastic Processes and their Applications*, 66:21–40 (1997).

[21] Lee, J. & Subba Rao, S. A note one general quadratic forms of nonstationary stochastic proceses. *American Journal of Theoretical and Applied Statistics*, 51:949–968 (2017).

[22] Leucht, A., Paporoditis, E. and Sapatinas, T. Testing equality of spectral density operators for functional linear processes. *arXiv:1804.03366* (2018).

[23] Liu, W. & Wu, W. B. Asymptotics of spectral density estimates. *Econometric Theory*, 26:1218–1245 (2010).

[24] Mikosh, T. Functional limit theorems for random quadratic forms. *Stochastic processes and their applications*, 37:81–98 (1991).

[25] Panaretos, V. & S. Tavakoli. Fourier analysis of stationary time series in function space. *The Annals of Statistics* 41: 568–603 (2013).

[26] Peligrad, M. & Wu, W. B. Central limit theorem for Fourier transforms of stationary processes. *The Annals of Probability.*, 38(5):2009–2022 (2010).

[27] Pham, T. and Panaretos, V. Methodology and convergence rates for functional time series regression. *Statistica Sinica*, 28:2521–2539 (2018).

[28] Rice, J. & Silverman, B. Estimating the mean and covariance structure nonparametrically when the data are curves. *The Journal of the Royal Statistical Society: Series.*, 53:233–243 (1991).

[29] Suquet, C. Tightness in Schauder decomposable Banach spaces. *Translations–American Mathematical Society*, 193:201–224 (1996).

[30] Wu, W. B. & Shao, X. A Limit Theorem for Quadratic Forms and Its Applications. *Econometric Theory*, 23(5):930–951 (2007).

[31] Wu, W. B. Fourier transforms of stationary processes. *Proc. Amer. Math. Soc.*, 133:285–293 (2005).

[32] Wu, W. B. Nonlinear system theory: Another look at dependence. *Proceedings of the National Academy of Sciences.*, 102:14150–14154 (2005).

A Inequalities for H-valued martingales and linear transforms

Let H be a Hilbert space. For a probability space $(\Omega, \mathscr{A}, \mathscr{G}_\infty, \mathbb{P})$ and $\mathscr{G} = \{\mathscr{G}_t\}_{t\geq 0}$ a nondecreasing sequence of sub-σ-fields of $\mathscr{G}_\infty$, let $\{M_t\} \in \mathscr{L}_H^p$ be a martingale with respect to $\mathscr{G}$ and note that we can write $M_n = \sum_{k=0}^n D_k$, where $\{D_k\}$ denotes its difference sequence. Additionally denote the variable

$$V_n(M) = (\sum_k \|D_k\|_H^2)^{1/2})$$

which we call the square function of M. It was shown [4, theorem 3.1] that for H-valued martingales, we have for $1 < p < \infty$

$$(p^\star - 1)^{-1}(\mathbb{E}|V(M)|^p)^{1/p} \leq (\mathbb{E}\|M\|_H^p)^{1/p} \leq (p^\star - 1)(\mathbb{E}|V(M)|^p)^{1/p} \tag{24}$$

where $p^\star = \max(p, \frac{p}{p-1})$.

As a consequence we have the following lemma, which extends lemma 1 of [30].

Lemma A.1. *Let $M_{i=1,\ldots,n} \in \mathscr{L}_H^p$ be a martingale with respect to $\mathscr{G}$ with $\{D_k\}$ denoting its difference sequence and let $\{A\}_{k=1,\ldots,n} \in S_\infty(H)$. Then, for $q = min(2, p)$,*

$$\Big\|\sum_{k=1}^n A_k(D_k)\Big\|_{\mathbb{H},p}^q \leq K_p^q \sum_{k=1}^n |||A_k|||_\infty^q \|D_k\|_{\mathbb{H},p}^q$$

where $K_p^q = (p^\star - 1)2^{q(p-1)/p}$ with $p^\star = \max(p, \frac{p}{p-1})$.

Proof of Lemma A.1. By Burkholder's inequality (24)

$$\Big\|\sum_{k=1}^n A_k(D_k)\Big\|_{\mathbb{H},p}^q = \Big(\mathbb{E}\Big\|\sum_{k=1}^n A_k(D_k)\Big\|_H^p\Big)^{q/p} \leq (p^\star - 1)\Big(\mathbb{E}\Big|\Big(\sum_{k=1}^n \|A_k(D_k)\|_H^2\Big)^{1/2}\Big|^p\Big)^{q/p}$$

and therefore by Minkowski's inequality on the l_2 norm and consequently using that $|\sum_k \|D_k\|_H|^p \leq 2^{p-1}\sum_k \|D_k\|_H^p$ for $p > 1$

$$\begin{aligned}
(p^\star - 1)\Big(\mathbb{E}\Big|\Big(\sum_{k=1}^n \|A_k(D_k)\|_H^2\Big)^{1/2}\Big|^p\Big)^{q/p} &\leq (p^\star - 1)\Big(\mathbb{E}\Big|\sum_{k=1}^n \|A_k(D_k)\|_H\Big|^p\Big)^{q/p} \\
&\leq (p^\star - 1)2^{q(p-1)/p}\Big(\mathbb{E}\sum_{k=1}^n \|A_k(D_k)\|_H^p\Big)^{q/p} \\
&\leq (p^\star - 1)2^{q(p-1)/p}\Big(\sum_{k=1}^n |||A_k|||_\infty^p \mathbb{E}\|D_k\|_H^p\Big)^{q/p} \\
&\leq (p^\star - 1)2^{q(p-1)/p}\sum_k |||A_k|||_\infty^p \|D_k\|_{\mathbb{H},p}^q
\end{aligned}$$

where the one before last inequality follows from Holder's inequality for operators and where the last inequality follows from subadditivity of the function $(\cdot)^{p/q}$ in case $q/p < 1$. □

Lemma A.2. *For $t = 1,\ldots,n$, let $\{X_t\}$ be a zero-mean stationary ergodic process in $\mathscr{L}_H^p$ and $\{A_t\} \in S_\infty(H)$. Then,*

$$\text{(i)} \Big\|\sum_{t=1}^n A_t(X_t)\Big\|_{\mathbb{H},p}^q \leq K_p^q |||A_n|||_{\ell_q}^q \mathbf{\Delta}_{p,q,0}, \quad \text{(ii)} \Big\|\sum_{t=1}^n A_t X_t^{(m)}\Big\|_{\mathbb{H},p}^q \leq K_p^q |||A_n|||_{\ell_q}^q \mathbf{\Delta}_{p,q,0},$$

$$\text{(iii)} \Big\|\sum_{t=1}^n A_t(X_t - X_t^{(m)})\Big\|_{\mathbb{H},p}^q \leq K_p^q |||A_n|||_{\ell_q}^q \mathbf{\Delta}_{p,q,m+1}.$$

where $\mathbf{\Delta}_{p,q,m} = \sum_{j=m}^\infty \nu_{\mathbb{H},p,j}^q$ and $|||A_n|||_{\ell_q}^q = \sum_{t=1}^n |||A_t|||_\infty^q$.

Proof. Using (3) and Lemma A.1 (i) directly follows. For (ii), by stationarity

$$\|P_j(X_t^{(m)})\|_{\mathbb{H},p} = \|\mathbb{E}[X_{t-j} - X_{t-j,\{0\}}|\mathscr{G}_{t,t-m,\{0\}}]\|_{\mathbb{H},p} \le \nu_{\mathbb{H},p}(X_{t-j})$$

and therefore (ii) follows from (i). Finally, we can write $X_t - X_t^{\le m} = \sum_{j=1+m}^{\infty} \mathbb{E}[X_t|\mathscr{G}_{t,t-j}] - \mathbb{E}[X_t|\mathscr{G}_{t,t-j+1}]$ and we note that $D_{t,j} := \mathbb{E}[X_t|\mathscr{G}_{t,t-j}] - \mathbb{E}[X_t|\mathscr{G}_{t,t-j+1}]$ for $t = n,\ldots,1$ defines a martingale with respect to the backward filtration $\mathscr{G}(\epsilon_t,\ldots,\epsilon_i)$, $i = 0,-1,\ldots$. (iii) now follows from noting by the contraction property and stationarity

$$\|D_{t,j}\|_{\mathbb{H},p} = \|\mathbb{E}[(X_t - X_{t,\{t-j\}})|\mathscr{G}_{t,t-j}]\|_{\mathbb{H},p} \le \|(X_t - X_{t,\{t-j\}})\|_{\mathbb{H},p} \tag{25}$$

$$= \|(X_j - X_{j,\{0\}})\|_{\mathbb{H},p} = \nu_{\mathbb{H},p}(X_j). \tag{26}$$

□

Proof of Proposition 3.2. (i) It is clear that, since the process $\{X_t^{(m)}\}$ is m-dependent, the $D_{m,k}$ are m-dependent. Hence, we may write $D_{m,0}^{\lambda} = \sum_{t=0}^{m} P_0(X_t^{(m)})e^{-\mathrm{i}\lambda t}$. By orthogonality, $\mathbb{E}\|D_{m,k}\|_H^2 \le \sum_{t=0}^{\infty} \mathbb{E}\|P_0(X_t)\|_H^2 < \infty$. Next, observe that

$$\mathbb{E}[D_{m,k}^{(\lambda)}|\mathscr{G}_{k-1}] = \frac{1}{\sqrt{2\pi}} \sum_{t=0}^{\infty} \mathbb{E}[\mathbb{E}[X_{t+k}^{(m)})|\mathscr{G}_k] - \mathbb{E}[X_{t+k}^{(m)})|\mathscr{G}_{k-1}|\mathscr{G}_{k-1}]e^{-\mathrm{i}t\lambda} = 0$$

by the properties of the conditional expectation.
(ii) Under Assumption 3.1 with $p = 4$, we obtain from Lemma A.1

$$\mathbb{E}|||D_{m,k}^{(\lambda)} \otimes D_{m,k}^{(\lambda)}|||_2^2 = \mathbb{E}\|D_{m,k}^{(\lambda)}\|_H^4 \le \big(\sum_{t=0}^{\infty} \|P_0(X_t^{(m)})\|_{\mathbb{H},4}^2\big)^2$$

$$\le \big(\sum_{t=0}^{\infty} \|P_0(X_t)\|_{\mathbb{H},4}^2\big)^2 \le \big(\sum_{t=0}^{\infty} (\nu_{\mathbb{H},4}^2(X_t)\big)^2 < \infty$$

Secondly, observe that for all $n_1, n_2 \in \mathbb{N}$ such that $n_2 \ge n_1$, we have using Lemma A.1

$$\mathbb{E}\|D_{m,k,n_2}^{(\lambda)} - D_{m,k,n_2}^{(\lambda)}\|_H^4 = \big((\mathbb{E}\|D_{m,k,n_2}^{(\lambda)} - D_{m,k,n_1}^{(\lambda)}\|_H^4)^{1/2}\big)^2 = \big(\|D_{m,k,n_2}^{(\lambda)} - D_{m,k,n_1}^{(\lambda)}\|_{\mathbb{H},4}^2\big)^2 \le \big(\sum_{t=n_1+1}^{n_2} \|P_0(X_t)\|_{\mathbb{H},4}^2\big)^2$$

from which it is clear that $\{D_{\infty,k,T}^{(\lambda)}\}_{\{T\ge1\}}$ is Cauchy in $\mathscr{L}_H^4$. Trivially, $\{D_{m,k,T}^{(\lambda)}\}_{\{T\ge1\}}$ is therefore Cauchy in $\mathscr{L}_H^4$, uniformly in m. To ease notation, let $Y_n := D_{m,k,n}^{(\lambda)}$. Now observe that for all $n_1, n_2 \in \mathbb{N}$,

$$\begin{aligned}\mathbb{E}|||Y_{n_2} \otimes Y_{n_2} - Y_{n_1} \otimes Y_{n_1}|||_2^2 &\le 2\mathbb{E}|||(Y_{n_2} - Y_{n_1}) \otimes Y_{n_2}|||_2^2 + 2\mathbb{E}|||Y_{n_1} \otimes (Y_{n_2} - Y_{n_1})|||_2^2 \\ &\le 2\mathbb{E}\|(Y_{n_2} - Y_{n_1})\|_H^2\|Y_{n_2}\|_H^2 + 2\mathbb{E}\|Y_{n_1}\|_H^2\|(Y_{n_2} - Y_{n_1})\|_H^2 \\ &\le 2(\mathbb{E}\|(Y_{n_2} - Y_{n_1})\|_H^4\mathbb{E}\|Y_{n_2}\|_H^4)^{1/2} + 2(\mathbb{E}\|Y_{n_1}\|_H^4\mathbb{E}\,(Y_{n_2} - Y_{n_1})\|_H^4)^{1/2} \\ &\le 4(\epsilon N)^{1/2}\end{aligned}$$

where we used, that since$\{Y_n\}$ is Cauchy in $\mathscr{L}_H^4$, for all $\epsilon > 0$ there exist an N such that for $n_1, n_2 \ge N$, $\mathbb{E}\|(Y_{n_2} - Y_{n_1})\|_H^4 < \epsilon$ and $\mathbb{E}\|Y_n\|_H^4 < N$. Next we prove (iii). First we need to prove that

$$\lim_{m\to\infty}\lim_{T\to\infty} \mathrm{Tr}(\mathrm{Var}(D_{m,0,T}^{(\lambda)}) = \mathrm{Tr}(\mathscr{F}^{(\lambda)}) < \infty \tag{27}$$

Recall that $\mathrm{Tr}(\mathrm{Var}(D^{(\lambda)}_{m,0,T}) = \mathbb{E}\|D^{(\lambda)}_{m,0,T}\|^2_H$, where the latter is finite uniformly in m and T since the limit satisfies $\mathbb{E}\|D^{(\lambda)}_0\|^2_H < \infty$ by property (ii). We shall therefore proceed similar to [26, 6]. By stationarity and the integral of the complex exponential yielding the constraint $t-s=h$

$$\int_{-\pi}^{\pi} \mathbb{E}\|D^{(\omega)}_{m,0,T}\|^2 e^{\mathrm{i}h\omega}\,d\omega = \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathbb{E}\langle \sum_{t=0}^{T} P_0(X^{(m)}_t), \sum_{s=0}^{T} P_0(X^{(m)}_s)\rangle e^{-\mathrm{i}(t-s-h)\omega}\,d\omega$$
$$= \mathbb{E}\sum_{t=h}^{T}\langle P_0(X^{(m)}_t), P_0(X^{(m)}_{t-h})\rangle.$$

Since $\mathscr{G}_{-t} \subseteq \mathscr{G}_{-h} \forall t \geq h$, we remark that, for any $m \geq 1$, we have by the properties of the conditional expectation that $\mathbb{E}[\mathbb{E}[X^{(m)}_0|\mathscr{G}_{-h}]|\mathscr{G}_{-t}] \overset{\mathscr{L}^2_H}{=} \mathbb{E}[X^{(m)}_0|\mathscr{G}_{-t}], \forall t \geq h$. Morevover, $X^{(m)}_{-h}$ is $\mathscr{G}_{-h}$-measurable. Therefore, we obtain by orthogonality of the projection operators and stationarity that

$$\mathbb{E}\sum_{t=h}^{T}\langle P_0(X^{(m)}_t), P_0(X^{(m)}_{t-h})\rangle = \mathbb{E}\langle \sum_{t=h}^{T} P_0(X^{(m)}_t), \sum_{s=h}^{T} P_0(X^{(m)}_{s-h})\rangle$$
$$= \mathbb{E}\langle \sum_{t=h}^{T} P_{-t}(\mathbb{E}[X^{(m)}_0|\mathscr{G}_{-h}]), \sum_{s=h}^{T} P_{-s}(\mathbb{E}[X^{(m)}_{-h}|\mathscr{G}_{-h}])\rangle$$

By ergodicity and from (ii), $\{D^{(\lambda)}_{m,0,T}\}_{\{T\geq 1\}}$ is Cauchy in $\mathscr{L}^2_H$. Thus, $\lim_{T\to\infty}\sum_{t=h}^{T} P_{-t}(\mathbb{E}[X^{(m)}_0|\mathscr{G}_{-h}]) \overset{\mathscr{L}^2_H}{=} \mathbb{E}[X^{(m)}_0|\mathscr{G}_{-h}])$ and $\lim_{T\to\infty}\sum_{s=h}^{T} P_{-s}(\mathbb{E}[X^{(m)}_{-h}|\mathscr{G}_{-h}] \overset{\mathscr{L}^2_H}{=} X^{(m)}_{-h}$. Therefore, continuity of the inner product yields

$$\lim_{T\to\infty}\mathbb{E}\langle \sum_{t=h}^{T} P_{-t}(\mathbb{E}[X^{(m)}_0|\mathscr{G}_{-h}]), \sum_{s=h}^{T} P_{-s}(\mathbb{E}[X^{(m)}_{-h}|\mathscr{G}_{-h}])\rangle$$
$$= \mathbb{E}\langle \mathbb{E}[X^{(m)}_0|\mathscr{G}_{-h}]), X^{(m)}_{-h}\rangle = \mathbb{E}\langle X^{(m)}_0, X^{(m)}_{-h}\rangle = \mathrm{Tr}(C^m_h),$$

where we used the tower property. Hence, $\lim_{T\to\infty}\frac{1}{2\pi}\int_{-\pi}^{\pi}\mathbb{E}\|D^{(\lambda)}_{m,0,T}\|^2_H e^{\mathrm{i}h\lambda}\,d\lambda = \mathrm{Tr}(C^{(m)}_h)$. But this holds in particular for $m=\infty$, i.e., for the process $\lim_{m\to\infty} X^m_t = X_t$. Now observe that the conditions of the classical Féjer-Lebesgue theorem are satisfied and therefore

$$\lim_T \mathrm{Tr}(\mathrm{Var}(\mathscr{D}^{\lambda}_T)) = \lim_T \sum_{h\leq T}(1-\frac{h}{T})\mathbb{E}\langle X_h, X_0\rangle e^{-\mathrm{i}h\omega} = \mathbb{E}\|D^{(\lambda)}_0\|^2_H = \mathrm{Tr}(\mathscr{F}^{(\lambda)}) < \infty, \tag{28}$$

where we used again property (ii) in order to obtain the finite trace. Let $\mathscr{D}^{\omega}_{m,T}$ denotes the functional DFT of $X^{(m)}_t$. Clearly, we have immediately from the above as well that

$$\lim_{m\to\infty}\lim_{T\to\infty}\mathrm{Tr}(\mathrm{Var}(\mathscr{D}^{\omega}_{m,T}) = \lim_{m\to\infty}\mathbb{E}\|D^{(\lambda)}_{m,0}\|^2_H = \lim_{m\to\infty}\mathrm{Tr}(\mathscr{F}^{(\lambda)}_m) = \mathrm{Tr}(\mathscr{F}^{(\lambda)}) \tag{29}$$

where $2\pi\mathscr{F}^{(\lambda)}_m = \sum_{|h|\leq m}\mathbb{E}(X^{(m)}_h \otimes X^{(m)}_0)e^{-\mathrm{i}h\lambda}$ and where we applied the dominated convergence theorem which is justified by (28). This proves (27). Consequently, non-negative definiteness allows us to conclude that $\mathscr{F}^{(\lambda)}_m \in S_1(H)$ for all $m \geq 1$ and any $\lambda \in (-\pi,\pi]$. Then, using the permutation operator is a unitary operator, Holders' inequality for operators yields

$$\left|\left|\left|\Pi_{ijkl}\big(\mathscr{F}^{(\lambda)}\otimes\mathscr{F}^{(\lambda)}\big)\right|\right|\right|_1 \leq |||\Pi_{ikjl}|||_\infty |||\mathscr{F}^{(\lambda)}\widetilde{\otimes}\mathscr{F}^{(\lambda)}|||_1 \leq |||\mathscr{F}^{(\lambda)}|||^2_1 = (\mathbb{E}\|D^{(\lambda)}_0\|^2_H)^2 \leq \mathbb{E}\|D^{(\lambda)}_0\|^4_H < \infty, \tag{30}$$

where we applied (28) in the equality and Jensen's inequality together with property (ii) in the last inequality. From continuity of $\widetilde{\otimes}$, Π and the dominated convergence theorem together with (29), we obtain

$$\lim_{m\to\infty}\lim_{T\to\infty}\mathrm{Tr}\big(\Pi_{ijkl}\mathrm{Var}(D^{(\lambda)}_{m,0,T})\widetilde{\otimes}\mathrm{Var}(D^{(\lambda)}_{m,0,T})\big)=\mathrm{Tr}(\Pi_{ijkl}\mathscr{F}^{(\lambda)}\widetilde{\otimes}\mathscr{F}^{(\lambda)})<\infty.$$

□

B Joint convergence of finite-dimensional distributions of $\xi^{\lambda}_{T,m}$

Proof of Theorem 3.2. We recall that

$$\xi^{\lambda}_{T,m}:=(|||\Phi_T|||_F^2)^{-1/2}(\mathscr{M}^{(\lambda)}_{T,m}+\mathscr{M}^{\dagger(\lambda)}_{T,m}).$$

We want to show that $\{\xi^{\lambda_1}_{T,m},\dots,\xi^{\lambda_d}_{T,m}\}$ are converging jointly to complex Gaussian elements of $S_2(H)$. From Proposition 3.3, we know that $\xi^{\lambda_j}_{T,m}$ define martingales in $\mathscr{L}^2_{S_2(H)}(\Omega,\mathscr{A},\mathbb{P})$ with respect to the filtration $\{\mathscr{G}_T\}$. Below we shall prove convergence of the finite-dimensional distributions via a martingale central limit theorem on the linear combinations. To make this precise, let $U=\{u_1,\dots,u_d,v_1,\dots,v_d\in H\}$. For any $u,v\in H$ note that we can define the natural filtration of the process $\{\langle X_t,u\rangle\}_t$ over $(\Omega,\mathscr{A},\mathbb{P})$ by $\{\mathscr{G}_t(u)\}$. In the following, we let $\{\mathscr{G}_t(u_j,v_j)\}=\sigma(\{\langle X_t,u_j\rangle,\langle X_s,v_j\rangle\}_{t,s:t\geq s})$ to be the natural filtration over $(\Omega,\mathscr{A},\mathbb{P})$ of the projected process process $\{\langle X_t\otimes X_s,u_j\otimes v_j\rangle_S\}_{t,s:t\geq s}$. Correspondingly, denote $P_0^{(u_j,v_j)}=\mathbb{E}[\cdot|\mathscr{G}_0(u_j,v_j)]-\mathbb{E}[\cdot|\mathscr{G}_{-1}(u_j,v_j)]$ the projection operator. More generally, let $\mathscr{G}_t(U)=\sigma(\langle X_{t_1},u_1\rangle,\dots,\langle X_{t_d},u_d\rangle,\dots,\langle X_{t_{d-1}},v_1\rangle,\langle X_{t_{2d}},u\rangle\})$ for all $t=t_1\geq\dots\geq t_{2d}$ and P_0^U the corresponding projection operator. Observe then that

$$\frac{1}{|||\Phi_T|||_F}\Big(\langle\mathscr{M}^{(\lambda_j)}_{T,m}(u_j),v_j\rangle\rangle+\langle\mathscr{M}^{\dagger(\lambda_j)}_{T,m}(u_j),v_j\rangle\Big).$$

defines a well-defined martingale process in $\mathscr{L}^2_{\mathbb{C}}(\Omega,\mathscr{A},\mathbb{P})$ with respect to the filtration $\{\mathscr{G}_T(u_j,v_j)\}$. In order to derive joint convergence of the finite-dimensional distributions, it suffices to show, using the Cramér-Wold device that, for any $a_1,\dots,a_d\in\mathbb{R}$ and $\lambda_i\pm\lambda_j\neq 0 \mod 2\pi$, the process

$$\frac{1}{|||\Phi_T|||_F}\sum_{j=1}^d a_j\Big(\langle\mathscr{M}^{(\lambda_j)}_{T,m}(u_j),v_j\rangle\rangle+\langle\mathscr{M}^{\dagger(\lambda_j)}_{T,m}(u_j),v_j\rangle\Big).$$

converges to a zero-mean complex normal random variable with covariance

$$\sum_{j=1}^d a_j\langle\Gamma_m(u_j),v_j\rangle=4\pi^2\sum_{j=1}^d a_j\Big(\langle\mathscr{F}^{(\lambda_j)}_m(v_j),v_j\rangle\langle u_j,\mathscr{F}^{(\lambda_j)}_m(u_j)\rangle+1_{\{0,\pi\}}(\langle\mathscr{F}^{(\lambda_j)}_m(u_j),v_j\rangle\langle\mathscr{F}^{(\lambda_j)}_m(u_j),v\rangle)\Big)$$

and pseudocovariance

$$\sum_{j=1}^d a_j\langle\Sigma_m(u_j),v_j\rangle=4\pi^2\sum_{j=1}^d a_j\Big(1_{\{0,\pi\}}\big(\langle\mathscr{F}^{(\lambda_j)}_m(u_j),u_j\rangle\langle\mathscr{F}^{(\lambda_j)}_m(v_j),v_j\rangle\big)+\langle\mathscr{F}^{(\lambda_j)}_m(u_j),v_j\rangle\langle\mathscr{F}^{(\lambda_j)}_m(u_j),v_j\rangle\Big).$$

Note that this process is adapted to the filtration $\mathscr{G}_T(U)$. We start with decomposing the functional processes $\mathscr{M}^{(\lambda_j)}_{T,m}$ as follows.

$$\mathscr{M}^{(\lambda)}_{T,m}=\sum_{t=2}^T D^{(\lambda)}_{m,t}\otimes\Big(\sum_{s=1}^{t-4m}\phi^{(\lambda)}_{T,s-t}D^{(\lambda)}_{m,s})+\sum_{s=t-4m+1\vee 1}^{t-1}\phi^{(\lambda)}_{T,s-t}D^{(\lambda)}_{m,s}\Big)$$

The following lemma shows the second sum is of lower order in norm.

Lemma B.1. *Under the conditions of Theorem 3.1*

$$\Big\|\sum_{t=2}^{T} D_{m,t}^{(\lambda)} \otimes \big(\sum_{s=t-4m+1\vee 1}^{t-1} \phi_{T,s-t}^{(\lambda)} D_{m,s}^{(\lambda)}\big)\Big\|_{S_2,2} = o(|||\Phi_T|||_F).$$

This implies in turn that we can focus on the distributional properties of the projections of the operators

$$\sum_{t=4m+1}^{T} D_{m,t}^{(\lambda)} \otimes N_{m,t}^{(\lambda)} \quad \text{and} \quad \Big(\sum_{t=4m+1}^{T} D_{m,t}^{(\lambda)} \otimes N_{m,t}^{(\lambda)}\Big)^{\dagger} \tag{31}$$

where

$$N_{m,t}^{(\lambda)} := \sum_{s=1}^{t-4m} \phi_{T,s-t}^{(\lambda)} D_{m,s}^{(\lambda)}. \tag{32}$$

From Proposition 3.3, it is immediate that both terms in (31) constitute well-defined martingales in $\mathscr{L}^2_{S_2(H)}(\Omega, \mathscr{A}, \mathscr{G}_T, \mathbb{P})$. Consequently, projecting these on fixed $u, v \in H$, we obtain the following two martingale processes with paths in $\mathbb{C}$

$$\Big\langle \sum_{t=4m+1}^{T} D_{m,t}^{(\lambda)} \otimes N_{m,t}^{(\lambda)}, v \otimes u\Big\rangle_S = \sum_{t=4m+1}^{T} \langle D_{m,t}^{(\lambda)}, v\rangle \overline{\langle N_{m,t}^{(\lambda)}, u\rangle} \tag{33}$$

$$\Big\langle \Big(\sum_{t=4m+1}^{T} D_{m,t}^{(\lambda)} \otimes N_{m,t}^{(\lambda)}\Big)^{\dagger}, v \otimes u\Big\rangle_S = \sum_{t=4m+1}^{T} \overline{\langle D_{m,t}^{(\lambda)}, u\rangle} \langle N_{m,t}^{(\lambda)}, v\rangle \tag{34}$$

In order to apply a martingale central limit theorem on the sum of (33) and (34) and over $j = 1,\dots,d$, We must verify the Lindeberg condition is satisfied. Without loss of generality we do this for (33) and for fixed u, v as the result is immediate to carry over to a finite sum over j. To ease notation in the following, we set $\langle D_{m,t}^{(\lambda)}, x\rangle := D_{m,t}^{(\lambda)}(x)$ and $\langle N_{m,t}^{(\lambda)}, x\rangle := N_t^{(\lambda)}(x)$ for any $x \in H$. Recall the inequality $\mathbb{E}\{\|Y\|_H^2 1_{\|Y\|_H} > \epsilon\} \le \frac{1}{\epsilon^2}\mathbb{E}\|Y\|_H^4$ which holds for any $Y \in \mathscr{L}_H^4$. Hence applying this for $H = \mathbb{C}$, we have

$$\sum_{t=1+4m}^{T} \mathbb{E}\Big\{|D_{m,t}^{(\lambda)}(v)\overline{N_t^{(\lambda)}(u)}|^2 1_{|D_{m,t}^{(\lambda)}(v)\overline{N_t^{(\lambda)}(u)}|>\epsilon}\Big\} \le \frac{1}{\epsilon^2}\sum_{t=1+4m}^{T} \mathbb{E}|D_{m,t}^{(\lambda)}(v)\overline{N_t^{(\lambda)}(u)}|^4.$$

The Lindeberg condition is therefore satisfied if we can show that the term on the right hand side is of order $o(\|\phi_T\|_F^4)$. Since the $\{D_{m,t}^{(\lambda)}\}$ are m-dependent and by definition of (32) $|t-s| \ge 4m$, $D_{m,t}^{(\lambda)}$ and N_t are uncorrelated. Therefore, by Lemma A.1 with $H = \mathbb{C}$

$$\begin{aligned}
\sum_{t=1+4m}^{T} \mathbb{E}|D_{m,t}^{(\lambda)}(v)\overline{N_t^{(\lambda)}(u)}|^4 &\le \|D_0\|_{\mathbb{H},4}^4 \|v\|_H^4 \sum_{t=1+4m}^{T} (\|\overline{N_t^{(\lambda)}(u)}\|_{\mathbb{C},4}^2)^2 \\
&\le \|D_0\|_{\mathbb{H},4}^4 \|v\|_H^4 \sum_{t=1+4m}^{T} (K_4^2 |||\Phi_T|||_{\ell_2}^2 \|\|u\|_H^2 \|D_0\|_{\mathbb{H},4}^2)^2 \\
&= O(T\varrho_T^4) = o(|||\Phi_T|||_F^4),
\end{aligned} \tag{35}$$

and similarly for (34) showing that the Lindeberg condition is satisfied. It therefore remains to verify the that the conditional variance satisfies

$$\frac{1}{|||\Phi_T|||_F^2} \sum_{t=1+4m}^{T} \mathbb{E}\Big(\Big|\sum_{j=1}^{d} a_j\big(D_{m,t}^{(\lambda_j)}(v_j)\overline{N_t^{(\lambda_j)}(u_j)} + \overline{D_{m,t}^{(\lambda_j)}(u_j)} N_t^{(\lambda_j)}(v_j)\big)\Big|^2 \Big| \mathscr{G}_{t-1}^{(U)}\Big) \xrightarrow{p} \sum_{j=1}^{d} a_j \Gamma_m^{\lambda_j} \tag{36}$$

and that the conditional pseudocovariance satisfies

$$\frac{1}{\|\Phi_T\|_F^2}\sum_{t=1+4m}^{T}\mathbb{E}\Big(\Big(\sum_{j=1}^{d}a_j\big(D_{m,t}^{(\lambda_j)}(v_j)\overline{N_t^{(\lambda_j)}(u_j)}+\overline{D_{m,t}^{(\lambda_j)}(u_j)}N_t^{(\lambda_j)}(v_j)\big)\Big)^2\Big|\mathscr{G}_{t-1}^{(U)}\Big)\xrightarrow{p}\sum_{j=1}^{d}a_j\Sigma_m^{\lambda_j}. \tag{37}$$

Moreover, observe that we can write $\mathbb{E}(\cdot|\mathscr{G}_{t-1}^{(U)})=\sum_{k=1}^{m}P_{t-k}^{(U)}(\cdot)+\mathbb{E}(\cdot|\mathscr{G}_{t-1}^{(U)})$. We shall show that the sum of projections are of lower order. Applying this to (37), one finds using the orthogonality of the $P_j^{(U)}(\cdot)$ and the contraction property of the expectation

$$\begin{aligned}
&\Big\|\sum_{t=1+4m}^{T}\sum_{k=1}^{m}P_{t-k}^{(U)}\Big([\sum_{j=1}^{d}a_j\big(D_{m,t}^{(\lambda_j)}(v_j)\overline{N_t^{(\lambda_j)}(u_j)}+\overline{D_{m,t}^{(\lambda_j)}(u_j)}N_t^{(\lambda_j)}(v_j)\big)]^2\Big)\Big\|_{\mathbb{C},2}^2\\
&\le\Big(\sum_{k=1}^{m}\Big\|\sum_{t=1+4m}^{T}P_{t-k}^{(U)}\Big([\sum_{j=1}^{d}a_j(D_{m,t}^{(\lambda_j)}(v_j)\overline{N_t^{(\lambda_j)}(u_j)}+\overline{D_{m,t}^{(\lambda_j)}(u_j)}N_t^{(\lambda_j)}(v_j))]^2\Big)\Big\|_{\mathbb{C},2}\Big)^2\\
&\le\Big(\sum_{k=1}^{m}\Big(\sum_{t=1+4m}^{T}\Big\|\Big(\sum_{j=1}^{d}a_j\big(D_{m,t}^{(\lambda_j)}(v_j)\overline{N_t^{(\lambda_j)}(u_j)}+\overline{D_{m,t}^{(\lambda_j)}(u_j)}N_t^{(\lambda_j)}(v_j)\big)\Big)\Big\|_{\mathbb{C},4}^4\Big)^{1/2}\Big)^2\\
&=O\big(m^2\sum_{t=1+4m}^{T}\max_j\|D_{m,t}^{(\lambda_j)}(v_j)\|_{\mathbb{C},4}^4\|N_t^{(\lambda_j)}(u_j)\|_{\mathbb{C},4}^4\big)\\
&=o(\|\Phi_T\|_F^4).
\end{aligned}$$

where we used again that $D_{m,t}^{(\lambda)}$ and $N_t^{(\lambda)}$ are uncorrelated for any λ and where the order follows in a similar manner to (35). Furthermore, for any $x\in U$ and any λ, observe that $N_t^{(\lambda)}(x)$ is $\mathscr{G}_{t-5m}^{(U)}$ and $D_{m,t}^{(\lambda)}(x)$ is $\mathscr{G}_{t,t-m}^{(U)}$ measurable, the left-hand side of (36) equals

$$\begin{aligned}
&\frac{1}{\|\Phi_T\|_F^2}\sum_{t=1+4m}^{T}\mathbb{E}\Big(\Big|\sum_{j=1}^{d}a_j\big(D_{m,t}^{(\lambda_j)}(v_j)\overline{N_t^{(\lambda_j)}(u_j)}+\overline{D_{m,t}^{(\lambda_j)}(u_j)}N_t^{(\lambda_j)}(v_j)\big)\Big|^2|\mathscr{G}_{t-m-1}^{(U)}\Big)+o(1)\\
&=\frac{1}{\|\Phi_T\|_F^2}\sum_{t=1+4m}^{T}\sum_{j=1}^{d}a_j^2\Big(\overline{|N_t^{(\lambda_j)}(u_j)|^2}\mathbb{E}|D_{m,t}^{(\lambda_j)}(v_j)|^2+|N_t^{(\lambda_j)}(v_j)|^2\mathbb{E}\overline{|D_{m,t}^{(\lambda_j)}(u_j)|^2}\\
&\qquad+\overline{N_t^{(\lambda_j)}(u_j)N_t^{(\lambda_j)}(v_j)}\mathbb{E}D_{m,t}^{(\lambda_j)}(v_j)D_{m,t}^{(\lambda_j)}(u_j)+N_t^{(\lambda_j)}(v_j)N_t^{(\lambda_j)}(u_j)\mathbb{E}\overline{D_{m,t}^{(\lambda_j)}(u_j)D_{m,t}^{(\lambda_j)}(v_j)}\Big)\\
&\qquad+\sum_{i\neq j}a_ia_j\big(\overline{N_t^{(\lambda_i)}(u_i)}N_t^{(\lambda_j)}(u_j)\mathbb{E}[D_{m,t}^{(\lambda_i)}(v_i)\overline{D_{m,t}^{(\lambda_j)}(v_j)}]+N_t^{(\lambda_i)}(v_i)N_t^{(\lambda_j)}(u_j)\mathbb{E}[\overline{D_{m,t}^{(\lambda_i)}(u_i)}\overline{D_{m,t}^{(\lambda_j)}(v_j)}]\\
&\qquad+\overline{N_t^{(\lambda_i)}(u_i)}\overline{N_t^{(\lambda_j)}}(v_j)\mathbb{E}[D_{m,t}^{(\lambda_i)}(v_i)D_{m,t}^{(\lambda_j)}(u_j)]+N_t^{(\lambda_i)}(v_i)\overline{N_t^{(\lambda_j)}}(v_j)\mathbb{E}[\overline{D_{m,t}^{(\lambda_i)}(u_i)}D_{m,t}^{(\lambda_j)}(u_j)]\big)
\end{aligned}$$

while (37) becomes

$$\begin{aligned}
&\frac{1}{\|\Phi_T\|_F^2}\sum_{t=1+4m}^{T}\mathbb{E}\Big(\Big(\sum_{j=1}^{d}a_j\big(D_{m,t}^{(\lambda_j)}(v_j)\overline{N_t^{(\lambda_j)}(u_j)}+\overline{D_{m,t}^{(\lambda_j)}(u_j)}N_t^{(\lambda_j)}(v_j)\big)\Big)^2|\mathscr{G}_{t-m-1}^{(U)}\Big)+o(1)\\
&=\frac{1}{\|\Phi_T\|_F^2}\sum_{t=1+4m}^{T}\sum_{j=1}^{d}a_j^2\big(\overline{(N_t^{(\lambda_j)}(u_j))^2}\mathbb{E}(D_{m,t}^{(\lambda_j)}(v_j))^2+(N_t^{(\lambda_j)}(v_j))^2\mathbb{E}\overline{(D_{m,t}^{(\lambda_j)}(u_j))^2}\\
&\qquad+2\overline{N_t^{(\lambda_j)}(u_j)}N_t^{(\lambda_j)}(v_j)\mathbb{E}D_{m,t}^{(\lambda_j)}(v_j)\overline{D_{m,t}^{(\lambda_j)}(u_j)}\big)
\end{aligned}$$

$$+\sum_{i\neq j} a_i a_j\big(\overline{N_t^{(\lambda_i)}(u_i)}\,\overline{N_t^{(\lambda_j)}(u_j)}\mathbb{E}[D_{m,t}^{(\lambda_i)}(v_i)D_{m,t}^{(\lambda_j)}(v_j)] + N_t^{(\lambda_i)}(v_i)\overline{N_t^{(\lambda_j)}(u_j)}\mathbb{E}[\overline{D_{m,t}^{(\lambda_i)}(u_i)}D_{m,t}^{(\lambda_j)}(v_j)]$$
$$+\overline{N_t^{(\lambda_i)}(u_i)}N_t^{(\lambda_j)}(v_j)\mathbb{E}[D^{(\lambda_i)}_{m,t}(v_i)\overline{D^{(\lambda_j)}_{m,t}(u_j)}] + N_t^{(\lambda_i)}(v_i)N_t^{(\lambda_j)}(v_j)\mathbb{E}[\overline{D^{(\lambda_i)}_{m,t}(u_i)}\,\overline{D^{(\lambda_j)}_{m,t}(u_j)}]\big).$$

We shall make use of the following lemma.

Lemma B.2. *Let $\{D_{m,t}^{(\lambda)}\} \in \mathscr{L}_H^4$ be a H-valued martingale difference process. Then, provided that conditions (i) and (iv) of Assumption 3.2 are satisfied*

$$\Big\| \frac{1}{|||\Phi_T|||_F^2} \sum_{t=2}^{T} M_t^{(\lambda_1)} \otimes M_t^{(\lambda_2)} - \mathbb{E} M_0^{(\lambda_1)} \otimes \overline{M}_0^{(\lambda_2}\Big\|_{S_2,2} = o\Big(\frac{1}{|||\Phi_T|||_F^2}\Big).$$

where $M_t^{(\lambda)} := \sum_{s=1}^{t-1} \phi_{T,s-t}^{(\lambda)} D_{m,s}^{(\lambda)}$.

Since norm convergence implies convergence in the weak operator topology we obtain for any $u, v \in H$,

$$\Big\| \frac{1}{|||\Phi_T|||_F^2} \sum_{t=1+4m}^{T} N_t^{(\lambda_1)}(u) N_t^{(\lambda_2)}(v) - \mathbb{E} N_t^{(\lambda_1)}(u) N_t^{(\lambda_2)}(v) \Big\|_{\mathbb{C},2} = o(1).$$

Therefore, we may replace them with their expectation in (36) and (37) in order to obtain for (36)

$$\frac{1}{|||\Phi_T|||_F^2} \sum_{t=1+4m}^{T} \sum_{j=1}^{d} a_j^2\Big(\mathbb{E}|\overline{N_t^{(\lambda_j)}(u_j)}|^2 \mathbb{E}|D_{m,t}^{(\lambda_j)}(v_j)|^2 + \mathbb{E}|N_t^{(\lambda_j)}(v_j)|^2 \mathbb{E}|\overline{D_{m,t}^{(\lambda_j)}(u_j)}|^2$$
$$+ \mathbb{E}[\overline{N_t^{(\lambda_j)}(u_j) N_t^{(\lambda_j)}(v_j)}] \mathbb{E}[D_{m,t}^{(\lambda_j)}(v_j) D_{m,t}^{(\lambda_j)}(u_j)] + \mathbb{E}[N_t^{(\lambda_j)}(v_j) N_t^{(\lambda_j)}(u_j)] \mathbb{E}\overline{D_{m,t}^{(\lambda_j)}(u_j) D_{m,t}^{(\lambda_j)}(v_j)}\Big)$$
$$+ \sum_{i\neq j} a_i a_j\big(\mathbb{E}[\overline{N_t^{(\lambda_i)}(u_i)} N_t^{(\lambda_j)}(u_j)] \mathbb{E}[D_{m,t}^{(\lambda_i)}(v_i) \overline{D_{m,t}^{(\lambda_j)}(v_j)}] + \mathbb{E}[N_t^{(\lambda_i)}(v_i) N_t^{(\lambda_j)}(u_j)] \mathbb{E}[\overline{D_{m,t}^{(\lambda_i)}(u_i)}\,\overline{D_{m,t}^{(\lambda_j)}(v_j)}]$$
$$+ \mathbb{E}[\overline{N_t^{(\lambda_i)}(u_i)}\,\overline{N_t^{(\lambda_j)}}(v_j)] \mathbb{E}[D_{m,t}^{(\lambda_i)}(v_i) D_{m,t}^{(\lambda_j)}(u_j)] + \mathbb{E}[N_t^{(\lambda_i)}(v_i) \overline{N_t^{(\lambda_j)}(v_j)}] \mathbb{E}[\overline{D_{m,t}^{(\lambda_i)}(u_i)} D_{m,t}^{(\lambda_j)}(u_j)]\big) \tag{38}$$

and for (37)

$$\frac{1}{|||\Phi_T|||_F^2} \sum_{t=1+4m}^{T} \sum_{j=1}^{d} a_j^2\Big(\mathbb{E}(\overline{N_t^{(\lambda_j)}(u_j)})^2 \mathbb{E}(D_{m,t}^{(\lambda_j)}(v_j))^2 + \mathbb{E}(N_t^{(\lambda_j)}(v_j))^2 \mathbb{E}(\overline{D_{m,t}^{(\lambda_j)}(u_j)})^2$$
$$+ 2\mathbb{E}\overline{N_t^{(\lambda_j)}(u_j)} N_t^{(\lambda_j)}(v_j) \mathbb{E} D_{m,t}^{(\lambda_j)}(v_j) \overline{D_{m,t}^{(\lambda_j)}(u_j)}\Big)$$
$$+ \sum_{i\neq j} a_i a_j\big(\mathbb{E}[\overline{N_t^{(\lambda_i)}(u_i)}\,\overline{N_t^{(\lambda_j)}(u_j)}] \mathbb{E}[D_{m,t}^{(\lambda_i)}(v_i) D_{m,t}^{(\lambda_j)}(v_j)] + \mathbb{E}[N_t^{(\lambda_i)}(v_i) \overline{N_t^{(\lambda_j)}(u_j)}] \mathbb{E}[\overline{D_{m,t}^{(\lambda_i)}(u_i)} D_{m,t}^{(\lambda_j)}(v_j)]$$
$$+ \mathbb{E}[\overline{N_t^{(\lambda_i)}(u_i)} N_t^{(\lambda_j)}(v_j)] \mathbb{E}[D_{m,t}^{(\lambda_i)}(v_i) \overline{D_{m,t}^{(\lambda_j)}(u_j)}] + \mathbb{E}[N_t^{(\lambda_i)}(v_i) N_t^{(\lambda_j)}(v_j)] \mathbb{E}[\overline{D_{m,t}^{(\lambda_i)}(u_i)}\,\overline{D_{m,t}^{(\lambda_j)}(u_j)}]\big). \tag{39}$$

Next, we make use of the following auxiliary result.

Lemma B.3. *Let* $\{D_{m,t}^{(\lambda)}\} \in \mathcal{L}_H^2$ *be a* H*-valued martingale difference process and let conditions* (ii) *and* (iii) *in Assumption 3.2 be satisfied. Furthermore assume,* $\lambda_1 \pm \lambda_2 \neq 0 \mod 2\pi$*. Then, for any* $u, v \in H$,

$$\sum_{t=1+4m}^{T} |\mathbb{E}N_t^{(\lambda_1)}(u)N_t^{(\lambda_2)}(v)| = o(|||\Phi_T|||_F^2),$$

where $N_t^{(\lambda)}$ *is as defined in* (32).

Suppose first that $d = 1$. Then, if $\lambda \neq 0, \pi$, it follows from this lemma that the third and fourth term of (38) and the first two terms of (39)) will be of lower order if $\lambda \neq 0, \pi$. Hence, from Proposition 3.2

$$\mathbb{E}\langle D_{m,t}^{(\lambda)}, u\rangle \overline{\langle D_{m,t}^{(\lambda)}, v\rangle} = \sum_{|k|\leq m} \langle C_k^{(m)}(v), u\rangle e^{-\mathrm{i}\lambda k} = 2\pi\langle \mathscr{F}_m^{(\lambda)}(v), u\rangle.$$

If $\lambda = 0 \mod \pi$, we also have $\mathbb{E}\langle D_{m,t}^{(\lambda)}, u\rangle\langle D_{m,t}^{(\lambda)}, v\rangle = 2\pi\langle \mathscr{F}_m^{(\lambda)}(v), (u)\rangle$. Note that the latter is real for $v = u$. Hence, we obtain for (38)and (39)

$$\begin{aligned}
\frac{1}{|||\Phi_T|||_F^2} \sum_{t=1+4m}^{T} & \mathbb{E}|\overline{N_t^{(\lambda)}(u)}|^2 \mathbb{E}|D_{m,t}^{(\lambda)}(v)|^2 + \mathbb{E}|N_t^{(\lambda)}(v)|^2 \mathbb{E}|\overline{D_{m,t}^{(\lambda)}(u)}|^2 \\
& + \mathbb{E}\overline{N_t^{(\lambda)}(u)N_t^{(\lambda)}(v)}\mathbb{E}D_{m,t}^{(\lambda)}(v)D_{m,t}^{(\lambda)}(u) + \mathbb{E}N_t^{(\lambda)}(v)N_t^{(\lambda)}(u)\mathbb{E}\overline{D_{m,t}^{(\lambda)}(u)D_{m,t}^{(\lambda)}(v)} = \\
& = \frac{8\pi^2 \sum_{t=1+4m}^{T}\sum_{s=1}^{t-4m} w_{s-t}^2}{|||\Phi_T|||_F^2}\Big(\langle \mathscr{F}_m^{(\lambda)}(v), v\rangle\langle u, \mathscr{F}_m^{(\lambda)}(u)\rangle + 1_{\{0,\pi\}}(\langle \mathscr{F}_m^{(\lambda)}(v), u\rangle\langle \mathscr{F}_m^{(\lambda)}(v), u\rangle)\Big) \\
& \to 4\pi^2\Big(\langle \mathscr{F}_m^{(\lambda)}(v), v\rangle\langle u, \mathscr{F}_m^{(\lambda)}(u)\rangle + 1_{\{0,\pi\}}(\langle \mathscr{F}_m^{(\lambda)}(u), v\rangle\langle \mathscr{F}_m^{(\lambda)}(u), v\rangle)\Big),
\end{aligned}$$

and

$$\begin{aligned}
\frac{1}{|||\Phi_T|||_F^2} \sum_{t=1+4m}^{T} & \mathbb{E}(\overline{N_t^{(\lambda)}(u)})^2 \mathbb{E}(D_{m,t}^{(\lambda)}(v))^2 + \mathbb{E}(N_t^{(\lambda)}(v))^2 \mathbb{E}(\overline{D_{m,t}^{(\lambda)}(u)})^2 + 2\mathbb{E}\overline{N_t^{(\lambda)}(u)}N_t^{(\lambda)}(v)\mathbb{E}D_{m,t}^{(\lambda)}(v)\overline{D_{m,t}^{(\lambda)}(u)} \\
& \to 4\pi^2\Big(1_{\{0,\pi\}}\big(\langle \mathscr{F}_m^{(\lambda)}(u), u\rangle\langle \mathscr{F}_m^{(\lambda)}(v), v\rangle\big) + \langle \mathscr{F}_m^{(\lambda)}(u), v\rangle\langle \mathscr{F}_m^{(\lambda)}(u), v\rangle\Big),
\end{aligned}$$

respectively as $T \to \infty$, which completes the proof for $d = 1$. Next suppose that $d > 1$. If $\lambda_i \pm \lambda_j \neq 0 \mod 2\pi$, then by Lemma B.3, the cross terms are of lower order and from the case $d = 1$, we therefore obtain the convergence in (36) and (37). ☐

B.1 Proofs auxiliary statements

Proof of Lemma B.2. Let $Y_{T-1} = \sum_{t=2}^{T} M_{t-1}^{(\lambda_1)} \otimes M_{t-1}^{(\lambda_2)} - |||\Phi_T|||_F^2 \mathbb{E}M_0^{(\lambda_1)} \otimes M_0^{(\lambda_2)}$. Observe that from the properties of $\{M_{t-1}^{(\lambda)}\}$, the process Y_{T-1} is $\mathscr{G}_{T-1}$ measurable, stationary and ergodic. Therefore, by ergodicity of the underlying process a telescoping argument allows us to write

$$\begin{aligned}
\mathbb{E}|||Y_{T-1}|||_2^2 &= \mathbb{E}\Big|\Big|\Big|\sum_{j=-\infty}^{T-1} P_j\Big(\sum_{t=2}^{T} M_{t-1}^{(\lambda_1)} \otimes M_{t-1}^{(\lambda_2)}\Big)\Big|\Big|\Big|_2^2 \\
&= \sum_{j=-\infty}^{0} \mathbb{E}\Big|\Big|\Big|P_j\Big(\sum_{t=2}^{T} M_{t-1}^{(\lambda_1)} \otimes M_{t-1}^{(\lambda_2)}\Big)\Big|\Big|\Big|_2^2 + \mathbb{E}\sum_{j=1}^{T-1}\Big|\Big|\Big|P_j\Big(\sum_{t=2}^{T} M_{t-1}^{(\lambda_1)} \otimes M_{t-1}^{(\lambda_2)}\Big)\Big|\Big|\Big|_2^2. \qquad (40)
\end{aligned}$$

where we used orthogonality of the projection operators $\{P_j\}$. We first consider the first term on the right hand side for which we have $j < t$. Since $\{D_{m,s}^{(\lambda)}\}$ has uncorrelated increments

$$P_j(M_{t-1}^{(\lambda_1)} \otimes M_{t-1}^{(\lambda_2)}) = \sum_{s,s'=1}^{t-1} P_j(\phi_{T,s,t}^{(\lambda_1)} D_{m,s}^{(\lambda)} \otimes \phi_{T,s',t}^{(\lambda_2)} D_{m,s'}^{(\lambda)}) = \sum_{s=1}^{t-1} P_j(\phi_{T,s,t}^{(\lambda_1)} D_{m,s}^{(\lambda)} \otimes \phi_{T,s,t}^{(\lambda_2)} D_{m,s}^{(\lambda)})$$

$$= \sum_{s=1}^{t-1} \big(\phi^{(\lambda_1)}_{T,s,t} \widetilde{\otimes} \phi^{(\lambda_2)}_{T,s,t}\big) P_j (D^{(\lambda)}_{m,s} \otimes D^{(\lambda)}_{m,s})$$

where we used that linear operators and expecation operators commute, i.e., $\mathbb{E}(AX \otimes AX) = \mathbb{E}((A \widetilde{\otimes} A)(X \otimes X)) = (A \widetilde{\otimes} A)\mathbb{E}((X \otimes X))$ for $A \in S_\infty(H), X \in \mathscr{L}^4_H$. Consequently, by linearity and orthogonality of the projections, Minkowsk's inequality and stationarity of $\{D^{(\lambda)}_{m,s}\}$

$$\begin{aligned}
\sum_{j=-\infty}^{0} \mathbb{E}\Big\|\Big|P_j\Big(\sum_{t=2}^{T} M^{(\lambda_1)}_{t-1} \otimes M^{(\lambda_2)}_{t-1}\Big)\Big\|\Big|_2^2 &\le \sum_{j=-\infty}^{0} \Big(\sum_{t=2}^{T} \Big\|P_j\Big(M^{(\lambda_1)}_{t-1} \otimes M^{(\lambda_2)}_{t-1}\Big)\Big\|_{S_2,2}\Big)^2 \\
&= \sum_{j=-\infty}^{0} \Big(\sum_{t=2}^{T} \Big(\Big\|\sum_{s=1}^{t-1} (\phi^{(\lambda_1)}_{T,s,t} \widetilde{\otimes} \phi^{(\lambda_2)}_{T,s,t}) P_0 (D^{(\lambda)}_{m,s-j} \otimes D^{(\lambda)}_{m,s-j}\Big\|_{S_2,2}\Big)^2 \\
&= \sum_{j=-\infty}^{0} \Big(\sum_{t=2}^{T}\sum_{s=1}^{t-1} \Big\|\Big|\phi^{(\lambda_1)}_{T,s,t} \widetilde{\otimes} \phi^{(\lambda_2)}_{T,s,t}\Big\|\Big|_\infty \Big\|P_0(D^{(\lambda)}_{m,s-j} \otimes D^{(\lambda)}_{m,s-j}\Big\|_{S_2,2}\Big)^2 \\
&\le \sum_{j=-\infty}^{0} \Big(\sum_{s=1}^{T-1}\sum_{t=s+1}^{T} |||\phi^{(0)}_{T,s,t}|||^2_\infty \Big\|P_0(D^{(\lambda)}_{m,s-j} \otimes D^{(\lambda)}_{m,s-j}\Big\|_{S_2,2}\Big)^2
\end{aligned}$$

From an application of Cauchy Schwarz' inequality, we obtain under Assumption 3.1

$$\begin{aligned}
&\sum_{j=-\infty}^{0} \Big(\sum_{s=1}^{T-1} \Big\|P_0(D^{(\lambda)}_{m,s-j} \otimes D^{(\lambda)}_{m,s-j})\Big\|_{S_2,2} \sum_{t=s+1}^{T} |||\phi^{(0)}_{T,s,t}|||^2_\infty\Big)^2 \\
&\quad\le \sum_{j=-\infty}^{0} \Big(\Big[\sum_{s=1}^{T-1} \Big\|P_0(D^{(\lambda)}_{m,s-j} \otimes D^{(\lambda)}_{m,s-j})\Big\|^2_{S_2,2}\Big]^{1/2} \Big[\sum_{s=1}^{T-1}\Big(\sum_{t=s+1}^{T} |||\phi^{(0)}_{T,s,t}|||^2_\infty\Big)^2\Big]^{1/2}\Big)^2 \\
&\quad\le \sum_{s=1}^{T-1}\Big(\sum_{t=s+1}^{T} |||\phi^{(0)}_{T,s,t}|||^2_\infty\Big)^2 \sum_{j=-\infty}^{0}\sum_{s=1}^{T-1} \Big\|P_0(D^{(\lambda)}_{m,s-j} \otimes D^{(\lambda)}_{m,s-j})\Big\|^2_{S_2,2} \\
&\quad= o(|||\Phi_T|||^4_F).
\end{aligned}$$

For the second term of (40), i.e.,

$$\sum_{j=1}^{T-1} \mathbb{E}\Big\|\Big|P_j\Big(\sum_{t=2}^{T} M^{(\lambda_1)}_{t-1} \otimes M^{(\lambda_2)}_{t-1}\Big)\Big\|\Big|_2^2,$$

we have to distinguish cases since $1 \le j \le T-1$. Firstly observe that if $1 \le t \le j$, then $P_j(M^{(\lambda_1)}_{t-1} \otimes M^{(\lambda_2)}_{t-1}) = O_H$ since $M^{(\lambda_1)}_{t-1} \otimes M^{(\lambda_2)}_{t-1}$ is $\mathscr{G}_{t-1}$ measurable and hence $\mathbb{E}[M^{(\lambda_1)}_{t-1} \otimes M^{(\lambda_2)}_{t-1} | G_j] = M^{(\lambda_1)}_{t-1} \otimes M^{(\lambda_2)}_{t-1}$. We can focus on $t > j$. To ease notation, denote $D^{(\lambda)}_{m,s} := D_s$. Since expectation and tensor operator commute, we obtain for the various cases:

- if $s_1 \le j-1$:
  - $s_2 > j$, $\mathbb{E}[D_{s_1} \otimes D_{s_2} | \mathscr{G}_j] = D_{s_1} \otimes \mathbb{E}[D_{s_2} | \mathscr{G}_j] = O_H$ and similarly $\mathbb{E}[D_{s_1} \otimes D_{s_2} | \mathscr{G}_{j-1}] = O_H$ and therefore $P_j(D_{s_1} \otimes D_{s_2}) = O_H$.
  - $s_2 = j$: We have $\mathbb{E}[D_{s_1} \otimes D_{s_2} | \mathscr{G}_j] = D_{s_1} \otimes \mathbb{E}[D_{s_2} | \mathscr{G}_j] = D_{s_1} \otimes D_{s_2}$ while $\mathbb{E}[D_{s_1} \otimes D_{s_2} | \mathscr{G}_{j-1}] = O_H$. Hence, $P_j(D_{s_1} \otimes D_{s_2}) = D_{s_1} \otimes D_{s_2}$
  - $s_2 > j-1$: We have again $P_j(D_{s_1} \otimes D_{s_2}) = O_H$.
- If $s_1 > s_2 \ge j$: using the tower property, we have

$$\mathbb{E}[D_{s_1} \otimes D_{s_2} | \mathscr{G}_j] = \mathbb{E}[\mathbb{E}[D_{s_1} \otimes D_{s_2} | \mathscr{G}_{s_2}] | \mathscr{G}_j] = \mathbb{E}[\mathbb{E}[D_{s_1} | \mathscr{G}_{s_2}] \otimes D_{s_2} | \mathscr{G}_j] = O_H.$$

Hence,

$$P_j(M_{t-1} \otimes M_{t-1}) = \sum_{s_1=1}^{j-1} \big(\phi_{T,s_1,t} \widetilde{\otimes} \phi_{T,j,t}\big)(D_{s_1} \otimes D_j) + \sum_{s_2=1}^{j-1} \big(\phi_{T,j,t} \widetilde{\otimes} \phi_{T,s_2,t}\big)(D_j \otimes D_{s_2})$$
$$+ \sum_{s=j+1}^{t-1} \big(\phi_{T,s,t} \widetilde{\otimes} \phi_{T,s,t}\big) P_j(D^{(\lambda)}_{m,s} \otimes D^{(\lambda)}_{m,s}) =: U_j + U_j^{\dagger} + V_j$$

and therefore

$$\sum_{j=1}^{T-1} \mathbb{E}\Big\|\Big| P_j\Big(\sum_{t=2}^{T} M_{t-1} \otimes M_{t-1}\Big)\Big\|\Big|_2^2 = \sum_{j=1}^{T-1} \mathbb{E}\Big\|\Big| P_j\Big(\sum_{t=j+1}^{T} M_{t-1} \otimes M_{t-1}\Big)\Big\|\Big|_2^2 \leq \sum_{j=1}^{T-1} \mathbb{E}\Big\|\Big| \sum_{t=j+1}^{T} U_j + U_j^{\dagger} + V_j\Big)\Big\|\Big|_2^2.$$

For the first term, orthogonal increments and stationarity of $\{D^{(\lambda)}_{m,s}\}$, the properties of $\widetilde{\otimes}$ and Lemma A.1 yield

$$\begin{aligned}
\sum_{j=1}^{T-1} \mathbb{E}\Big\|\Big| \sum_{t=j+1}^{T} U_j\Big)\Big\|\Big|_2^2 &= \sum_{j=1}^{T-1} \mathbb{E}\Big\|\Big| \sum_{t=j+1}^{T} \sum_{s_1=1}^{j-1} \phi_{T,s_1,t} \widetilde{\otimes} \phi_{T,j,t}(D_{s_1} \otimes \overline{D}_j)\Big)\Big\|\Big|_2^2 \\
&= \sum_{j=1}^{T-1} \mathbb{E}\Big\|\Big| \sum_{s_1=1}^{j-1} \sum_{t=j+1}^{T} \big(\phi_{T,s_1,t} \widetilde{\otimes} \phi_{T,j,t}\big)(D_{s_1} \otimes \overline{D}_j)\Big)\Big\|\Big|_2^2 \\
&= \sum_{j=1}^{T-1} \mathbb{E}\Big\|\Big| \sum_{s_1=1}^{j-1} \sum_{t=j+1}^{T} \phi_{T,s_1,t}(D_{s_1}) \otimes \phi_{T,j,t}(D_j)\Big)\Big\|\Big|_2^2 \\
&= \sum_{j=1}^{T-1} K_2^2 \sum_{s_1=1}^{j-1} \Big\| \sum_{t=j+1}^{T} \phi_{T,s_1,t}\Big\|_\infty^2 \mathbb{E}\|D_{s_1}\|_H^2 |||\phi_{T,j,t}|||_\infty^2 \mathbb{E}\|D_j\|_H^2 \\
&= K_2^2 \sum_{j=1}^{T-1} \sum_{s_1=1}^{j-1} \Big\| \sum_{t=j+1}^{T} \phi_{T,s_1,t} \widetilde{\otimes} \phi_{T,j,t}\Big\|_\infty^2 \|D_0\|_{\mathbb{H},2}^2 \|D_0\|_{\mathbb{H},2}^2 \\
&= o(|||\Phi_T|||_F^4)
\end{aligned}$$

which follows from Assumption 3.2(iv). The same order applies to the second term. For the third term we find

$$\begin{aligned}
\sum_{j=1}^{T-1} \Big\| \sum_{t=j+1}^{T} V_j \Big\|_{S_2,2} &= \sum_{j=1}^{T-1} \Big\| \sum_{s=j+1}^{T-1} \sum_{t=s+1}^{T} \big(\phi_{T,s,t} \widetilde{\otimes} \phi_{T,s,t}\big) P_0(D^{(\lambda)}_{m,s-j} \otimes D^{(\lambda)}_{m,s-j})\Big\|_{S_2,2} \\
&= \sum_{j=1}^{T-1} \sum_{s=j+1}^{T-1} \sum_{t=s+1}^{T} |||\phi_{T,s,t}|||_\infty^2 \Big\| P_0(D^{(\lambda)}_{m,s-j} \otimes D^{(\lambda)}_{m,s-j})\Big\|_{S_2,2}
\end{aligned}$$

For convenience denote $A_s = \sum_{t=s+1}^{T} |||\phi_{T,s,t}|||_\infty^2$ and $\vartheta_{s-j} = \|P_0(D^{(\lambda)}_{m,s-j} \otimes D^{(\lambda)}_{m,s-j}\|_{S_2,2}$. Then we split the sum over j in a sum with terms $1, \ldots T-1-k$ and with terms $T-k, \ldots, T-1$ where $k = \lfloor T^{0.25} \rfloor$. Additionally, we split the inner sum of the first. We then find via tedious calculations

$$\begin{aligned}
\sum_{j=1}^{T-1}\Big(\sum_{s=j+1}^{T-1} A_s \vartheta_{s-j}\Big)^2 &\leq \sum_{j=T-k}^{T-1}\Big(\sum_{s=j+1}^{T-1} A_s \vartheta_{s-j}\Big)^2 + 2\sum_{j=1}^{T-1-k}\Big(\sum_{s=j+1}^{j+k-1} A_s \vartheta_{s-j}\Big)^2 + 2\sum_{j=1}^{T-1-k}\Big(\sum_{s=j+k}^{T-1} A_s \vartheta_{s-j}\Big)^2 \\
&\leq \sum_{j=T-k}^{T-1} \sum_{s=j+1}^{T-1} A_s^2 \sum_{s=j+1}^{T-1} \vartheta_{s-j}^2 + 2\sum_{j=1}^{T-1-k} \sum_{s=j+1}^{j+k-1} A_s^2 \sum_{s=j+1}^{j+k-1} \vartheta_{s-j}^2 + 2\sum_{j=1}^{T-1-k} \sum_{s=j+k}^{T-1} A_s^2 \sum_{s=j+k}^{T-1} \vartheta_{s-j}^2 \\
&\leq \sum_{s=j+1}^{T-1} A_s^2 \sum_{j=T-k}^{T-1} \sum_{s=1}^{T-j-1} \vartheta_s^2 + 2\sum_{j=1}^{T-1-k} \sum_{s=j+1}^{j+k-1} A_s^2 \sum_{s=1}^{k-1} \vartheta_s^2 + 2\sum_{j=1}^{T-1-k} \sum_{s=j+k}^{T-1} A_s^2 \sum_{s=k}^{T-1-j} \vartheta_s^2
\end{aligned}$$

$$\leq \sum_{s=1}^{T-1} A_s^2 k \sum_{s=1}^{k-1} \vartheta_s^2 + 2 \sum_{j=1}^{T-1-k} \sum_{s=j+1}^{j+k-1} A_s^2 \sum_{s=1}^{k-1} \vartheta_s^2 + 2 \sum_{j=1}^{T-1-k} \sum_{s=j+k}^{T-1} A_s^2 \sum_{s=k}^{T-1} \vartheta_s^2$$

$$\leq \sum_{s=1}^{T-1} A_s^2 k \sum_{s=1}^{k-1} \vartheta_s^2 + 2 \sum_{j=1}^{T-1-k} \sum_{s=j+1}^{j+k-1} A_s^2 \sum_{s=1}^{k-1} \vartheta_s^2 + 2 \sum_{j=1}^{T-1-k} \sum_{s=j+k}^{T-1} A_s^2 \sum_{s=k}^{T-1} \vartheta_s^2$$

$$= O(T^{1/4} \sum_{s=1}^{T-1} A_s^2) + O(T(\sum_{s=1}^{T-1} A_s^2)^{0.25}) + O(T \sum_{s=1}^{T-1} A_s^2 \mathbf{\Delta}_{4,2,k}) = o(\|\!|\Phi_T|\!\|_F^4).$$

□

Proof of Lemma B.1. By orthogonality of the martingale processes, Lemma A.1 and Jensen's inequality, we obtain for fixed m,

$$\Big\|\sum_{t=2}^{T} D_{m,t}^{(\lambda)} \otimes \big(\sum_{s=t-4m+1\vee 1}^{t-1} \phi_{T,s-t}^{(\lambda)} D_{m,s}^{(\lambda)}\big)\Big\|_{S_2,2}^2 \leq \sum_{t=2}^{T} \mathbb{E}\|D_{m,t}^{(\lambda)}\|_H^2 \mathbb{E}\Big\| \sum_{s=t-4m+1\vee 1}^{t-1} \phi_{T,s-t}^{(\lambda)} D_{m,s}^{(\lambda)}\Big\|_H^2$$

$$\leq K_2^2 \|D_{m,0}^{(\lambda)}\|_{\mathbb{H},2}^2 \sum_{t=2}^{T} \sum_{s=t-4m+1\vee 1}^{t-1} \|\!|\phi_{T,s-t}^{(\lambda)}|\!\|_\infty^2 \|D_{m,s}^{(\lambda)}\|_{\mathbb{H},2}^2$$

$$= 2K_4^2 \|D_{m,0}^{(\lambda)}\|_{\mathbb{H},4}^4 \Big(\sum_{t=2}^{4m} \sum_{s=1}^{t-1} \|\!|\phi_{T,s-t}^{(\lambda)}|\!\|_\infty^2 + \sum_{t=4m+1}^{T} \sum_{s=t-4m+1}^{t-1} \|\!|\phi_{T,s-t}^{(\lambda)}|\!\|_\infty^2\Big)$$

$$\leq 2K_4^2 \|D_{m,0}^{(\lambda)}\|_{\mathbb{H},4}^4 O\Big((4m)\varrho_{t-1}^2 + T4m (\max_t \|\!|A_{T,t}|\!\|_\infty^2 \Big) = o(\|\!|\Phi_T|\!\|_F^2) + To(\varrho_T^2) = o(\|\!|\Phi_T|\!\|_F^2).$$

□

Proof of Lemma B.3. Denote $\lambda = \lambda_1 \pm \lambda_2$ and recall that $N_{m,t}^{(\lambda)} = \sum_{s=1}^{t-4m} \phi_{T,s-t}^{(\lambda)} D_{m,s}^{(\lambda)}$. Since the increments of $\{D_{m,s}^{(\lambda)}\}$ are uncorrelated, we have

$$\sum_{t=1+4m}^{T} \Big|\mathbb{E}\langle N_{m,t}^{(\lambda_1)}, u\rangle \langle N_{m,t}^{(\lambda_2)}, v\rangle\Big| = \sum_{t=1+4m}^{T} \Big| \sum_{s=1}^{t-4m} e^{\mathrm{i}\lambda(s-t)} \mathbb{E}\langle A_{T,s-t}(D_{m,s}^{(\lambda_1)}), u\rangle \langle A_{T,s-t}(D_{m,s}^{(\lambda_2)}), v\rangle \Big|$$

$$= \sum_{t=1+4m}^{T} \Big| \sum_{s=1}^{t-4m} e^{\mathrm{i}\lambda(s-t)} \mathbb{E}\Big\langle A_{T,s-t}(D_{m,s}^{(\lambda_1)}) \otimes \overline{A_{T,s-t}(D_{m,s}^{(\lambda_2)})}, u \otimes \overline{v}\Big\rangle_S \Big|$$

$$= \sum_{t=1+4m}^{T} \Big| e^{-\mathrm{i}\lambda t} \sum_{s=1}^{t-4m} e^{\mathrm{i}\lambda s} \Big\langle (A_{T,s-t} \widetilde{\otimes} \overline{A}_{T,s-t}) \mathbb{E}(D_{m,0}^{(\lambda_1)} \otimes \overline{D_{m,0}^{(\lambda_2)}}), u \otimes \overline{v}\Big\rangle_S \Big|$$

$$\leq \sum_{t=1+4m}^{T} \Big| \sum_{s=1}^{t-4m} e^{\mathrm{i}\lambda s} \Big\langle (A_{T,s-t} \mathbb{E}(D_{m,0}^{(\lambda_1)} \otimes \overline{D_{m,0}^{(\lambda_2)}}) \overline{A}_{T,s-t}^{\dagger}), u \otimes \overline{v}\Big\rangle_S \Big|$$

To ease notation, set $W_{s-t} = \Big\langle (A_{T,s-t} \widetilde{\otimes} \overline{A}_{T,s-t}) \mathbb{E}(D_{m,0}^{(\lambda_1)} \otimes \overline{D_{m,0}^{(\lambda_2)}}), u \otimes \overline{v}\Big\rangle$ and write $B_j = \sum_{k=1}^{j} e^{\mathrm{i}\lambda j}$. Summation by parts, and Holder's inequality for bounded operators yield

$$\sum_{t=1+4m}^{T} \Big| \sum_{s=1}^{t-4m} W_{s-t}(B_s - B_{s-1}) \Big| = \sum_{t=1+4m}^{T} W_{t-4m} B_{t-4m} + \sum_{t=1+4m}^{T} \sum_{s=1}^{t-4m-1} B_s \big(W_{s-t} - W_{s-1-t}\big)$$

$$\leq \sum_{t=1+4m}^{T} |W_{t-4m} B_{t-4m}| + \sum_{t=1+4m}^{T} \sum_{s=1}^{t-4m-1} |B_s| \Big|\!\Big|\!\Big| \Big(A_{T,s-t} \widetilde{\otimes} (\overline{A}_{T,s-t} - \overline{A}_{T,s-t-1}) \mathbb{E}(D_{m,0}^{(\lambda_1)}) \otimes \overline{D_{m,0}^{(\lambda_2)}}\Big) \Big|\!\Big|\!\Big|_2 \|\!|u \otimes \overline{v}|\!\|_2$$

$$+ \sum_{t=1+4m}^{T} \sum_{s=1}^{t-4m-1} |B_s| \Big|\!\Big|\!\Big| \Big((A_{T,s-t} - A_{T,s-t-1}) \widetilde{\otimes} \overline{A}_{T,s-t-1} \Big) \mathbb{E}(D_{m,0}^{(\lambda_1)}) \otimes \overline{D_{m,0}^{(\lambda_2)}}\Big) \Big|\!\Big|\!\Big|_2 \|\!|u \otimes \overline{v}|\!\|_2$$

$$\leq \sum_{t=1+4m}^{T} |W_{t-4m} B_{t-4m}| + \sum_{t=1+4m}^{T} \sum_{s=1}^{t-4m-1} |B_s| \|\!|A_{T,s-t}|\!\|_\infty \|\!|\overline{A}_{T,s-t} - \overline{A}_{T,s-t-1}|\!\|_\infty \|\!|\mathbb{E}(D_{m,0}^{(\lambda_1)}) \otimes \overline{D_{m,0}^{(\lambda_2)}})|\!\|_2 \|u\|_H \|\overline{v}\|_H$$

$$+\sum_{t=1+4m}^{T}\sum_{s=1}^{t-4m-1}|B_s|\|\|A_{T,s-t}-A_{T,s-t-1}\|\|_\infty\|\|\overline{A}_{T,s-t-1}\|\|_\infty\|\|\mathbb{E}(D_{m,0}^{(\lambda_1)}\otimes\overline{D_{m,0}^{(\lambda_2)}})\|\|_2\|u\|_H\|\overline{v}\|_H$$

Then, using Jensen's inequality and the Cauchy Schwarz inequality twice, we obtain

$$\le C_1\|D_{m,0}^{(\lambda_1)}\|_{\mathbb{H},2}\|\overline{D_{m,0}^{(\lambda_2)}}\|_{\mathbb{H},2}|1/\sin(\lambda/2)|\Big(\sum_{t=1+4m}^{T}\|\|A_{t-4m}\|\|_\infty^2+\sum_{t=1+4m}^{T}\sum_{s=1}^{t-4m-1}\|\|A_{T,s-t}\|\|_\infty\|\|\overline{A}_{T,s-t}-\overline{A}_{T,s-t-1}\|\|_\infty$$
$$+\sum_{t=1+4m}^{T}\sum_{s=1}^{t-4m-1}\|\|A_{T,s-t}-A_{T,s-t-1}\|\|_\infty\|\|\overline{A}_{T,s-t-1}\|\|_\infty\Big)$$
$$\le C_1C_2|1/\sin(\lambda/2)|\Big(\sum_{t=1+4m}^{T}\|\|A_{t-4m}\|\|_\infty^2+\sum_{t=1+4m}^{T}\Big(\sum_{s=1}^{t-4m-1}\|\|A_{T,s-t}\|\|_\infty^2\sum_{s=1}^{t-4m-1}\|\|\overline{A}_{T,s-t}-\overline{A}_{T,s-t-1}\|\|_\infty^2\Big)^{1/2}$$
$$+\sum_{t=1+4m}^{T}\Big(\sum_{s=1}^{t-4m-1}\|\|A_{T,s-t}-A_{T,s-t-1}\|\|_\infty^2\sum_{s=1}^{t-4m-1}\|\|\overline{A}_{T,s-t-1}\|\|_\infty^2\Big)^{1/2}\Big)$$
$$\le O(|1/\sin(\lambda/2)|)\Big(O(\varrho_T^2)+O(T)o(\varrho_T)O(\varrho_T)+O(T)O(\varrho_T)o(\varrho_T)\Big)=o(\|\|\Phi_T\|\|_F^2),$$

where we used that $\max_{1\le t\le T}|B_t|\le|1/(\sin(\lambda/2))|$. □

C Operator approximations

Proof of Lemma 3.3. We can decompose the quadratic form

$$\hat{\mathscr{Q}}_T^\lambda=\mathcal{V}_T^\lambda+\mathcal{V}_T^{\dagger\lambda}+\sum_{1\le t\le T}\phi_{T,t,t}(X_t\otimes X_t)$$

Set $\hat{\mathscr{C}}_T=\sum_{t=1}^T X_t\otimes X_t$. Then, using linearity of the operator $\Phi_{T,t,t}$

$$\Big\|\hat{\mathscr{Q}}_T^\lambda-\mathbb{E}\hat{\mathscr{Q}}^\lambda\Big\|_{S_2,2}\le\Big\|\mathcal{V}_T^\lambda+\mathcal{V}_T^{\dagger\lambda}-\mathbb{E}\mathcal{V}_T^\lambda-\mathbb{E}\mathcal{V}_T^{\dagger\lambda}\Big\|_{S_2,2}+\Big\|\Phi_{T,t,t}(\hat{\mathscr{C}}_T-\mathbb{E}\hat{\mathscr{C}}_T)\Big\|_{S_2,2},$$

For the last term, Holder's inequality for linear operators,Lemma C.1

$$\frac{1}{\|\|\Phi_T\|\|_F}\Big\|\Phi_{T,t,t}(\hat{\mathscr{C}}_T-\mathbb{E}\hat{\mathscr{C}}_T)\Big\|_{S_2,2}\le\frac{1}{\|\|\Phi_T\|\|_F}\|\Phi_{T,t,t}\|_\infty\Big\|\hat{\mathscr{C}}_T-\mathbb{E}\hat{\mathscr{C}}_T\Big\|_{S_2,2}=o(\varrho_T)O\big(\frac{\sqrt{T}}{\|\|\Phi_T\|\|_F}\big)=o(1),$$

For the first term, we find using Lemma C.2 and Lemma C.3, respectively

$$\begin{aligned}\frac{1}{\|\|\Phi_T\|\|_F}\Big\|\mathcal{V}_T^\lambda-\mathbb{E}\mathcal{V}_T^\lambda\Big\|_{S_2,2}&\le\frac{1}{\|\|\Phi_T\|\|_F}\Big\|\mathcal{V}_T^\lambda-\mathbb{E}\mathcal{V}_T^\lambda-(\mathcal{V}_T^{(m),\lambda}-\mathbb{E}\mathcal{V}_T^{(m),\lambda})\Big\|_{S_2,2}\\&+\frac{1}{\|\|\Phi_T\|\|_F}\Big\|\mathcal{V}_T^{(m),\lambda}-\mathbb{E}\mathcal{V}_T^{(m),\lambda}-\mathscr{M}_{T,m}^{(\lambda)}\Big\|_{S_2,2}\\&=K_4\Upsilon_{4,m}\sum_{t=0}^{\infty}\nu_{\mathbb{H},4}(X_t)\\&+\nu_{\mathbb{H},4}(X_0)^2\frac{m^2\sqrt{T}}{\|\|\Phi_T\|\|_F}(\max_t\|\|A_{T,t}\|\|_\infty^2+m\sum_{s=1}^{t-4m}\|\|A_{T,s-t}-A_{T,s-(t+1)}\|\|_\infty^2)^{1/2}).\end{aligned}$$

Therefore, under Assumption 3.2 the results follows. □

Lemma C.1. *Let X_t satisfy Assumption 3.1 with $p=4$. Then*

$$\Big\|\sum_{1\le t\le T}(X_t\otimes X_t)-T\mathbb{E}(X_0\otimes X_0)\Big\|_{S_2,2}=O(\sqrt{T}\sum_{j=0}^{\infty}\nu_{\mathbb{H},4}(X_j))=O(\sqrt{T}).$$

Proof of (C.1). By stationarity and ergodicity and orthogonality of the projection operators

$$\Big\| \sum_{1\le t\le T} (X_t \otimes X_t) - T\mathbb{E}(X_0\otimes X_0)\Big\|^2_{S_2,2} = \Big\|\sum_{j=-\infty}^{T}\sum_{t=1}^{T} P_j(X_t\otimes X_t)\Big\|^2_{S_2,2} \le \sum_{j=-\infty}^{T}\Big\|\sum_{t=1}^{T} P_j(X_t\otimes X_t)\Big\|^2_{S_2,2}$$

Then, by Minkowski's inequality Cauchy Schwarz's inequality and stationarity

$$\begin{aligned}
\Big\|\sum_{t=1}^{T} P_j(X_t\otimes X_t)\Big\|_{S_2,2} &= \sum_{j=-\infty}^{T}\sum_{t=1}^{T}\Big\|P_0(X_{t-j}\otimes X_{t-j})\Big\|_{S_2,2}\\
&\le \sum_{t=1}^{T}\Big\|X_{t-j}\otimes(X_{t-j}-X_{t-j,\{0\}})\Big\|_{S_2,2} + \Big\|(X_{t-j}-X_{t-j,\{0\}})\otimes X_{t-j,\{0\}}\Big\|^2_{S_2,2}\\
&\le \sum_{t=1}^{T}(\mathbb{E}\|X_{t-j}\|^4_H)^{1/4}(\mathbb{E}\|(X_{t-j}-X_{t-j,\{0\}})\|^4_H)^{1/4} + (\mathbb{E}\|X_{t-j,\{0\}}\|^4_H)^{1/4}(\mathbb{E}\|(X_{t-j}-X_{t-j,\{0\}})\|^4_H)^{1/4}\\
&\le 2\|X_0\|_{\mathbb{H},4}\sum_{t=1}^{T}\nu_{\mathbb{H},4}(X_{t-j}).
\end{aligned}$$

Consequently,

$$\sum_{j=-\infty}^{T}\Big\|\sum_{t=1}^{T} P_j(X_t\otimes X_t)\Big\|^2_{S_2,2} \le 4\|X_0\|^2_{\mathbb{H},4}\sum_{j=-\infty}^{T}\Big(\sum_{t=1}^{T}\nu_{\mathbb{H},4}(X_{t-j})\Big)^2 \le 4\|X_0\|^2_{\mathbb{H},4}T(\sum_{j=0}^{\infty}\nu_{\mathbb{H},4}(X_j))^2.$$

The result follows by taking the square root. □

Lemma C.2 (M-dependence approximation)**.** *Suppose* (4) *with* $2p$ *is satisfied for some* $p\ge 2$. *Then*

$$\frac{\|\mathcal{V}^{\lambda}_T - \mathbb{E}\mathcal{V}^{\lambda}_T - (\mathcal{V}^{(m),\lambda}_T - \mathbb{E}\mathcal{V}^{(m),\lambda}_T)\|_{S_2,p}}{\sqrt{T}|\!|\!|\phi_T|\!|\!|_{\ell_2}\sum_{t=0}^{\infty}\nu_{\mathbb{H},2p}(X_t)} \le K_p\Upsilon_{2p,m}$$

where $\Upsilon_{2p,m} = 2\sum_{t=0}^{\infty}\min(\nu_{\mathbb{H},2p}(X_t),\mathbf{\Delta}_{2p,2,m+1})$ *and*

$$\mathcal{V}^{\lambda}_T := \sum_{s=2}^{T}\sum_{t=1}^{s-1}\Phi^{(\lambda)}_{T,s,t}(X_s\otimes X_t) \quad \textit{and} \quad \mathcal{V}^{(m),\lambda}_T := \sum_{s=2}^{T}\sum_{t=1}^{s-1}\Phi^{(\lambda)}_{T,s,t}(X^{(m)}_s\otimes X^{(m)}_t)$$

Proof of Lemma C.2. Let $N_{T,s} = \sum_{s=1}^{t-1}\phi^{(\lambda)}_{T,s-t}(X^{(m)}_s)$ and $N_{T,s} = \sum_{s=1}^{t-1}\phi^{(\lambda)}_{T,s-t}(X_s)$ and observe this these are $\mathcal{G}_s$-measurable. By orthogonality of the projections and minkowski's inequality

$$\|\mathcal{V}^{\lambda}_T - \mathbb{E}\mathcal{V}^{\lambda}_T - (\mathcal{V}^{(m),\lambda}_T - \mathbb{E}\mathcal{V}^{(m),\lambda}_T)\|^2_{S_2,p} \le 2\sum_{j=-\infty}^{T}\|P_j(\mathcal{V}^{\lambda}_T - \tilde{\mathcal{V}}^{(m),\lambda}_T)\|^2_{S_2,p} + \|P_j(\tilde{\mathcal{V}}^{(m),\lambda}_T - \mathcal{V}^{(m),\lambda}_T)\|^2_{S_2,p}$$

where $\tilde{\mathcal{V}}^{(m),\lambda}_T = \sum_{s=2}^{T}X_s\otimes N^{(m)}_{T,s}$. We shall focus on bounding the first term as the second is similar and has the same upperbound. A similar trick as in Lemma A.2 shows

$$\mathbb{E}[\mathcal{V}^{\lambda}_T - \tilde{\mathcal{V}}^{(m),\lambda}_T|\mathcal{G}_{j-1}] = \mathbb{E}[\mathcal{V}^{\lambda}_T - \tilde{\mathcal{V}}^{(m),\lambda}_T|\mathcal{G}_{j,\{j\}}] = \mathbb{E}[\mathcal{V}_{T,\{j\}} - \tilde{\mathcal{V}}^{(m)}_{T,\{j\}}|\mathcal{G}_j],$$

so that by the contraction property of the conditional expectation

$$\Big\|P_j(\mathcal{V}^{\lambda}_T - \tilde{\mathcal{V}}^{(m),\lambda}_T)\Big\|_{S_2,p} \le \Big\|\mathcal{V}^{\lambda}_T - \tilde{\mathcal{V}}^{(m),\lambda}_T - (\mathcal{V}_{T,\{j\}} - \tilde{\mathcal{V}}^{(m)}_{T,\{j\}})\Big\|_{S_2,p}$$

$$
\begin{aligned}
&\leq \Big\| \sum_{s=2}^{T} (X_s - X_{s,\{j\}}) \otimes (N_{T,s} - N_{T,s}^{(m)}) \Big\|_{S_2,p} \\
&+ \Big\| \sum_{s=2}^{T} X_{s,\{j\}} \otimes (N_{T,s} - N_{T,s}^{(m)} - N_{T,s,\{j\}} + N_{s-1,T,\{j\}}^{(m)}) \Big\|_{S_2,p} := J_1 + J_2,
\end{aligned}
$$

where we added and subtracted $X_{s,\{j\}} \otimes (N_{T,s} - N_{T,s}^{(m)})$ and applied Minkowski's inequality. From Lemma A.2(iii)

$$
\| N_{T,s} - N_{T,s}^{(m)} \|_{\mathbb{H},2p} \leq (K_p^q \|\!| \phi_T \|\!|_{\ell_q}^q \mathbf{\Delta}_{2p,q,m+1})^{1/q}.
$$

and $\| X_s - X_{s,\{j\}} \|_{\mathbb{H},2p} \leq \nu_{\mathbb{H},2p}(X_{s-j})$. Then, by Cauchy-schwarz inequality and that $q = \min(2,2p) = 2$

$$
\begin{aligned}
\sum_{j=-\infty}^{T} J_1^2 &\leq \sum_{j=-\infty}^{T} \Big(\sum_{s=2}^{T} \nu_{\mathbb{H},2p}(X_{s-j}) \Big)^2 \Big((K_p^2 \|\!| \phi_T \|\!|_{\ell_2}^2 \mathbf{\Delta}_{2p,2,m+1})^{1/2} \Big)^2 \\
&\leq K_p^2 \|\!| \phi_T \|\!|_{\ell_2}^2 \mathbf{\Delta}_{2p,1,m+1}^2 \sum_{s=2}^{T} \sum_{j=-\infty}^{T} \nu_{\mathbb{H},2p}(X_{s-j}) \sum_{s=2}^{T} \nu_{\mathbb{H},2p}(X_{s-j}) \\
&\leq T K_p^2 \|\!| \phi_T \|\!|_{\ell_q}^2 \mathbf{\Delta}_{2p,1,m+1}^2 \mathbf{\Delta}_{2p,1,0}^2
\end{aligned}
$$

where we used that $\mathbf{\Delta}_{2p,2,m+1} \leq \mathbf{\Delta}_{2p,1,m+1}^2$ and $\sum_{s=2}^{T} \nu_{\mathbb{H},2p}(X_{s-j}) \leq \mathbf{\Delta}_{2p,1,0}$. Secondly, noting that from Lemma A.2, (26) and Minkowski's inequality

$$
\begin{aligned}
&\| X_s - X_s^{(m)} + X_{s,\{j\}}^{(m)} - X_{s,\{j\}} \|_{\mathbb{H},2p} \\
&\leq \min(\| X_s - X_s^{(m)} \|_{\mathbb{H},2p} + \| X_{s,\{j\}}^{(m)} - X_{s,\{j\}} \|_{\mathbb{H},2p}, \| X_s - X_{s,\{j\}} \|_{\mathbb{H},2p} + \| X_{s,\{j\}}^{(m)} - X_s^{(m)} \|_{\mathbb{H},2p}) \\
&\leq 2 \min\Big(\sum_{j=m+1}^{\infty} \| D_{t,j} \|_{\mathbb{H},2p}, \nu_{\mathbb{H},2p}(X_{s-j}) \Big).
\end{aligned}
$$

Hence, changing the order of summation and from property (8) of $\Phi_{T,t,s}^{(\lambda)}$

$$
\begin{aligned}
\sum_{j=-\infty}^{T} J_2^2 &\leq \sum_{j=-\infty}^{T} \Big\| \sum_{s=2}^{T} \sum_{t=1}^{s-1} X_{s,\{j\}} \otimes \phi_{T,t-s}^{(\lambda)} (X_t - X_t^{(m)} + X_{t,\{j\}}^{(m)} - X_{t,\{j\}}) \Big\|_{S_2,p}^2 \\
&= \sum_{j=-\infty}^{T} \Big(\sum_{t=1}^{T-1} \Big\| \sum_{s=t+1}^{T} \phi_{T,s-t}^{(-\lambda)} (X_{s,\{j\}}) \otimes (X_t - X_t^{(m)} + X_{t,\{j\}}^{(m)} - X_{t,\{j\}}) \Big\|_{S_2,p} \Big)^2 \\
&\leq (K_p^2 \|\!| \phi_T \|\!|_{\ell_2}^2 \mathbf{\Delta}_{2p,2,0}) 2 \sum_{j=-\infty}^{T} \mathbf{\Delta}_{2p,1,0} \sum_{t=1}^{T-1} \min\big(\mathbf{\Delta}_{2p,m+1}, \nu_{\mathbb{H},2p}(X_{t-j}) \big) \\
&\leq (K_p^2 \|\!| \phi_T \|\!|_{\ell_2}^2 \mathbf{\Delta}_{2p,2,0}) 2 T \mathbf{\Delta}_{2p,1,0} \Upsilon_{2p,m}
\end{aligned}
$$

Noting again that $\mathbf{\Delta}_{2p,2,m+1} \leq \mathbf{\Delta}_{2p,1,m+1}^2$ and that $\Upsilon_{2p,m} \geq \mathbf{\Delta}_{2p,1,m+1}$, we obtain

$$
\begin{aligned}
&\sum_{j=-\infty}^{T} \| P_j (\mathcal{V}_T^{\lambda} - \tilde{\mathcal{V}}_T^{(m),\lambda}) \|_{S_2,p}^2 + \| P_j (\tilde{\mathcal{V}}_T^{(m),\lambda} - \mathcal{V}_T^{(m),\lambda}) \|_{S_2,p}^2 \\
&\quad \leq 2 \Big(T K_p^2 \| \phi_T \|_{\ell_2}^2 \mathbf{\Delta}_{2p,1,0}^2 \mathbf{\Delta}_{2p,2,m+1} + 2 T K_p^2 \|\!| \phi_T \|\!|_{\ell_2}^2 \mathbf{\Delta}_{2p,2,0} \mathbf{\Delta}_{2p,1,0} \Upsilon_{2p,m} \Big) \\
&\quad \leq 2 K_p^2 \|\!| \phi_T \|\!|_{\ell_2}^2 T \mathbf{\Delta}_{2p,1,0}^2 \Upsilon_{2p,m}^2.
\end{aligned}
$$

□

Lemma C.3 (martingale approximation to m dependent process)**.** *Let* $\mathscr{M}^{(\lambda)}_{T,m}$ *as defined in* (10) *and* $\mathcal{V}^{(m),\lambda}_{T}$ *as in Lemma C.2. Under Assumption 3.1 with* $p=4$*, then*

$$\frac{\|\mathcal{V}^{(m),\lambda}_{T}-\mathbb{E}\mathcal{V}^{(m),\lambda}_{T}-\mathscr{M}^{(\lambda)}_{T,m}\|_{S_2,2}}{m^2\sqrt{T}\|X_0\|^2_{\mathbb{H},4}} \le (\max_t |\!|\!| A_{T,t}|\!|\!|^2_\infty + m\sum_{s=1}^{t-4m}|\!|\!| A_{T,s-t}-A_{T,s-(t+1)}|\!|\!|^2_\infty)^{1/2}.$$

Proof of Lemma C.3. By construction $D^{(\lambda)}_{m,k}\in\mathscr{L}^p_H$ defines an m-dependent martingale difference and therefore we can write $D^{(\lambda)}_{m,k}=\sum_{t=0}^m P_k(X^{(m)}_{t+k})e^{-\mathrm{i}t\lambda}$ since the terms $t>m$ are zero. We decompose the difference $\mathcal{V}^{(m),\lambda}_{T}-\mathscr{M}^{(\lambda)}_{T,m}$ as follows

$$\begin{aligned}
&\sum_{t=2}^{T}X^{(m)}_t\otimes\sum_{s=1}^{t-1}\phi^{(\lambda)}_{T,s-t}(X^{(m)}_s-D^{(\lambda)}_{m,s})+\sum_{t=2}^{T}(X^{(m)}_t-D^{(\lambda)}_{m,t})\otimes\sum_{s=1}^{t-1}\phi^{(\lambda)}_{T,s-t}D^{(\lambda)}_{m,s}\\
&\quad=\sum_{t=2}^{T}X^{(m)}_t\otimes\Big(\sum_{s=1}^{t-4m}\phi^{(\lambda)}_{T,s-t}(X^{(m)}_s-D^{(\lambda)}_{m,s})+\sum_{s=t-4m+1}^{t-1}\phi^{(\lambda)}_{T,s-t}(X^{(m)}_s-D^{(\lambda)}_{m,s})\Big)\\
&\quad+\sum_{t=2}^{T}(X^{(m)}_t-D^{(\lambda)}_{m,t})\otimes\sum_{s=1}^{t-1}\phi^{(\lambda)}_{T,s-t}D^{(\lambda)}_{m,s}:=\sum_t M^\star_t+Y_t+Z_t
\end{aligned}\tag{41}$$

Note that $\mathcal{V}^{(m),\lambda}_{T}-\mathbb{E}\mathcal{V}^{(m),\lambda}_{T}-\mathscr{M}^{(\lambda)}_{T,m}=\sum_t M^\star_t+Y_t-\mathbb{E}Y_t+Z_t-\mathbb{E}Z_t$. We treat the above terms separately. Firstly, we consider $M^\star_t:=X^{(m)}_t\otimes\sum_{s=1}^{t-4m}\phi^{(\lambda)}_{T,s-t}(X^{(m)}_s-D^{(\lambda)}_{m,s})$, for which the process $\{M^\star_{t+4mk}\}_{k\in\mathbb{N}}$ is itself a martingale in $\mathscr{L}^2_{S_2}$. Let $W_k=\sum_{t=0}^m\mathbb{E}[X^{(m)}_{t+k}|\mathscr{G}_k]e^{-\mathrm{i}t\lambda}$ and observe that

$$\begin{aligned}
X^{(m)}_k&=\sum_{t=0}^{m}\mathbb{E}[X^{(m)}_{t+k}|\mathscr{G}_k]e^{-\mathrm{i}t\lambda}-\sum_{t=1}^{m}\mathbb{E}[X^{(m)}_{t+k}|\mathscr{G}_k]e^{-\mathrm{i}t\lambda}\\
&=\sum_{t=0}^{m}\mathbb{E}[X^{(m)}_{t+k}|\mathscr{G}_k]e^{-\mathrm{i}t\lambda}-\sum_{t=0}^{m-1}\mathbb{E}[X^{(m)}_{t+k+1}|\mathscr{G}_k]e^{-\mathrm{i}(t+1)\lambda}=W_k-\mathbb{E}[W_{k+1}|\mathscr{G}_k]e^{-\mathrm{i}\lambda}
\end{aligned}$$

and that $D^{(\lambda)}_{m,k}=W_k-\mathbb{E}[W_k|\mathscr{G}_{k-1}]$. Therefore

$$\begin{aligned}
\Big\|\sum_{s=1}^{t-4m}\phi^{(\lambda)}_{T,s-t}(X^{(m)}_s-D^{(\lambda)}_{m,s})\Big\|_{\mathbb{H},2}&=\Big\|\sum_{s=1}^{t-4m}A_{T,s-t}e^{-\mathrm{i}(s-t)\lambda}(\mathbb{E}[W_s|\mathscr{G}_{s-1}]-\mathbb{E}[W_{s+1}|\mathscr{G}_s]e^{-\mathrm{i}\lambda})\Big\|_{\mathbb{H},2}\\
&\Big\|\sum_{s=1}^{t-4m}A_{T,s-t}\Big(e^{-\mathrm{i}(s-t)\lambda}\mathbb{E}[W_s|\mathscr{G}_{s-1}]-e^{\mathrm{i}(t-(s+1))\lambda}\mathbb{E}[W_{s+1}|\mathscr{G}_s]\Big)\Big\|_{\mathbb{H},2}.
\end{aligned}$$

Set $V_s=e^{\mathrm{i}(t-s)\lambda}\mathbb{E}[W_s|\mathscr{G}_{s-1}]$. Summation by parts, Holder's inequality for operators and Lemma A.1 together yield

$$\begin{aligned}
\Big\|\sum_{s=1}^{t-4m}A_{T,s-t}\Big(V_s-V_{s+1}\Big)\Big\|_{\mathbb{H},2}&\le\Big\|A_{T,t-4m}(V_{t-4m})\Big\|_{\mathbb{H},2}+\Big\|\sum_{s=1}^{t-4m+1}(A_{T,s-t}-A_{T,s-t-1})V_s\Big\|_{\mathbb{H},2}\\
&\le\max_t|\!|\!| A_{T,t}|\!|\!|_\infty\|V_{t-4m}\|_{\mathbb{H},2}+\Big\|\sum_{s=1}^{t-4m+1}(A_{T,s-t}-A_{T,s-t-1})(\sum_{l=1}^{m}P_{s-l}V_s)\Big\|_{\mathbb{H},2}\\
&\le\max_t|\!|\!| A_{T,t}|\!|\!|_\infty\|V_{t-4m}\|_{\mathbb{H},2}+\sqrt{\Big\|\sum_{l=1}^{m}\sum_{s=1}^{t-4m+1}(A_{T,s-t}-A_{T,s-t-1})(P_{s-l}V_s)\Big\|^2_{\mathbb{H},2}}\\
&\le 2m\|X_0\|_{\mathbb{H},2}\max_t|\!|\!| A_{T,t}|\!|\!|_\infty+\Big(\sum_{s=1}^{t-4m}|\!|\!| A_{T,s-t}-A_{T,s-t-1}|\!|\!|^2_\infty\Big)^{1/2}Cm^{3/2}\|X_0\|_{\mathbb{H},2}
\end{aligned}$$

Hence, some constant C,

$$\|M_t^\star\|_{\mathbb{H},2} \le Cm\|X_0\|^2_{\mathbb{H},2}\big(\max_t |||A_{T,t}|||_\infty + (\sum_{s=1}^{t-4m} |||A_{T,s-t} - A_{T,s-t-1}|||^2_\infty m)^{1/2}\big)$$

where we used that, the contraction property and stationarity to find $\|V_{t-4m}\|_{\mathbb{H},2} \le 2\|W_k\|_{\mathbb{H},2} \le 2m\|X_0\|_{\mathbb{H},2}$ and that, since $P_j(\cdot)$ form martingale differences, we have $((\sum_{l=1}^m \|P_{-l}V_0\|_{\mathbb{H},2})^2)^{1/2} \le (\sum_{l=1}^m \|P_{-l}V_0\|^2_{\mathbb{H},2})^{1/2} \le \sqrt{m}2m\|X_0\|_{\mathbb{H},2}$. Consequently, using a martingale decomposition of the sum

$$\Big\|\sum_{t=1}^T M_t^\star\Big\|_{\mathbb{H},2} \le \sum_{t=1}^{4m}\Big\|\sum_{s=0}^{\lfloor T-t/(4m)\rfloor} M^\star_{s+4mt}\Big\|_{\mathbb{H},2}$$
$$\le 4m^2T^{1/2}C\|X_0\|^2_{\mathbb{H},2}\big(\max_t |||A_{T,t}|||_\infty + (\sum_{s=1}^{t-4m} |||A_{T,s-t} - A_{T,s-t-1}|||^2_\infty m)^{1/2}\big)$$

For Y_t, we note that $\sum_{s=t-4m+1}^{t-1}\phi^{(\lambda)}_{T,s-t}(X_s^{(m)} - D^{(\lambda)}_{m,s})$ is $\mathscr{G}_{t-1}$ measurable and that Y_t is $t-5m$-dependent. Therefore, via the Minkow'ski's inequality, the Cauchy Schwarz inequality and a similar decomposition as above shows

$$\Big\|\sum_t Y_t - \mathbb{E}Y_t\Big\|_{\mathbb{H},2} = C\sqrt{T}m\|X_0\|^2_{\mathbb{H},4}(\max_t |||A_{T,t}|||_\infty + \sum_{s=1}^{t-4m} |||A_{T,s-t} - A_{T,s-t-1}|||^2_\infty m^{1/2})$$

and similarly for $\|\sum_t Z_t - \mathbb{E}Z_t\|_{\mathbb{H},2}$. □

D Proofs of Section 4

Proof of Theorem 4.1. We first prove part (i) and consider the following bias variance decomposition

$$|||\hat{\mathscr{F}}^{(\lambda)}_T - \mathscr{F}^{(\lambda)}|||_{S_2,p} \le |||\hat{\mathscr{F}}^{(\lambda)}_T - \mathbb{E}\hat{\mathscr{F}}^{(\lambda)}_T|||_{S_2,2} + |||\mathbb{E}\hat{\mathscr{F}}^{(\lambda)}_T - \mathscr{F}^{(\lambda)}|||_{S_2,2}$$

We shall first focus on the first term. We decompose the error as

$$\hat{\mathscr{F}}^{(\lambda)}_T - \mathbb{E}\hat{\mathscr{F}}^{(\lambda)}_T = \frac{1}{T}\Big(\sum_{s=2}^T X_s \otimes N^{(\lambda)}_{T,s} - \mathbb{E}\sum_{s=2}^T X_s \otimes N^{(\lambda)}_{T,s} + (\sum_{s=2}^T X_s \otimes N^{(\lambda)}_{T,s})^\dagger - \mathbb{E}(\sum_{s=2}^T X_s \otimes N^{(\lambda)}_{T,s})^\dagger \tag{42}$$
$$+ \sum_{1\le t\le T} A_{T,0}(X_t \otimes X_t) - \mathbb{E}\sum_{1\le t\le T} A_{T,0}(X_t \otimes X_t)\Big) \tag{43}$$

where $N^{(\lambda)}_{T,s} = \sum_{t=1}^{s-1}\phi^{(\lambda)}_{T,t,s}X_t$. We shall first derive the order of (42) in $\mathscr{L}^2_{S_2(H)}$. By ergodicity and othogonality of the projection operator

$$\Big\|\sum_{t=2}^T X_t \otimes N^{(\lambda)}_{T,t-1} - \mathbb{E}\sum_{t=2}^T X_t \otimes N^{(\lambda)}_{T,t-1}\Big\|^2_{S_2,2} = \sum_{j=-\infty}^T \Big\|P_j\Big(\sum_{t=2}^T\sum_{s=1}^{t-1}(X_t \otimes \phi^{(\lambda)}_{T,s,t}X_s)\Big)\Big\|^2_{S_2,2}$$

By the contraction property of the conditional expectation and Cauchy Schwarz inequality

$$\Big\|P_j\Big(\sum_{t=2}^T\sum_{s=1}^{t-1}\phi^{(\lambda)}_{T,t,s}(X_t \otimes X_s)\Big)\Big\|_{S_2,2} \le \Big\|\sum_{t=2}^T\sum_{s=1}^{t-1}\phi^{(\lambda)}_{T,t,s}\Big((X_t \otimes X_s) - (X_{t,\{j\}} \otimes X_{s,\{j\}})\Big)\Big\|_{S_2,2}$$
$$= \Big\|\sum_{t=2}^T\sum_{s=1}^{t-1}\phi^{(\lambda)}_{T,t,s}\Big(X_t \otimes (X_s - X_{s,\{j\}}) + (X_t - X_{t,\{j\}}) \otimes X_{s,\{j\}}\Big)\Big\|_{S_2,2}$$

$$\leq \Big\| \sum_{t=2}^{T}\sum_{s=1}^{t-1} \phi_{T,t,s}^{(\lambda)}\Big(X_t \otimes (X_s - X_{s,\{j\}})\Big)\Big\|_{S_2,2} + \Big\| \sum_{t=2}^{T}\sum_{s=1}^{t-1} \phi_{T,t,s}^{(\lambda)}\Big((X_t - X_{t,\{j\}}) \otimes X_{s,\{j\}}\Big)\Big\|_{S_2,2}$$

$$\leq \sum_{s=1}^{T-1} \Big\| \sum_{t=s+1}^{T} \phi_{T,t,s}^{(\lambda)} X_t \otimes (X_s - X_{s,\{j\}})\Big\|_{S_2,2} + \sum_{t=2}^{T} \Big\| \sum_{s=1}^{t-1} \phi_{T,t,s}^{(\lambda)}\Big((X_t - X_{t,\{j\}}) \otimes X_{s,\{j\}}\Big)\Big\|_{S_2,2}$$

$$\leq \sum_{s=1}^{T-1} \Big(\Big\| \sum_{t=s+1}^{T} \phi_{T,t,s}^{(\lambda)} X_t \Big\|_{\mathbb{H},4}^2 \Big\| X_s - X_{s,\{j\}} \Big\|_{\mathbb{H},4}^2 \Big)^{1/2} + \sum_{t=2}^{T} \Big\| \Big((X_t - X_{t,\{j\}}) \otimes \sum_{s=1}^{t-1} \phi_{T,s,t}^{(\lambda)} X_{s,\{j\}} \Big) \Big\|_{S_2,2}$$

$$\leq \sum_{s=1}^{T-1} \Big(\Big\| \sum_{t=s+1}^{T} \phi_{T,t,s}^{(\lambda)} X_t \Big\|_{\mathbb{H},4}^2 \Big\| X_s - X_{s,\{j\}} \Big\|_{\mathbb{H},4}^2 \Big)^{1/2} + \sum_{t=2}^{T} \Big(\| X_t - X_{t,\{j\}} \|_{\mathbb{H},4}^2 \Big\| \sum_{s=1}^{t-1} \phi_{T,s,t}^{(\lambda)} X_{s,\{j\}} \Big\|_{\mathbb{H},4}^2 \Big)^{1/2}$$

Hence, using Lemma A.2 we obtain

$$\frac{1}{T^2} \sum_{j=-\infty}^{T} \Big\| P_j \Big(\sum_{t=2}^{T}\sum_{s=1}^{t-1} \phi_{T,t,s}^{(\lambda)} (X_t \otimes X_s) \Big) \Big\|_{S_2,2}^2$$

$$\leq \frac{1}{T^2} \sum_{j=-\infty}^{T} \Big(K_4 \max_{1\leq s\leq T-1} |||\phi_{s,T}|||_{\ell_2} \mathbf{\Delta}_{4,2,0}^{1/2} \sum_{s=1}^{T-1} \nu_{\mathbb{H},4}(X_{s-j}) + \sum_{t=2}^{T} \nu_{\mathbb{H},4}(X_{t-j}) K_4 \max_{2\leq t\leq T} |||\phi_{t,T}|||_{\ell_2} \mathbf{\Delta}_{4,2,0}^{1/2} \Big)^2$$

$$\leq \frac{1}{T^2} K_4^2 (\max_{1\leq t\leq T} |||\phi_{t,T}|||_{\ell_2})^2 \mathbf{\Delta}_{4,2,0} \sum_{j=-\infty}^{T} \Big(\sum_{s=2}^{T-1} \nu_{\mathbb{H},4}(X_{s-j}) + \sum_{t=2}^{T} \nu_{\mathbb{H},4}(X_{t-j}) \Big)^2$$

$$\leq \frac{1}{T^2} 4K_4^2 (\max_{1\leq t\leq T} |||\phi_{t,T}|||_{\ell_2})^2 T \mathbf{\Delta}_{4,0}^4) = O((b_T T)^{-1}).$$

From Lemma C.1 it is immediate that (42) is of order $O(T^{-1})$ in $\mathscr{L}_{S_2}^2$. It therefore follows by Minkowski's inequality that

$$\| \hat{\mathscr{F}}_T^{(\lambda)} - \mathbb{E}\hat{\mathscr{F}}_T^{(\lambda)} \|_{S_2,2}^2 = O((b_T T)^{-1}).$$

Let us then consider the bias. Observe that from (22), stationarity yields

$$\mathbb{E}\hat{\mathscr{F}}_T^{\lambda} = \frac{1}{2\pi T} \sum_{s,t=1}^{T} \mathbb{E}(X_s \otimes X_t) w(b_T(t-s)) e^{\mathrm{i}\lambda(t-s)} = \frac{1}{2\pi} \sum_{|h|\leq T} w(b_T h) \frac{1}{T} \sum_{t=1}^{T-|h|} \mathrm{Cov}(X_{t+h} \otimes X_t) e^{-\mathrm{i}\lambda h}$$

$$= \frac{1}{2\pi} \sum_{|h|<T} w(bh)(1 - \frac{|h|}{T}) C_h e^{-\mathrm{i}\lambda h}.$$

Hence, using Minkowski's inequality we can bound the error by

$$|||\mathbb{E}\hat{\mathscr{F}}_T^{(\lambda)} - \mathscr{F}^{\lambda}|||_2 \leq \underbrace{\Big|\Big|\Big| \frac{1}{2\pi} \sum_{|h|<T} (w(b_T h) - 1) C_h e^{-\mathrm{i}\lambda h} \Big|\Big|\Big|_2}_{R_{o,\lambda}} + \underbrace{\Big|\Big|\Big| \frac{1}{2\pi} \sum_{|h|\geq T} C_h e^{-\mathrm{i}\lambda h} \Big|\Big|\Big|_2}_{R_{1,\lambda}}$$

$$+ \underbrace{\Big|\Big|\Big| \frac{1}{2\pi} \sum_{|h|<T} w(b_T h) \frac{|h|}{T} C_h e^{-\mathrm{i}\lambda h} \Big|\Big|\Big|_2}_{R_{2,\lambda}} \tag{44}$$

It is immediate that, since $\sum_h |||C_h||| < \infty$,

$$\sup_{\lambda} |||R_{1,\lambda}|||_2 \leq \sum_{|h|\geq T} |||C_h|||_2 \to 0 \text{ as } T \to \infty.$$

For the final term, note that $\sum_{h\in\mathbb{Z}} w(bh) |||C_h|||_2 \leq \sup_x |w(x)| \sum_{h\in\mathbb{Z}} |||C_h|||_2 < \infty$ since $w(\cdot)$ is bounded. Hence by Kronecker's lemma $\sup_{\lambda} |||R_{2,\lambda}|||_2 \to 0$ as $T \to \infty$. Finally, provided that $b_T \to 0$ as $T \to \infty$

and since $\lim_{x\to 0} w(x) = w(0)$, we obtain $w(b_T h) \to w(0) = 1$ as $T \to \infty$. Hence $\sup_\lambda |||R_{0,\lambda}|||_2 \to 0$, from which asymptotic unbiasedness follows. Next we prove part (ii), for which it remain to derive the order of the bias under the additional conditions. Firstly, from Proposition 3.1, the condition $\sum_{h\in\mathbb{Z}} h\|P_0(X_h)\|_{\mathbb{H},2} < \infty$ implies that $\sum_{h\in\mathbb{Z}} h|||C_h|||_{\mathbb{H},2} < \infty$. Now, observe that we can decompose the bias as follows.

$$|||\mathbb{E}\hat{\mathscr{F}}_T^{(\lambda)} - \mathscr{F}^\lambda|||_2 \le \underbrace{\Big|\Big|\Big|\frac{1}{2\pi}\sum_{|h|<1/b_T}(w(b_T h)-1)C_h e^{-\mathrm{i}\lambda h}\Big|\Big|\Big|_2}_{R_{o,\lambda}} + \underbrace{\Big|\Big|\Big|\frac{1}{2\pi}\sum_{|h|\ge 1/b_T} C_h e^{-\mathrm{i}\lambda h}\Big|\Big|\Big|_2}_{R_{1,\lambda}}$$
$$+ \underbrace{\Big|\Big|\Big|\frac{1}{2\pi}\sum_{|h|<1/b_T} w(b_T h)\frac{|h|}{T} C_h e^{-\mathrm{i}\lambda h}\Big|\Big|\Big|_2}_{R_{2,\lambda}}$$

For the final term, the fact that $\sup_x |w(x)| = O(1)$ and $\sum_h |h||||C_h||| < \infty$, yield

$$\sup_\lambda |||R_{2,\lambda}|||_2 = \Big|\Big|\Big|\frac{1}{2\pi}\sum_{|h|<1/b_T}|w(b_T h)|\frac{|h|}{T}C_h\Big|\Big|\Big|_2 \le \sup_x |w(x)|\frac{1}{2\pi}\sum_{|h|<1/b_T}\frac{|h|}{T}|||C_{|h|}|||_2 = O(\frac{1}{T}),$$

while moreover $\sup_\lambda |||R_{1,\lambda}|||_2 = O(b_T)$. Finally, since $|w(x)-1| = O(x)$ as $x \to 0$ and $\sum_h |h||||C_h||| < \infty$, we find for the first term

$$\sup_\lambda |||R_{0,\lambda}|||_2 = |||\frac{1}{2\pi}\sum_{|h|<1/b_T}(w(b_T h)-1)C_{|h|}e^{-\mathrm{i}\lambda h}|||_2 \le \sum_h O(b_T h)|||C_h|||_2 = O(b_T).$$

□

Proof of Theorem 4.2. From (22) we have $\hat{\mathscr{F}}^\lambda = (2\pi T)^{-1}\hat{\mathscr{Q}}^\lambda$ with $\phi^\omega_{T,(t-s)} = w(b_T(t-s))e^{\mathrm{i}\omega(t-s)}$. Note that in this case $|||\Phi_T|||_F^2 := \sum_{t=1}^T\sum_{s=1}^T w^2(b_T(t-s))$ and that $\varrho_T^2 = \sum_{s=1}^T w^2(b_T s)$. Moreover, under Assumption 4.1

$$\sum_{t=1}^T\sum_{s=1}^T |\phi_{T,(t-s)}|^2 = \sum_{t=1}^T\sum_{s=1}^T \Big|\frac{1}{2\pi}w(b_T(t-s))e^{-\mathrm{i}\omega(t-s)}\Big|^2 = \sum_{t=1}^T\sum_{s=1}^T w^2(b_T(t-s))$$
$$\to \frac{T}{b_T}\int w^2(x)dx = \frac{T}{b_T}\kappa^2. \tag{45}$$

Therefore it suffices to verify the conditions of Theorem 3.1, which is given here for completeness but follows a standard argument [see e.g., 23]. From Assumption 4.1, it is obvious that (ii) of Assumption 3.2 holds. Additionally, from (45), it is immediate that $\varrho_T^2 = \kappa O(b_T^{-1})$ and $|||\Phi_T|||_F^2 = O(T\varrho_T^2)$ so that (i) is also satisfied. For (iii), observe that we can write

$$\sum_{t=1}^{M/b_T}|w(b_T t) - w(b_T(t-1))|^2 + \sum_{t=M/b_T}^{T}|w(b_T t) - w(b_T(t-1))|^2$$

By assumption, the function w is bounded and continuous except on a set of measure zero. Oserve that there must exist $\epsilon > 0$ such that $|w(b_T t) - w(b_T(t-1))| < \epsilon$, for some constant $\varepsilon > 0$ uniformly for $|t| \le M/b_T$, except for a finite number of points, ε/b_T. Noting that $w(\cdot)$ is bounded, the first term will thus be of order $o(1/b_T)$. For the second term, the length of the interval I_M, converges to zero for fixed b_T, T as $M \to \infty$. Since the summand is at most of order b_T^{-1} under the stated conditions the second term is of order $O(I_M/b_T) = o(\varrho_T^2)$. To verify (iv), we decompose again

$$\sum_{j=1}^{T-1}\sum_{s=1}^{j-M/b_T}\Big(\sum_{t=j+1}^{T} w(b_T(s-t))w(b_T(j-t)\Big)^2 + \sum_{j=1}^{T-1}\sum_{s=1\vee j-M/b_T}^{j-1}\Big(\sum_{t=j+1}^{T} w(b_T(s-t))w(b_T(j-t))\Big)^2$$

For the first term Cauchy-Schwarz' inequality and the condition $\sup_{0\le b\le 1} b\sum_{h\ge M/b} w^2(bh) \to 0$ as $M\to\infty$ yield

$$\begin{aligned}
&\sum_{j=1}^{T-1}\sum_{s=1}^{j-M/b_T}\sum_{t=j+1}^{T} w^2(b_T(s-t))\sum_{t=j+1}^{T} w^2(b_T(j-t))\\
&=\sum_{j=1}^{T-1}\sum_{t=j+1}^{T} w^2(b_T(j-t))\sum_{s=1}^{j-M/b_T}\sum_{t=M/b_T}^{T} w^2(b_T(s-t))\\
&= O(Tb_T^{-1}I_M b_T^{-1}) = o(|||\Phi_T|||_F^4).
\end{aligned}$$

Secondly,

$$\sum_{j=1}^{T-1}\sum_{s=1\vee j-M/b_T}^{j-1}\sum_{t=j+1}^{T} w^2(b_T(s-t))\sum_{t=j+1}^{T} w^2(b_T(j-t)) = O(Tb_T^{-1}\varrho_T^4) = O(Tb_T^{-3}) = o(|||\Phi_T|||_F^4)$$

where we used that $1/b_T = o(T)$. □

Proof of Lemma 4.1. Note that we can write the spectral density operator as

$$\begin{aligned}
2\pi T\hat{\hat{\mathscr{F}}}^{(\lambda)} &= \sum_{s,t=1}^{T}\Phi_{t,s,T}^{\lambda}\big((X_s-\mu+\mu-\hat{\mu})\otimes(X_t-\mu+\mu-\hat{\mu})\big)\\
&= 2\pi T\hat{\mathscr{F}}^{(\lambda)} + \sum_{s,t=1}^{T}\Phi_{t,s,T}^{\lambda}(\mu-\hat{\mu})\otimes(X_t-\mu) + \sum_{s,t=1}^{T}\Phi_{t,s,T}^{\lambda}(X_s-\mu)\otimes(\mu-\hat{\mu}) \qquad (46)\\
&+\sum_{s,t=1}^{T}\Phi_{t,s,T}^{\lambda}((\mu-\hat{\mu})\otimes(\mu-\hat{\mu}). \qquad (47)
\end{aligned}$$

hence we shall show that the last two terms in (46) and the term (47) are of lower order. For the second term of (46), a change of variables, the properties of the tensor product and the Cauchy-Schwarz inequality yield

$$\begin{aligned}
\Big\|\sum_{|h|\le T}\sum_{t=1}^{T-h} w(b_T(h))e^{\mathrm{i}(h)\lambda}(\mu-\hat{\mu})\otimes(X_t-\mu)\Big\|_{S_2,2} &= \Big\|\sum_{|h|\le T} w(b_T(h))e^{\mathrm{i}(h)\lambda}(\mu-\hat{\mu})\otimes\sum_{t=1}^{T-h}(X_t-\mu)\Big\|_{S_2,2}\\
&\le \sum_{|h|\le T} w(b_T(h))\Big(\|\mu-\hat{\mu}\|_{\mathbb{H},4}^2\Big\|\sum_{t=1}^{T-h}(X_t-\mu)\Big\|_{\mathbb{H},4}^2\Big)^{1/2}\\
&\le \sum_{|h|\le T} w(b_T(h))\Big(O(T^{-1})O(T)\Big)^{1/2} = O(\frac{1}{b_T}),
\end{aligned}$$

where we used Lemma A.2(i) in order to obtain $\|\mu-\hat{\mu}\|_{\mathbb{H},4}^2 = \|\frac{1}{T}\sum_{t=1}^{T}(X_t-\mathbb{E}X_T)\|_{\mathbb{H},4}^2 \le \frac{1}{T^2}T\mathbf{\Delta}_{4,2,0} = O(\frac{1}{T})$ and to obtain $\Big\|\sum_{t=1}^{T-h}(X_t-\mu)\Big\|_{\mathbb{H},4}^2 = O(T)$. The third term in (46) is similar. For (47), Holder's inequality and Lemma A.2(i) yield

$$\Big\|\sum_{s,t=1}^{T}\Phi_{t,s,T}^{\lambda}((\mu-\hat{\mu})\otimes(\mu-\hat{\mu})\Big\|_{S_2,2} \le \sum_{s,t=1}^{T}|||\Phi_{t,s,T}^{\lambda}|||_\infty\Big(\mathbb{E}|||(\mu-\hat{\mu})\otimes(\mu-\hat{\mu})|||_2^2\Big)^{1/2}$$

$$\leq \sum_{s,t=1}^{T} |||\phi^{\lambda}_{t,s,T}|||_{\infty} \|\mu - \hat{\mu}\|^2_{\mathbb{H},4} = O(Tb_T^{-1}\frac{1}{T}) = O(\frac{1}{b_T})$$

where we used that $\sum_{s,t=1}^{T} |||\phi^{\lambda}_{t,s,T}|||_{\infty} = \sum_{s,t=1}^{T} w(b_T(t-s)) = O(Tb_T^{-1})$. Hence, $\|\hat{\tilde{\mathscr{F}}}^{(\lambda)} - \hat{\mathscr{F}}^{(\lambda)}\|_{S_2,2} = O(T^{-1}b_T^{-1})$ and the exact argument shows that $\|\hat{\tilde{\mathscr{F}}}^{(\lambda)} - \hat{\mathscr{F}}^{(\lambda)}\|_{S_2,\frac{p}{2}} = O(T^{-1}b_T^{-1})$ for $p \geq 2$ provided the process is in $\mathscr{L}^p_H$. □