Available online at www.sciencedirect.com

SciVerse ScienceDirect

Stochastic Processes and their Applications 123 (2013) 2779–2807

stochastic processes and their applications

www.elsevier.com/locate/spa

# Cramér–Karhunen–Loève representation and harmonic principal component analysis of functional time series[☆]

Victor M. Panaretos*, Shahin Tavakoli

*Ecole Polytechnique Fédérale de Lausanne, Switzerland*

Available online 28 March 2013

**Abstract**

We develop a doubly spectral representation of a stationary functional time series, and study the properties of its empirical version. The representation decomposes the time series into an integral of uncorrelated frequency components (Cramér representation), each of which is in turn expanded in a Karhunen–Loève series. The construction is based on the spectral density operator, the functional analogue of the spectral density matrix, whose eigenvalues and eigenfunctions at different frequencies provide the building blocks of the representation. By truncating the representation at a finite level, we obtain a harmonic principal component analysis of the time series, an optimal finite dimensional reduction of the time series that captures both the temporal dynamics of the process, as well as the within-curve dynamics. Empirical versions of the decompositions are introduced, and a rigorous analysis of their large-sample behaviour is provided, that does not require any prior structural assumptions such as linearity or Gaussianity of the functional time series, but rather hinges on Brillinger-type mixing conditions involving cumulants.


*MSC:* 60G10; 62M15; 62M10

*Keywords:* Spectral representation; Spectral density operator; Functional data analysis; Functional principal components; Discrete Fourier transform; Cumulants; Mixing

## 0. Introduction

Though spectral decompositions can play an important role in the statistical analysis of many classes of stochastic processes, it may not be an exaggeration to claim that in functional data

☆ Research supported by an ERC Starting Grant Award.
* Corresponding author.
*E-mail address:* victor.panaretos@epfl.ch (V.M. Panaretos).

0304-4149/$ - see front matter 
http://dx.doi.org/10.1016/j.spa.2013.03.015

analysis in particular, they are not simply important, but *essential*. Functional data analysis consists in drawing inferences pertaining to the law of a continuous time stochastic process $\{X(\tau); \tau \in [0, 1]\}$ with mean function and covariance operator

$$m(\tau) = \mathbb{E}[X(\tau)] \quad \text{and} \quad \mathscr{R}_0 := \mathbb{E}[(X - m) \otimes (X - m)],$$

respectively, on the basis of a collection of $T$ (independent identically distributed) realizations of this stochastic process, $\{X_t(\tau)\}_{t=0}^{T-1}$. The process $\{X(\tau); \tau \in [0, 1]\}$ is typically modelled as a random element of a separable Hilbert space of functions, often that of square integrable complex functions defined on [0, 1], say $L^2([0, 1], \mathbb{C})$. As such, it admits a Karhunen–Loève decomposition, a spectral representation of the form

$$X(\tau) = m(\tau) + \sum_{n=1}^{\infty} \xi_n \varphi_n(\tau), \tag{1}$$

where $\{\varphi_n\}_{n=1}^{\infty}$ are the orthonormal eigenfunctions of the operator $\mathscr{R}_0$, and $\{\xi_n\}_{n=1}^{\infty}$ are the corresponding uncorrelated Fourier coefficients, $\xi_n = \int_0^1 \varphi_n(\tau)[X(\tau) - m(\tau)]d\tau$, with variance equal to the respective eigenvalue of $\mathscr{R}_0$, say $\lambda_n$. Convergence is in mean square, and can in fact be seen to be uniform over $\tau$, provided $X$ is continuous in mean square. The decomposition is essentially unique (assuming no multiplicities in the eigenvalues), and characterizes the law of $X$.

The functional principal component decomposition (1) is fundamental for a number of reasons. First and foremost, it yields a separation of variables: the stochastic part of $X$, represented by the countable collection $\{\xi_n\}$, is separated from the functional part, represented by the deterministic functions $\{\varphi_n\}$. Furthermore, it provides insight into the smoothness properties of the random function, which are encapsulated in the smoothness of the functions $\varphi_n$, each "relatively contributing" according to the ratio $\lambda_n / \sum_{k \geq 1} \lambda_k$. Finally, it allows for an optimal finite-dimensional approximation of the random function $X$, a *functional Principal Component Analysis*, in that the projection of $X$ onto the space spanned by the first $K$ eigenfunctions $\{\varphi_n\}_{n=1}^{K}$ provides the best $K$-dimensional approximation of $X$ in mean square. As a consequence, the Karhunen–Loève representation has become both the object of and the means for much of the statistical methodology developed for functional data. It has defined what today is accepted as the canonical framework for functional data analysis and has provided a bridge allowing for a technology transfer of tools from multivariate statistics to problems of functional statistics.

As the name suggests, the Karhunen–Loève expansion can be traced back to the work of Karhunen [24] and Loève [27], the former in the context of series representations of Wiener measures and the latter in the context of linear filtering of stochastic processes. From the statistical perspective, Grenander [16] used the countable representation afforded by the expansion as a coordinate system to construct inferential procedures for random functions (perhaps marking the birth of functional data analysis; see also Grenander [17]). Large sample asymptotic properties of the empirical functional principal components, constructed on the basis of an i.i.d. sample $\{X_t(\tau); \tau \in [0, 1]\}_{t=0}^{T-1}$, were considered by Kleffe [25], who proved their consistency for the true functional principal components, and Dauxois et al. [13], who determined their asymptotic distributions. The empirical functional principal components were subsequently put to use to generalize finite-dimensional methods to the functional case, notably by Besse and Ramsay [2] and Rice and Silverman [35], leading on the one hand to a surge in methodological work on functional principal components: smooth components (e.g. Silverman [36]), higher order theory (Hall and Hosseini-Nasab [18,19]), nonparametric and conditional components (e.g. Cardot [10,11]) and components for irregularly sampled functional data (e.g. Yao et al. [38], Hall et al. [20] and

Amini and Wainwright [1]); and on the other hand, to a surge in methodology for functional data *hinging* on Karhunen–Loève representations and the corresponding functional principal component decompositions: functional regression, functional classification, functional testing, and functional robustness, to name a few (see Ramsay and Silverman [34], Ferraty and Romain [15], and Horváth and Kokoszka [22] for detailed overviews). It is interesting to note that, in this body of work, functional principal components arose both as a basis for motivating methodology, but also as a tool to apply regularization (via spectral truncation) to problems such as prediction, testing and regression which are ill-posed in the infinite dimensional case.

Parallel to the development of statistical work for i.i.d. functional samples, an important body of literature on dependent but stationary sequences of random functions $\{X_t(\tau); \tau \in [0, 1]\}_{t=0}^{T-1}$ was developed. In the dependent case, one needs to consider both the *within curve dynamics*, described by the covariance operator $\mathscr{R}_0$ and its spectral decomposition, as well as the *between curve dynamics*, captured (up to second order) by autocovariance operators

$$\mathscr{R}_t := \mathbb{E}[(X_t - m) \otimes (X_0 - m)] \tag{2}$$

and their respective singular value decompositions. Work on this front was pioneered by Bosq [3,4], who focused on the estimation of spectral decompositions of autocovariance operators in the special case of stationary AR(1) functional process, work later extended to more general *linear* stationary functional processes (Mas [28,29] and Bosq [5]). As in the i.i.d. case, these decompositions are interesting in themselves, as they provide variable separation, smoothness information and dimension reduction for the different orders of dependence of the functional process; but they are also fundamental as a stepping stone for the development of time series methodology used in prediction, filtering, order estimation and change detection, to name only a few (see Bosq [6] and Bosq and Blanke [7] for an overview, and the recent review by Mas and Pumo [31]). However, many of these decompositions concern *isolated* aspects of the functional process: no single autocovariance operator captures the global dynamics of the functional process, and the "sum" of all separate spectral decompositions of each autocovariance operator does not provide a coherent simultaneous spectral decomposition of the entire second-order dynamics of the process; for example, there is no clear representation of the original series in terms of these separate decompositions. Furthermore, most of the work carried out thus far has assumed the linearity of the underlying process—an assumption that is often reasonable, but by no means a weak one. More recent work has attempted to move away from the linear process model. Hörmann and Kokoszka [21] consider the estimation of decompositions of $\mathscr{R}_0$ under general mixing conditions, without assuming linearity. Under similar assumptions, Horváth et al. [23] estimate the long-run covariance operator, an average of all covariance operators. Panaretos and Tavakoli [32] introduce a frequency-domain approach and estimate the complete second order structure of the process, by means of estimators of the *spectral density operator*. See Kokoszka [26] for a review of recent developments in dependent functional data.

The purpose of this paper is to develop spectral representations for stationary sequences that simultaneously capture both the *within curve dynamics* (the dynamics of the curve $\{X_0(\tau) : \tau \in [0, 1]\}$) as well as the *between curve dynamics* (the dynamics of the sequence $\{X_t : t \in \mathbb{Z}\}$) and to estimate them without assuming any prior structural properties for the stationary sequence (e.g. linearity or Gaussianity) except some weak dependence conditions based on cumulants. Our approach is a frequency domain one, exploiting the notion of a spectral density operator

$$\mathscr{F}_\omega := \frac{1}{2\pi} \sum_t \mathscr{R}_t e^{-\mathrm{i}\omega t}, \quad \omega \in [-\pi, \pi],$$

the Fourier transform of the autocovariance operators (with respect to the lag argument) introduced and investigated in Panaretos and Tavakoli [32].

Assuming without loss of generality that $m = 0$, we develop end estimate a doubly spectral representation of the functional process $\{X_t(\tau)\}$ of the form

$$X_t(\tau) = \left[\int_{-\pi}^{\pi} e^{\mathrm{i}\omega t}\left(\sum_{n=1}^{\infty}\varphi_n^{\omega}\otimes\varphi_n^{\omega}\right)dZ_{\omega}\right](\tau),$$

which can be formally represented as

$$X_t(\tau) = \int_{-\pi}^{\pi} e^{\mathrm{i}\omega t}\sum_{n=1}^{\infty}\langle\varphi_n^{\omega}, dZ_{\omega}\rangle\varphi_n^{\omega}(\tau),$$

and which we call a *Cramér–Karhunen–Loève* representation; here, for each frequency $\omega$, $\{\varphi_n^{\omega}(\tau)\}_{n=1}^{\infty}$ is an orthonormal basis of $L^2([0,1],\mathbb{C})$ comprised of eigenfunctions of the spectral density operator $\mathscr{F}_{\omega}$, and the Fourier coefficients $\{\langle\varphi_n^{\omega}, dZ_{\omega}\rangle\}_{n=1}^{\infty}$ are uncorrelated (both w.r.t. $n$ and $\omega$) random variables with variance equal to the $n$th eigenvalue of $\mathscr{F}_{\omega}$, say $\mu_n(\omega)$.

Similarly with the Karhunen–Loève decomposition, such a representation yields a separation of variables (separating the functional from the stochastic component); it provides insight on the smoothness properties of the random functions, and how these interact with the dependence structure of the sequence: not only does it decompose the process into uncorrelated functional frequency components, but it reveals which basis is optimal to represent each frequency component (by looking at the corresponding eigenfunctions $\{\varphi_n(\omega)\}$), and what the effective dimensionality of each frequency component is (by looking at the relative decay of the corresponding eigenvalues $\{\mu_n(\omega)\}$); finally, it serves to yield an optimal reduction of the process $\{X_t\}$, to a process with only $K$ degrees of freedom which nevertheless captures its temporal dynamics, by truncation of the series inside the integral at a finite level $K$,

$$\int_{-\pi}^{\pi} e^{\mathrm{i}\omega t}\sum_{n=1}^{K}\langle\varphi_n^{\omega}, dZ_{\omega}\rangle\varphi_n^{\omega}(\tau),$$

providing a *Harmonic Functional Principal Component Analysis* of the process $\{X_t\}$.

Further to developing these formal representations, we consider the explicit construction of their empirical counterparts, on the basis of a finite stretch of length $T$ of the time series $\{X_t\}$. We derive the asymptotic distributions of the different elements of the empirical representation, obtaining analogues of the results that Dauxois et al. [13] for i.i.d. functional data. The paper is organized in two parts. The first part (Part (I)) provides an accessible heuristic motivation and overview of the main results. The second part (Part (II)) provides the corresponding rigorous statements and their formal derivations.

(I) *Heuristic outline of the main results*

The autocovariance operators $\{\mathscr{R}_t\}_{t\in\mathbb{Z}}$ of a second-order stationary time series $X_t \in L^2([0,1],\mathbb{R})$ (defined in (2)) encode the complete second-order structure of the time series $\{X_t\}_{t\in\mathbb{Z}}$, assumed to have mean zero. Corresponding to this sequence of operators is a collection of *spectral density operators* $\{\mathscr{F}_{\omega}\}_{\omega\in[-\pi,\pi]}$, defined as their discrete-time Fourier transform, and yielding the Fourier pair

$$\mathscr{F}_{\omega} = \frac{1}{2\pi}\sum_{t\in\mathbb{Z}} e^{-\mathrm{i}\omega t}\mathscr{R}_t, \quad \mathscr{R}_t = \int_{-\pi}^{\pi}\mathscr{F}_{\omega}e^{\mathrm{i}t\omega}d\omega, \tag{3}$$

provided that the $\{\mathscr{R}_t\}$ are summable in an appropriate sense (a rigorous definition – and sufficient conditions for its validity – will be given in Section 1.2). Now, assume that we can approximate the integral in the second equation in (3) by a Riemann sum, to get

$$\mathscr{R}_s = \int_{-\pi}^{\pi} e^{\mathrm{i}s\omega}\mathscr{F}_\omega d\omega \approx \sum_{j=1}^{J} \mathscr{F}_{\omega_j} e^{\mathrm{i}s\omega_j}(\omega_{j+1}-\omega_j),$$

where $-\pi = \omega_1 < \cdots < \omega_{J+1} = \pi$ is a partition. Then, we are naturally tempted to conjecture that $X_t$ ought to be decomposable into a sum of distinct and uncorrelated frequency components,

$$X_t \approx \sum_{j=1}^{J} e^{\mathrm{i}\omega_j t} X_t(\omega_j) \tag{4}$$

where each $X_t(\omega_j)$ would be a mean-zero functional time series in $L^2([0,1],\mathbb{C})$ with covariance operator close to $\mathscr{F}_{\omega_j}(\omega_{j+1}-\omega_j)$, since, in this case, $X_t$ would indeed have covariance $\mathscr{R}_t = \sum_{j=1}^{J}\mathscr{F}_{\omega_j}e^{\mathrm{i}t\omega_j}(\omega_{j+1}-\omega_j)$. We pursue such a decomposition in Section 2, where we formalize it as the functional Cramér representation (Theorem 2.1),

$$X_t = \int_{-\pi}^{\pi} e^{\mathrm{i}\omega t} dZ_\omega, \quad \text{a.s. a.e.,}$$

for a functional orthogonal increment process $Z$ (independent of $t$), thus extending the classical Cramér representation of multivariate stationary processes (e.g. Brillinger [8]).

The Cramér representation provides a spectral decomposition with respect to frequency. Nevertheless, we may pursue a second "layer" of spectral decomposition, one in terms of dimension. Going back to the heuristic form (4), we notice that, for each $j = 1,\ldots,J$, $X_t(\omega_j)$ is a random element of $L^2([0,1],\mathbb{C})$. We may thus represent it through its Karhunen–Loève (KL) expansion (1), leading to the heuristic representation

$$X_t \approx \sum_{j=1}^{J} e^{\mathrm{i}\omega_j t}\sum_{i=1}^{\infty}\xi_{i,j}\varphi_{i,j}(\tau), \tag{5}$$

with $\{\varphi_{i,j}\}_{i\geq 1}$ being the eigenfunctions of the covariance operator of $X_t(\omega_j)$ and $\{\xi_{i,j}\}_{i\geq 1}$ the corresponding Fourier coefficients. Truncating the second series at some $K < \infty$, will yield a decomposition into distinct frequency elements that are uncorrelated, and finite dimensional,

$$X_t \approx \sum_{j=1}^{J} e^{\mathrm{i}\omega_j t}\sum_{i=1}^{K}\xi_{i,j}\varphi_{i,j}(\tau). \tag{6}$$

The finite dimensional subspace in which each frequency component takes its values need not be the same for distinct $j$'s, even though each of them is of dimension $K$. In fact, it will turn out that this truncated representation only possesses $K$ degrees of freedom. One would then hope that this reduced version of $X_t(\omega_j)$ would retain the property of being the optimal (in the $L^2$ sense) $K$-dimensional reduction of the process $X_t$. Rigorous versions of decomposition (5), and its truncated version (6), are formally carried out in Section 2. Specifically, we derive the *Cramér–Karhunen–Loève* decomposition

$$X_t = \int_{-\pi}^{\pi} e^{\mathrm{i}\omega t}\left(\sum_{n=1}^{\infty}\varphi_n^\omega\otimes\varphi_n^\omega\right)dZ_\omega = \int_{-\pi}^{\pi} e^{\mathrm{i}\omega t}\sum_{n=1}^{\infty}\langle\varphi_n^\omega, dZ_\omega\rangle\varphi_n^\omega, \tag{7}$$

where the last equality is understood formally (Remarks 2.4 and 3.10). This is a Cramér representation with respect to frequency, but also a Karhunen–Loève expansion in terms of dimension, since it can be seen that $\{\varphi_n^\omega\}_{n\geq 1}$ is the basis of eigenfunctions of $\mathscr{F}_\omega$ (the covariance operator of $dZ_\omega$). Furthermore, by considering the bounded operator-valued function $\sum_{n=1}^K \varphi_n^\omega(\tau)\otimes\varphi_n^\omega(\sigma)$ as a function over $[-\pi,\pi]$, and defining the notion of its stochastic integral (Section 3.1), we show that the truncated representation

$$X_t^* = \int_{-\pi}^{\pi} e^{\mathrm{i}\omega t}\left(\sum_{n=1}^{K}\varphi_n^\omega\otimes\varphi_n^\omega\right) dZ_\omega \tag{8}$$

is well defined, possesses $K$ degrees of freedom, and converges to $X_t$ in mean square as $K\to\infty$. More importantly, by considering the process $X_t^*$ for different values of $K$, we obtain a *harmonic principal component analysis* of $X_t$ (Section 3.2). That is, we prove (Theorem 3.7 and Remark 3.10) that, among all linear reductions $X_t$ to a process $W_t$ of only $K$ degrees of freedom, we have

$$\mathbb{E}\|X_t - X_t^*\|^2 \leq \mathbb{E}\|X_t - W_t\|^2.$$

Section 3.2 explains how the process $\{X_t^*\}$ can be constructed explicitly, when the spectral density estimator $\mathscr{F}_\omega$ is known, and how it can be represented as a stationary vector valued process *with uncorrelated coordinates* in $\mathbb{R}^K$.

Parallel to the rank $K$ reduction, one may want to have a better finite dimensional approximation of $X_t(\omega_j)$ for some $j$'s, and a cruder one for other $j$'s, depending on how much each $\omega_j$ contributes to the power of the signal and/or the effective dimension of each $X_t(\omega_j)$. This can be done by letting the dimension $K$ vary with $j$, leading to the heuristic approximation

$$X_t \approx \sum_{j=1}^{J} e^{\mathrm{i}\omega_j t} X_t^{K_j}(\omega_j) = \sum_{j=1}^{J} e^{\mathrm{i}\omega_j t}\sum_{i=1}^{K_j}\xi_{i,j}\varphi_{i,j}, \tag{9}$$

where $X_t^{K_j}(\omega_j)$ is $K_j$-dimensional. It will turn out that such a representation is also rigorously valid (Section 3.1), and of the form

$$X_t^{**} = \int_{-\pi}^{\pi} e^{\mathrm{i}\omega t}\left(\sum_{n=1}^{K(\omega)}\varphi_n^\omega\otimes\varphi_n^\omega\right) dZ_\omega$$

provided that the function $K:[-\pi,\pi]\to\{0,1,\ldots\}$ yielding the desired finite rank for each frequency component is cadlag. In fact, it will be shown, that among all linear transformations of the process $\{X_t\}$ having finite rank $K(\omega)$ at each frequency component, this is the optimal one, in the $L^2$ sense (Theorem 3.7).

The final part of the paper (Section 4) is devoted to the construction of empirical versions of the representations presented, on the basis of a finite stretch of the time series, say $\{X_t\}_{t=0}^T$, $T<\infty$. These require estimates of the eigenfunctions and eigenvalues of the spectral density operator. We estimate these by the eigenfunctions and eigenvalues of the estimate $\mathscr{F}_\omega^{(T)}$ of the spectral density operator $\mathscr{F}_\omega$ introduced by Panaretos and Tavakoli [32], constructed via the discrete Fourier transform of the observed time series. We derive the asymptotic distributions of the estimated eigenfunctions and eigenvalues, showing that they are jointly asymptotically normal for a finite number of distinct frequencies $\omega\in[0,\pi]$. Moreover, the estimators are independent for distinct frequencies, and, in fact, the estimators of the eigenvalues are independent between different orders at the same frequency $\omega$, and independent of the eigenfunctions (see Theorem 4.3). Some technical material is contained in Section 5 (and in Panaretos and Tavakoli [33]).

(II) *Formal statements and their proofs*

## 1. Background material

### 1.1. Notation

We shall denote by $H$ the Hilbert space $L^2([0,1],\mathbb{C})$, and by $\mathbb{H}$ the Hilbert space $L^2(\Omega, H, \mathbb{P})$ of $H$-valued random variables with finite second moment, where $\mathbb{P}$ is the underlying probability measure. Their norms and inner products will be denoted by $\langle\cdot,\cdot\rangle$, $\|\cdot\|$ and $\langle\cdot,\cdot\rangle_{\mathbb{H}}$, $\|\cdot\|_{\mathbb{H}}$, respectively. We denote by $\mathcal{S}_p(H)$ the Schatten $p$-class of operators on $H$, and denote its norm by $|\!|\!|\cdot|\!|\!|_p$. The cases $p=1$, $p=2$ and $p=\infty$ correspond to the spaces of nuclear (or trace-class), Hilbert–Schmidt, and bounded operators on $H$, respectively. If $T \in \mathcal{S}_1(H)$, we will denote its trace by trace $(T)$. The space $\mathcal{S}_2(H)$ being a Hilbert space, we will denote its inner product by $\langle\cdot,\cdot\rangle_{\mathcal{S}_2}$. We will denote the identity operator on $H$ by $I$, and, for operators $A$ and $B$, the term $AB$ will denote the composition of $A$ and $B$, and $A^\dagger$ will denote the adjoint of $A$. If $A$ is an integral operator, we will sometimes abuse notation and denote its kernel by $A(\tau,\sigma)$. Hence if $A$, $B$ are integral operators on $H$, then the kernel of $AB$ is

$$AB(\tau,\sigma) = \int_0^1 A(\tau,x)B(x,\sigma)dx.$$

We will denote the imaginary number by $\mathring{\imath} \in \mathbb{C}$. For a set $S \subset [-\pi,\pi]$, we define the indicator function of $S$ by $\mathbf{1}_S$. For $u, v \in H$, we define the *tensor product* between $u$ and $v$, $u \otimes v \in \mathcal{S}_2(H)$, by $u \otimes v(f) = \langle f, v\rangle u$, for $f \in H$. Notice that for $A, B \in \mathcal{S}_2(H)$, we can also define $A \otimes B \in \mathcal{S}_2(\mathcal{S}_2(H))$. To avoid confusion, we will denote the latter by $A \bigotimes B$. We also define the *Kronecker product* $A \widetilde{\otimes} B \in \mathcal{S}_2(\mathcal{S}_2(H))$ by $A\widetilde{\otimes} B(C) = ACB^\dagger$, for $C \in \mathcal{S}_2(H)$. If $H_{\mathbb{R}}$ is a real Hilbert space, then for two operators $A, B \in \mathcal{S}_2(H_{\mathbb{R}})$, we also define their *transpose Kronecker product* $A \widetilde{\otimes}_{\mathsf{T}} B \in \mathcal{S}_2(\mathcal{S}_2(H_{\mathbb{R}}))$, by $A\widetilde{\otimes}_{\mathsf{T}} B(C) = (A\widetilde{\otimes} B)(C^{\mathsf{T}}) = AC^{\mathsf{T}}B^{\mathsf{T}}$, for $C \in \mathcal{S}_2(H_{\mathbb{R}})$. Here, $A^{\mathsf{T}}$ also denotes the adjoint of $A$, but stresses that it operates on a real Hilbert space. Useful properties of the tensor, Kronecker and transpose Kronecker product are given in Section 5.1.

### 1.2. Basic definitions and assumptions

Our basic assumptions concerning the smoothness of the curves $\{X_t(\tau)\}$ and the strength of dependence between the elements of the sequence $\{X_t\}$ follow those in Panaretos and Tavakoli [32]. In particular, we will assume that the following conditions hold:

**Conditions 1.1** (*And Definitions*). *$X_t$ is a second order stationary (temporally translation invariant only up to the second order characteristics of its law), times series in $L^2([0,1],\mathbb{R})$, with mean zero, and $\mathbb{E}\|X_0\|^2 < \infty$. Its autocovariance kernel at lag $t$ will be defined as*

$$r_t(\tau,\sigma) = \mathbb{E}[X_t(\tau)X_0(\sigma)], \quad \tau,\sigma \in [0,1] \text{ and } t \in \mathbb{Z},$$

*inducing a corresponding operator* $\mathscr{R}_t : L^2([0,1],\mathbb{R}) \to L^2([0,1],\mathbb{R})$ *by right integration, the autocovariance operator at lag t,*

$$\mathscr{R}_t h = \mathbb{E}\left[X_t \langle h, X_0\rangle\right], \quad h \in L^2([0,1],\mathbb{R}).$$

*Furthermore, we assume the following conditions hold:*

(i) $\sum_{t\in\mathbb{Z}} |||\mathscr{R}_t|||_1 < \infty$.
(ii) $(\tau,\sigma) \mapsto r_t(\tau,\sigma)$ *is continuous* $\forall t \in \mathbb{Z}$, *and* $\sum_{t\in\mathbb{Z}} \|r_t\|_\infty < \infty$.

Panaretos and Tavakoli [32] discuss the role of these conditions, and at what cost they may be weakened. Under these conditions, we define the spectral density kernel,

$$f_\omega(\cdot,\cdot) = \frac{1}{2\pi}\sum_{t\in\mathbb{Z}} r_t(\cdot,\cdot),$$

where the convergence is uniform, and we denote by $\mathscr{F}_\omega$ the operator on $L^2([0,1],\mathbb{C})$ induced by $f_\omega$. The autocovariance kernel can be recovered from the spectral density kernels by means of the following inversion formula:

$$\int_{-\pi}^{\pi} f_\omega(\tau,\sigma)e^{\mathrm{i}\omega t}d\omega = r_t(\tau,\sigma), \quad \forall t \in \mathbb{Z} \text{ and } \tau,\sigma \in [0,1].$$

Each $\mathscr{F}_\omega$ is trace-class, $\omega \mapsto |||\mathscr{F}_\omega|||_1$ is uniformly continuous, $f_\omega(\tau,\sigma)$ is continuous in $\omega,\tau,\sigma$; $\mathscr{F}_\omega$ is non-negative definite, and thus

$$\int_0^1 f_\omega(\tau,\tau)d\tau = |||\mathscr{F}_\omega|||_1.$$

Furthermore, $\mathscr{F}_\omega$ is self-adjoint, $2\pi$-periodic with respect to $\omega$, and

$$f_{-\omega}(\tau,\sigma) = \overline{f_\omega(\tau,\sigma)} = f_\omega(\sigma,\tau).$$

The reader is referred to Panaretos and Tavakoli [32] for proofs of these assertions. The spectral density operator thus admits a singular value decomposition of the form

$$\mathscr{F}_\omega = \sum_{j=1}^{\infty} \mu_j(\omega)\varphi_j^\omega \otimes \varphi_j^\omega, \tag{10}$$

where $\{\mu_j(\omega)\}_{j\geq 1}$ is a non-increasing sequence of positive real numbers, tending to zero and $\{\varphi_n^\omega\}_{n\geq 1}$ is an orthonormal system $L^2([0,1],\mathbb{C})$. When $\mathscr{F}_\omega$ is strictly positive-definite, the orthonormal system $\{\varphi_n^\omega\}_{n\geq 1}$ is, in fact, complete for $L^2([0,1],\mathbb{C})$.

## 2. Towards a Cramér–Karhunen–Loève representation

We begin by deriving a functional version of the classical spectral representation of a stationary time series, thus extending Cramér's representation [12] to the infinite dimensional case.

**Theorem 2.1** (*Functional Cramér Representation*)**.** *Under Conditions 1.1,* $X_t$ *admits the representation*

$$X_t = \int_{-\pi}^{\pi} e^{\mathrm{i}\omega t} dZ_\omega, \quad \textit{a.s. a.e.,} \tag{11}$$

*where for fixed $\omega$, $Z_\omega$ is random element of $L^2([0,1],\mathbb{C})$ with $\mathbb{E}\,\|Z_\omega\|_2^2 = \int_{-\pi}^{\omega} |||\mathscr{F}_\alpha|||_1 d\alpha$, and the process $\omega \mapsto Z_\omega$ has orthogonal increments:*

$$\mathbb{E}\left\langle Z_{\omega_1} - Z_{\omega_2}, Z_{\omega_3} - Z_{\omega_4}\right\rangle = 0, \quad \text{if } \omega_1 > \omega_2 \geq \omega_3 > \omega_4. \tag{12}$$

*The representation (11) is called the Cramér representation of $X_t$, and the stochastic integral involved can be understood as a Riemann–Stieltjes limit, in the sense that*

$$\mathbb{E}\left\| X_t - \sum_{j=1}^{J} e^{\mathrm{i}\omega_j t}(Z_{\omega_{j+1}} - Z_{\omega_j}) \right\|^2 \to 0, \quad \text{as } J \to \infty, \tag{13}$$

*where $-\pi = \omega_1 < \cdots < \omega_{J+1} = \pi$ and $\max_{j=1,\ldots,J} |\omega_{j+1} - \omega_j| \to 0$ as $J \to \infty$.*

**Remark 2.2.** Consider the particular case where $\{X_t\}$ is a linear process, i.e. $X_t = \sum_{t\in\mathbb{Z}} A_{t-s}\varepsilon_s$, with $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ an i.i.d. sequence of random elements of $L^2([0,1],\mathbb{R})$ and $\{A_t\}_{t\in\mathbb{Z}}$ a sequence of linear operators in $\mathcal{S}_2(H)$; then, Conditions 1.1 will be satisfied if: (1) the $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ are mean square continuous and $\mathbb{E}\|\varepsilon_t\|^2 < \infty$, $\forall t \in \mathbb{Z}$, (2) the $\{A_t\}_{t\in\mathbb{Z}}$ admit continuous kernels and satisfy $\sum_{t\in\mathbb{Z}} |||A_t|||_2 < \infty$.

**Proof of Theorem 2.1.** Let $\mathbb{M}_0$ be the complex linear space spanned by all finite linear combinations of the $X_t$'s,

$$\mathbb{M}_0 := \left\{ \sum_{j=1}^{n} a_j X_{t_j} : n = 1, 2, \ldots; a_j \in \mathbb{C}, t_j \in \mathbb{Z} \right\} \subset \mathbb{H},$$

and $\mathbb{M}$ be the closure of $\mathbb{M}_0$ in $\mathbb{H}$. Let $H'$ be the space of all measurable complex functions $g : [0,1] \to \mathbb{C}$ such that

$$\int_0^1 |g(\alpha)|^2 |||\mathscr{F}_\alpha|||_1 d\alpha < \infty,$$

where $|||\mathscr{F}_\alpha|||_1$ is the nuclear norm of the spectral density operator $\mathscr{F}_\alpha$. We endow $H'$ with the inner product

$$\langle g, h\rangle_{H'} = \int_0^1 g(\alpha)\overline{h(\alpha)} |||\mathscr{F}_\alpha|||_1 d\alpha, \quad g, h \in H',$$

which makes $H'$ a Hilbert space. Now, let $e_t \in H'$ denote the function $\alpha \mapsto e^{\mathrm{i}t\alpha}$. We define an operator $T : \mathbb{M}_0 \to H'$ by linear extension of the mapping $X_t \mapsto e_t$, that is

$$T\left(\sum_{j=1}^{n} a_j X_{t_j}\right) = \sum_{j=1}^{n} a_j e_{t_j}, \tag{14}$$

for any $a_j \in \mathbb{C}$, and $t_j \in \mathbb{Z}$. Using the inversion formula, $\langle X_t, X_s\rangle_{\mathbb{H}} = \langle e_t, e_s\rangle_{H'}$, $t, s \in \mathbb{Z}$, and hence for any $A, B \in \mathbb{M}_0$, $\langle T(A), T(B)\rangle_{H'} = \langle A, B\rangle_{\mathbb{H}}$. In particular, $T$ is well defined and is a linear isometry. We extend its domain to $\mathbb{M}$ in the following way: for any $A \in \mathbb{M}$, let $T(A)$ be the limit in $H'$ of $T(A_n)$, where $\{A_n\}_{n\geq 1} \subset \mathbb{M}_0$ is a sequence converging to $A$. If $\{A'_n\}_{n\geq 1} \subset \mathbb{M}_0$ is another sequence converging to $A$, then

$$\left\|T(A_n) - T(A'_n)\right\|_{H'} = \left\|A_n - A'_n\right\|_{\mathbb{H}} \to 0 \quad \text{as } n \to \infty,$$

and hence the extension of $T$ is well defined. It is also linear, maintains the isometry property on the entire $\mathbb{M}$, and satisfies $T(\mathbb{M}) = H'$ since the linear space of the functions $e_t$ is dense in $H'$. Hence $T$ admits a well-defined inverse, say $T^{-1} : H' \to \mathbb{M}$.

For any $\omega \in (-\pi, \pi]$, we define $Z_\omega = T^{-1}(\mathbf{1}_{[-\pi,\omega)}) \in \mathbb{M}$, and also $Z_{-\pi} \equiv 0 \in \mathbb{M}$. By the isometry property of $T$,

$$\left\langle Z_\omega, Z_\beta \right\rangle_{\mathbb{H}} = \left\langle T^{-1}\mathbf{1}_{[-\pi,\omega)}, T^{-1}\mathbf{1}_{[-\pi,\beta)} \right\rangle_{\mathbb{H}} = \int_{-\pi}^{\min(\omega,\beta)} |||\mathscr{F}_\alpha|||_1 d\alpha. \tag{15}$$

It follows that $\omega \mapsto Z_\omega$ is an orthogonal increment process. We shall now define the integral with respect to $Z_\omega$. Let $D \subset H'$ be the space of simple functions of the form

$$g = \sum_{j=1}^{n} g_j \mathbf{1}_{[\omega_j,\omega_{j+1})},$$

where $g_j \in \mathbb{C}$ and $-\pi = \omega_1 < \omega_2 < \cdots < \omega_{n+1} = \pi$. We equip $D$ with the scalar product of $H'$, and define $\phi : D \to \mathbb{H}$ by

$$\phi\left(\sum_{j=1}^{n} g_j \mathbf{1}_{[\omega_j,\omega_{j+1})}\right) = \sum_{j=1}^{n} g_j (Z_{\omega_{j+1}} - Z_{\omega_j}).$$

By (15), $\phi$ is isomorphism, whose domain is $\overline{D} = H'$. Moreover, it is straightforward to see that $\phi$ is equal to $T^{-1}$ on $D$, hence $\phi = T^{-1}$. This in turn implies $X_t = T^{-1}(e_t) = \phi(e_t)$. If $g$ is cadlag with a finite number of jumps, then $\phi(g)$ is in fact the Riemann–Stieltjes integral (in the mean square sense) with respect to the orthogonal increment process $Z_\omega$:

$$\phi(g) = \int_{-\pi}^{\pi} g(\alpha) dZ_\alpha, \quad g \in H'. \tag{16}$$

In other words, we have shown that $X_t = \int_{-\pi}^{\pi} e^{it\alpha} dZ_\alpha$, as claimed. □

As expected from the finite-dimensional version of the Cramér representation, the second-order properties of the orthogonal increment process $Z$ are inextricably linked with the spectral density operator of $\{X_t\}$. In fact, we may prove that these properties extend to the functional case not only in an $L^2$ sense, but in a stronger sense:

**Proposition 2.3** (*Second-Order Properties of the Orthogonal Increment Process*)**.** *Let $\alpha \mapsto Z_\alpha$ be the process defined in Theorem 2.1, and assume Conditions 1.1 hold. If $\pi \geq \omega_1 > \omega_2 \geq \omega_3 > \omega_4 \geq -\pi$, then*

$$\mathbb{E}\left[(Z_{\omega_1}(\tau) - Z_{\omega_2}(\tau))\overline{(Z_{\omega_3}(\sigma) - Z_{\omega_4}(\sigma))}\right] = 0 \quad a.e.,$$

*and for $\pi \geq \omega > \beta > -\pi$, we have*

$$\mathbb{E}\left[(Z_\omega(\tau) - Z_\beta(\tau))\overline{(Z_\omega(\sigma) - Z_\beta(\sigma))}\right] = \int_\beta^\omega f_\alpha(\tau, \sigma) d\alpha, \quad a.e.$$

**Proof.** It suffices to show that

$$\mathbb{E}\left[Z_\omega(\tau)\overline{Z_\beta(\sigma)}\right] = \int_{-\pi}^{\min(\omega,\beta)} f_\alpha(\tau, \sigma) d\alpha, \quad \text{a.e.}$$

Let $B' = L^1([0,1]^2, \mathbb{C})$ with norm

$$\|g\|_{B'} = \iint_{[0,1]^2} |g(\tau,\sigma)| d\tau d\sigma, \quad g \in B'.$$

We will show that for any $A_1, A_2 \in \mathbb{M}$,

$$\left\| \mathbb{E}[A_1 \otimes A_2] - \int_{-\pi}^{\pi} T(A_1)(\alpha)\overline{T(A_2)(\alpha)} f_\alpha \right\|_{B'} = 0, \tag{17}$$

where

$$\mathbb{E}[A_1 \otimes A_2](\tau,\sigma) = \mathbb{E}[A_1 \otimes A_2(\tau,\sigma)] = \mathbb{E}[A_1(\tau)A_2(\sigma)].$$

First, let us mention two properties that will simplify the proof:

(i) $\|\mathbb{E}[A_1 \otimes A_2]\|_{B'} \le \|A_1\|_{\mathbb{H}} \|A_2\|_{\mathbb{H}}$,
(ii) $\left\| \int_{-\pi}^{\pi} T(A_1)(\alpha)\overline{T(A_2)(\alpha)} f_\alpha \right\|_{B'} \le \|A_1\|_{\mathbb{H}} \|A_2\|_{\mathbb{H}}$.

Property (i) is a consequence of the Cauchy–Schwarz inequality. Property (ii) follows from

$$\begin{aligned}
&\left\| \int_{-\pi}^{\pi} T(A_1)(\alpha)\overline{T(A_2)(\alpha)} f_\alpha \right\|_{B'} \\
&= \iint_{[0,1]^2} \left| \int_{-\pi}^{\pi} T(A_1)(\alpha)\overline{T(A_2)(\alpha)} f_\alpha(\tau,\sigma) d\alpha \right| d\tau d\sigma \\
&\le \int_{-\pi}^{\pi} \left( \iint_{[0,1]^2} \left| f_\alpha(\tau,\sigma) \right| d\tau d\sigma \right) \left| T(A_1)(\alpha)\overline{T(A_2)(\alpha)} \right| d\alpha \\
&\le \int_{-\pi}^{\pi} |||\mathscr{F}_\alpha|||_1 \left| T(A_1)(\alpha)\overline{T(A_2)(\alpha)} \right| d\alpha \\
&\le \|T(A_1)\|_{H'} \|T(A_2)\|_{H'} \\
&= \|A_1\|_{\mathbb{H}} \|A_2\|_{\mathbb{H}},
\end{aligned}$$

where we have used the Cauchy–Schwarz inequality and the fact that the Hilbert–Schmidt norm is dominated by the nuclear norm. The operator $T$ and the Hilbert space $H' = L^2([-\pi,\pi], \mathbb{C}, |||\mathscr{F}_\alpha|||_1 d\alpha)$ are defined in the proof of Theorem 2.1. We can now show (17). Let $A_{1,n}, A_{2,n} \subset \mathbb{M}_0$ be sequences converging to $A_1$, respectively $A_2$ in the norm of $\mathbb{H}$. Using the triangle inequality, (17) is bounded above by the sum of

$$\left\| \mathbb{E}[A_1 \otimes A_2] - \mathbb{E}\left[A_{1,n} \otimes A_{2,n}\right] \right\|_{B'} \tag{18}$$

and

$$\left\| \mathbb{E}\left[A_{1,n} \otimes A_{2,n}\right] - \int_{-\pi}^{\pi} T(A_1)(\alpha)\overline{T(A_2)(\alpha)} f_\alpha \right\|_{B'}. \tag{19}$$

By (i), we may make (18) sufficiently small by choice of a sufficiently large $n$. Indeed,

$$\begin{aligned}
\left\| \mathbb{E}[A_1 \otimes A_2] - \mathbb{E}\left[A_{1,n} \otimes A_{2,n}\right] \right\|_{B'} \le{}& \left\|A_1 - A_{1,n}\right\|_{\mathbb{H}} \|A_2\|_{\mathbb{H}} \\
&+ \left\|A_{1,n}\right\|_{\mathbb{H}} \left\|A_2 - A_{2,n}\right\|_{\mathbb{H}} < \varepsilon/2,
\end{aligned}$$

for $n$ large enough. Let us now show that (19) is bounded by $\varepsilon/2$ for large $n$. Notice that

$$\mathbb{E}[X_t(\tau)X_s(\sigma)] = \int_{-\pi}^{\pi} T(X_t)(\alpha)\overline{T(X_s)(\alpha)} f_\alpha(\tau,\sigma) d\alpha,$$

hence by linearity, we have

$$\mathbb{E}\left[A_{1,n}(\tau)A_{2,n}(\sigma)\right] = \int_{-\pi}^{\pi} T(A_{1,n})(\alpha)\overline{T(A_{2,n})(\alpha)}f_\alpha(\tau,\sigma)d\alpha.$$

Thus (19) is bounded by

$$\left\|\int_{-\pi}^{\pi} T(A_{1,n}-A_1)(\alpha)\overline{T(A_{2,n})(\alpha)}f_\alpha\right\|_{B'} + \left\|\int_{-\pi}^{\pi} T(A_1)(\alpha)\overline{T(A_{2,n}-A_2)(\alpha)}f_\alpha\right\|_{B'}.$$

Using (ii), this is in turn bounded by

$$\left\|A_{1,n}-A_1\right\|_{\mathbb{H}}\left\|A_{2,n}\right\|_{\mathbb{H}} + \|A_1\|_{\mathbb{H}}\left\|A_{2,n}-A_2\right\|_{\mathbb{H}} < \varepsilon/2$$

for $n$ large enough. □

**Remark 2.4** (*A Cramér–Karhunen–Loève Representation*)**.** Assuming Conditions 1.1 and that, for all $\omega \in [-\pi,\pi]$, the spectral density operator $\mathscr{F}_\omega$ is strictly positive-definite, we may now straightforwardly write

$$X_t = \int_{-\pi}^{\pi} e^{\mathring{\mathrm{i}}\omega t}\left(\sum_{n=1}^{\infty}\varphi_n^\omega\otimes\varphi_n^\omega\right)dZ_\omega, \tag{20}$$

where $\{Z_\omega\}$ is as in Theorem 2.1, and $\{\varphi_n^\omega\}_{n\geq 1}$ are the eigenfunctions of $\mathscr{F}_\omega$.

Though the representation in Remark 2.4 is a mere reformulation of Theorem 2.1, it sets the scene for the natural question of the nature of the approximation of $\{X_t\}$ that might arise if we where able to truncate the identity operator $\sum_{n=1}^{\infty}\varphi_n^\omega\otimes\varphi_n^\omega$ to have finite rank

$$X_t^* := \int_{-\pi}^{\pi} e^{\mathring{\mathrm{i}}\omega t}\left(\sum_{n=1}^{K}\varphi_n^\omega\otimes\varphi_n^\omega\right)dZ_\omega, \tag{21}$$

i.e. to consider the limiting behaviour of $\mathbb{E}\|X_t - X_t^*\|^2$ as $K \to \infty$ (see Remark 3.10). It is such truncations (and their approximation error) that are at the essence of representations of the Karhunen–Loève type. This is the subject of the next section, where we provide a formalism to make sense of an integral of the form (21) (it is not a priori clear that it is well-defined, since now the operator in the integrand depends on $\omega$), and prove that it provides an optimal rank $K$ approximation of the original process, yielding a *harmonic principal component analysis* of the process $X_t$ (in fact, we will not be requiring that $\mathscr{F}_\omega$ be strictly positive).

**Remark 2.5.** Note that the action of the operator $\sum_{n=1}^{\infty}\varphi_n^\omega\otimes\varphi_n^\omega$ on an element $g \in L^2([0,1],\mathbb{C})$ is described by $\left[\sum_{n=1}^{\infty}\varphi_n^\omega\otimes\varphi_n^\omega\right]g = \sum_{n=1}^{\infty}\langle\varphi_n^\omega, g\rangle\varphi_n^\omega$. Therefore, we may formally interpret the Cramér–Karhunen–Loève representation as

$$X_t = \int_{-\pi}^{\pi} e^{\mathring{\mathrm{i}}\omega t}\sum_{n=1}^{\infty}\langle\varphi_n^\omega, dZ_\omega\rangle\varphi_n^\omega,$$

a form which emphasizes the doubly spectral decomposition of $\{X_t\}$ as discussed in Part (I) and the introduction.

## 3. Harmonic principal component analysis

### 3.1. Stochastic integrals of operators

Let $\omega \mapsto A_\omega$ be a mapping $[-\pi, \pi] \to \mathcal{S}_\infty(H)$. In order to make sense of expressions such as (21), we wish to give a meaning to the stochastic integral $\int_{-\pi}^{\pi} A_\omega dZ_\omega$. This is done in a fashion similar to the Itô integral (e.g. Steele [37]); however the major differences here are that no filtration is involved, $Z_\omega$ is a random element of $L^2([0,1], \mathbb{C})$, and the functions $A_\omega$ are operator-valued.

Let

$$\mathcal{B} = \{A : [-\pi, \pi] \to \mathcal{S}_\infty(H) \text{ such that } \|A\|_{\mathcal{B}} < \infty\},$$

where

$$\|A\|_{\mathcal{B}}^2 = \int_{[-\pi,\pi]} |||A_\omega|||_\infty^2 |||\mathscr{F}_\omega|||_1 d\omega.$$

The space $\mathcal{B}$ is, in fact, the Bochner space $L^2([-\pi, \pi], \mathcal{S}_\infty(H))$, equipped with the measure $\mu(E) = \int_E |||\mathscr{F}_\omega|||_1 d\omega$; in particular, it is a Banach space (see e.g. Dinculeanu [14]). Let $M_0 \subset \mathcal{B}$ be the subspace of step functions, spanned by the elements $A\mathbf{1}_{[\alpha,\beta)}$, where either $A \in \mathcal{S}_2(H)$ or $A = I$, and $\alpha, \beta \in [-\pi, \pi]$, $\alpha < \beta$. We also define $M = \overline{M_0} \subset \mathcal{B}$, where the closure is taken in $\mathcal{B}$. We first define $\mathcal{I} : M_0 \to \mathbb{H}$ by linear extension of the mapping

$$\mathcal{I}(A\mathbf{1}_{[\alpha,\beta)}) = A(Z_\beta - Z_\alpha), \quad A \in \mathcal{S}_2(H) \text{ or } A = I.$$

In order to show that the image of $\mathcal{I}$ is indeed in $\mathbb{H}$, we need a Lemma:

**Lemma 3.1.** *Assume Conditions 1.1 hold, and let $T_1 = A_1 + \gamma_1 I$, $T_2 = A_2 + \gamma_2 I$, where $A_i \in \mathcal{S}_2(H)$, $\gamma_i \in \mathbb{C}$ for $i = 1, 2$. Then*

$$\langle T_1 Z_\alpha, T_2 Z_\beta \rangle_{\mathbb{H}} = \operatorname{trace}\left( T_1 \left[ \int_{-\pi}^{\min(\alpha,\beta)} \mathscr{F}_\omega d\omega \right] T_2^\dagger \right). \tag{22}$$

$$\|T_1 Z_\alpha\|_{\mathbb{H}}^2 \leq |||T_1|||_\infty^2 \int_{-\pi}^{\alpha} |||\mathscr{F}_\omega|||_1 d\omega. \tag{23}$$

*Hence if $\alpha_1 < \alpha_2 \leq \alpha_3 < \alpha_4$,*

$$\langle T_1(Z_{\alpha_2} - Z_{\alpha_1}), T_2(Z_{\alpha_4} - Z_{\alpha_3}) \rangle_{\mathbb{H}} = 0 \tag{24}$$

*and*

$$\left\| T_1(Z_{\alpha_2} - Z_{\alpha_1}) \right\|_{\mathbb{H}}^2 \leq |||T_1|||_\infty^2 \int_{\alpha_1}^{\alpha_2} |||\mathscr{F}_\omega|||_1 d\omega. \tag{25}$$

**Proof.** First notice that

$$\langle T_1 Z_\alpha, T_2 Z_\beta \rangle_{\mathbb{H}} = \langle A_1 Z_\alpha, A_2 Z_\beta \rangle_{\mathbb{H}} + \gamma_1 \langle Z_\alpha, A_2 Z_\beta \rangle_{\mathbb{H}} + \overline{\gamma_2} \langle A_1 Z_\alpha, Z_\beta \rangle_{\mathbb{H}} + \gamma_1 \overline{\gamma_2} \langle Z_\alpha, Z_\beta \rangle_{\mathbb{H}}.$$

By linearity, it suffices to show that the formula (22) holds for each of the terms on the right hand side. We will only show that if $\alpha \leq \beta$,

$$\langle A_1 Z_\alpha, A_2 Z_\beta \rangle_{\mathbb{H}} = \operatorname{trace}\left( A_1 \left[ \int_{-\pi}^{\alpha} \mathscr{F}_\omega d\omega \right] A_2^\dagger \right),$$

as the other equalities follow in a similar fashion. We have

$$\begin{aligned}
\langle A_1 Z_\alpha, A_2 Z_\beta\rangle_{\mathbb{H}} &= \mathbb{E}\int_0^1 T_1 Z_\alpha(\tau)\overline{T_2 Z_\beta(\tau)}d\tau \\
&= \mathbb{E}\int_0^1 \iint_{[0,1]^2} A_1(\tau,\sigma_1)\overline{A_2(\tau,\sigma_2)}Z_\alpha(\sigma_1)\overline{Z_\beta(\sigma_2)}d\sigma_1 d\sigma_2 d\tau \\
&= \int_0^1 \iint_{[0,1]^2} A_1(\tau,\sigma_1)\overline{A_2(\tau,\sigma_2)}\mathbb{E}\left[Z_\alpha(\sigma_1)\overline{Z_\beta}(\sigma_2)\right]d\sigma_1 d\sigma_2 d\tau \\
&= \int_0^1 A_1\left[\int_{-\pi}^{\alpha}\mathscr{F}_\omega d\omega\right]A_2^\dagger(\tau,\tau)d\tau \\
&= \operatorname{trace}\left(A_1\left[\int_{-\pi}^{\alpha}\mathscr{F}_\omega d\omega\right]A_2^\dagger\right),
\end{aligned}$$

where the last equality is justified by Brislawn [9, Proposition 3.3]. The permutation of integrals is justified by Fubini's Theorem, since

$$\begin{aligned}
&\int_0^1 \iint_{[0,1]^2} \mathbb{E}|Z_\alpha(\sigma_1)Z_\beta(\sigma_2)A_1(\tau,\sigma_1)A_2(\tau,\sigma_2)|d\sigma_1 d\sigma_2 d\tau \\
&\le \int_0^1 \iint_{[0,1]^2} \sqrt{\int_{-\pi}^{\alpha} f_\omega(\sigma_1,\sigma_1)d\omega \int_{-\pi}^{\beta} f_\omega(\sigma_2,\sigma_2)d\omega}\,|A_1(\tau,\sigma_1)A_2(\tau,\sigma_2)|d\sigma_1 d\sigma_2 d\tau \\
&\le \int_0^1 \left\{\iint_{[0,1]^2}\left[\int_{-\pi}^{\alpha} f_\omega(\sigma_1,\sigma_1)d\omega \int_{-\pi}^{\beta} f_\omega(\sigma_2,\sigma_2)d\omega\right]d\sigma_1 d\sigma_2 \right.\\
&\quad \times \left.\iint_{[0,1]^2}|A_1(\tau,\sigma_1)|^2|A_2(\tau,\sigma_2)|^2 d\sigma_1 d\sigma_2\right\}^{1/2} d\tau \\
&= \sqrt{\int_{-\pi}^{\alpha} |||\mathscr{F}_\omega|||_1 d\omega \int_{-\pi}^{\beta} |||\mathscr{F}_\omega|||_1 d\omega}\int_0^1\left(\int_0^1 |A_1(\tau,\sigma_1)|^2 d\sigma_1\right.\\
&\quad \times \left.\int_0^1 |A_2(\tau,\sigma_2)|^2 d\sigma_2\right)^{1/2} d\tau \\
&\le \sqrt{\int_{-\pi}^{\alpha} |||\mathscr{F}_\omega|||_1 d\omega \int_{-\pi}^{\beta} |||\mathscr{F}_\omega|||_1 d\omega}\,|||A_1|||_2|||A_2|||_2 < \infty,
\end{aligned}$$

by the Cauchy–Schwarz inequality. The proof of (23) follows from (22), Lemma 5.4 and Hölder's inequality. Statements (24) and (25) follow then directly from (22) and (23). □

We can now show that the image of $\mathcal{I}$ is indeed in $\mathbb{H}$. Since any $T \in M_0$ can be written $T = \sum_{j=1}^n T_j \mathbf{1}_{[\omega_j,\omega_{j+1})}$, where $T_j = A_j + \gamma_j I$ for some $A_j \in \mathcal{S}_2(H)$, $\gamma_j \in \mathbb{C}$ and $-\pi \le \omega_1 < \omega_2 < \cdots < \omega_{n+1} \le \pi$, Lemma 3.1 yields

$$\begin{aligned}
\|\mathcal{I}(T)\|_{\mathbb{H}}^2 &= \sum_{j,l=1}^n \left\langle T_j(Z_{\omega_{j+1}} - Z_{\omega_j}), T_l(Z_{\omega_{l+1}} - Z_{\omega_l})\right\rangle_{\mathbb{H}} = \sum_{j=1}^n \left\|T_j(Z_{\omega_{j+1}} - Z_{\omega_j})\right\|_{\mathbb{H}}^2 \\
&\le \sum_{j=1}^n |||T_j|||_\infty^2 \int_{\omega_j}^{\omega_{j+1}} |||\mathscr{F}_\omega|||_1 d\omega = \|T\|_{\mathcal{B}}^2.
\end{aligned}$$

Hence $\mathcal{I} : M_0 \to \mathbb{H}$ is continuous. We can therefore extend its domain to $M = \overline{M_0}$ (with the closure taken in $\mathcal{B}$) in the following way. Fix $T \in M$. For any sequence $\{T_n\}_{n\geq 1} \subset M_0$ converging to $T$, notice that $\{\mathcal{I}(T_n)\}_{n\geq 1}$ is also a Cauchy sequence in $\mathbb{H}$, by continuity of the operator $\mathcal{I}$. We then define

$$\mathcal{I}(T) = \lim_{n\to\infty} \mathcal{I}(T_n), \quad \text{in } \mathbb{H},$$

where the definition does not depend on the choice of the sequence $\{T_n\}_{n\geq 1}$ because $\mathcal{I}$ is linear and continuous. Moreover, $\mathcal{I} : M \to \mathbb{H}$ is linear, and

$$\|\mathcal{I}(T)\|_{\mathbb{H}} \leq \|T\|_{\mathcal{B}}$$

is valid for any $T \in M$. The precise characterization of the space $M$ on which the stochastic integral is defined is rather involved, and beyond the scope of this paper; however, $M$ does contain elements of the form

$$\omega \in [-\pi, \pi] \mapsto g(\omega)I + A_\omega, \tag{26}$$

where $g : [-\pi, \pi] \to \mathbb{C}$ is a cadlag function with a finite number of jumps and $A : [-\pi, \pi] \to \mathcal{S}_2(H)$ is cadlag (where continuity is meant with respect to $||| \cdot |||_\infty$) with a finite number of jumps, such that $\int_{-\pi}^{\pi} |||A_\omega|||_2^2 |||\mathscr{F}_\omega|||_1 d\omega < \infty$. This is all that will be required for our results. A noteworthy property of the stochastic integral is that $\mathcal{I}(T)$ can be seen as a Riemann–Stieltjes limit, for elements $T$ of the form (26). The next proposition gives the pointwise covariance of the stochastic integral.

**Proposition 3.2.** *Assume Conditions 1.1 hold, and let $T, S \in M$. Then*

$$\mathbb{E}\left[\int_{-\pi}^{\pi} T_\omega dZ_\omega(\tau) \cdot \overline{\int_{-\pi}^{\pi} S_\omega dZ_\omega(\sigma)}\right] = \int_{-\pi}^{\pi} T_\omega \mathscr{F}_\omega S_\omega^\dagger(\tau, \sigma) d\omega, \quad a.e.$$

**Proof.** The proof is similar to the proofs of Proposition 2.3 and Lemma 3.1 and so we only provide a sketch. Let $T_1 = A_1 + \gamma_1 I$, $T_2 = A_2 + \gamma_2 I$, where $A_i \in \mathcal{S}_2(H)$, $\gamma_i \in \mathbb{C}$ for $i = 1, 2$. We first show, similarly to Lemma 3.1, that

$$\mathbb{E}\left\langle T_1 Z_\alpha \otimes T_2 Z_\beta, u \otimes v\right\rangle_{\mathcal{S}_2} = \left\langle T_1 \int_{-\pi}^{\min(\alpha,\beta)} \mathscr{F}_\omega d\omega T_2^\dagger, u \otimes v\right\rangle_{\mathcal{S}_2},$$

for all $u, v \in H$. Using this, we then extend to $M_0$ by linearity:

$$\mathbb{E}\left\langle \mathcal{I}(T) \otimes \mathcal{I}(S), u \otimes v\right\rangle_{\mathcal{S}_2} = \left\langle \int_{-\pi}^{\pi} T_\omega \mathscr{F}_\omega S_\omega^\dagger d\omega, u \otimes v\right\rangle_{\mathcal{S}_2}, \quad T, S \in M_0.$$

Hence

$$\mathbb{E}\left[\mathcal{I}(T)(\tau)\mathcal{I}(S)(\sigma)\right] = \int_{-\pi}^{\pi} T_\omega \mathscr{F}_\omega S_\omega^\dagger(\tau, \sigma) d\omega, \quad \text{a.e., } T, S \in M_0. \tag{27}$$

The rest is similar to Proposition 2.3: let $B' = L^1([0, 1]^2, \mathbb{C})$, with norm

$$\|g\|_{B'} = \iint_{[0,1]^2} |g(\tau, \sigma)| d\tau d\sigma, \quad g \in B'.$$

Notice that for $T, S \in M$, the Cauchy–Schwarz inequality gives

(i) $\mathbb{E}\,\|\mathcal{I}(T)\otimes\mathcal{I}(S)\|_{B'} \le \|T\|_{\mathcal{B}}\,\|S\|_{\mathcal{B}}$,

(ii) $\left\|\int_{-\pi}^{\pi} T_\omega \mathscr{F}_\omega S_\omega^\dagger d\omega\right\|_{B'} \le \|T\|_{\mathcal{B}}\,\|S\|_{\mathcal{B}}$.

The proof is then completed by writing

$$\left\|\mathbb{E}\left[\mathcal{I}(T)\otimes\mathcal{I}(S)\right] - \int_{-\pi}^{\pi} T_\omega \mathscr{F}_\omega S_\omega^\dagger d\omega\right\|_{B'} \le \left\|\mathbb{E}\left[\mathcal{I}(T)\otimes\mathcal{I}(S)\right] - \mathbb{E}\left[\mathcal{I}(T_n)\otimes\mathcal{I}(S_n)\right]\right\|_{B'} + \left\|\mathbb{E}\left[\mathcal{I}(T_n)\otimes\mathcal{I}(S_n)\right] - \int_{-\pi}^{\pi} T_\omega \mathscr{F}_\omega S_\omega^\dagger d\omega\right\|_{B'},$$

where $\{T_n\}_{n\ge1}$ and $\{S_n\}_{n\ge1}$ are sequences in $M_0$ converging to $T$ and $S$, respectively. The terms on the right hand side can be made arbitrarily small using (i), (ii) and (27). □

As a direct Corollary, we obtain a formula for the scalar product between two stochastic integrals:

**Corollary 3.3.** *Assume Conditions 1.1 hold, and let $T, S \in M$. Then*

$$\left\langle \int_{-\pi}^{\pi} T_\omega dZ_\omega, \int_{-\pi}^{\pi} S_\omega dZ_\omega \right\rangle_{\mathbb{H}} = \int_0^1 \int_{-\pi}^{\pi} T_\omega \mathscr{F}_\omega S_\omega^\dagger(\tau,\tau)\, d\omega\, d\tau.$$

### *3.2. Optimal dimension reduction*

It follows from the discussion in the previous section that the stochastic integral (21) defined by a truncation of the Cramér–Karhunen–Loève representation is well-defined. The purpose of this section is to prove that it provides a version of $X_t$ with only $K$ degrees of freedom, which optimally approximates $X_t$ in mean square among all other linear transformations of $X_t$ with $K$ degrees of freedom. First we make sense of linear transformations of $X_t$, or as they are known in time series, filtered versions of $X_t$.

Given the stationary time series $\{X_t\}_{t\in\mathbb{Z}}$ in $L^2([0,1],\mathbb{R})$ and a sequence $\{A_s\}_{s\in\mathbb{Z}}$ of Hilbert–Schmidt operators on $L^2([0,1],\mathbb{R})$, we can construct a new time series

$$Y_t = \sum_{s\in\mathbb{Z}} A_{t-s} X_s,$$

where $Y_t$ is a random element of $L^2([0,1],\mathbb{R})$, which is said to be obtained by *linear filtering* of $X_t$. Notice that the rank $K$ approximation of $X_t$ based on the PCA of $\mathscr{R}_0$ can also be expressed as a filtered version of $X_t$.

The following proposition formalizes the construction of filtered series $Y_t$, and gives their Cramér representation:

**Proposition 3.4.** *Assume Conditions 1.1 hold, and let $\{A_s\}_{s\in\mathbb{Z}} \subset \mathcal{S}_2(H)$ such that $\sum_{s\in\mathbb{Z}} |||A_s|||_2 < \infty$. Then $Y_t = \sum_{s\in\mathbb{Z}} A_s X_{t-s}$ converges in $\mathbb{H}$, and for each $t \in \mathbb{Z}$,*

$$Y_t = \int_{-\pi}^{\pi} e^{\mathrm{i}t\omega} \widetilde{A}_\omega dZ_\omega, \quad a.s.\ a.e.,$$

*where $\widetilde{A}_\omega = \sum_s e^{-\mathrm{i}s\omega} A_s$.*

**Proof.** First notice that

$$A_s X_{t-s} = \int_{-\pi}^{\pi} e^{\mathrm{i}(t-s)\omega} A_s dZ_\omega = \mathcal{I}(B_s),$$

where $B_s$ denotes the mapping $\omega \mapsto e^{\mathrm{i}(t-s)\omega} A_s$. Now $B_s \in \mathcal{B}$ since

$$\|B_s\|_{\mathcal{B}} \le \sqrt{\int_{-\pi}^{\pi} \|\mathscr{F}_\omega\|_1 d\omega} \|A_s\|_2.$$

Hence,

$$\sum_s \|B_s\|_{\mathcal{B}} \le \sqrt{\int_{-\pi}^{\pi} \|\mathscr{F}_\omega\|_1 d\omega} \sum_s \|A_s\|_2 < \infty,$$

and the partial sums $B^{(N)} := \sum_{|s|\le N} B_s$ converge in $\mathcal{B}$ to

$$B = \sum_s B_s = e^{\mathrm{i}t\omega} \sum_s e^{-\mathrm{i}\omega s} A_s = e^{\mathrm{i}t\omega} \widetilde{A}_\omega,$$

where $\widetilde{A}_\omega = \sum_s e^{-\mathrm{i}\omega s} A_s$. Hence, by continuity, we obtain

$$\sum_s A_s X_{t-s} = \sum_s \mathcal{I}(B_s) = \mathcal{I}(B) = \int_{-\pi}^{\pi} e^{\mathrm{i}t\omega} \widetilde{A}_\omega dZ_\omega. \quad \square$$

Writing $Z^X$ for the orthogonal increment process associated with $X_t$ (and similarly $Z^Y$ for that of $Y_t$), we see that the Cramér representation of $Y$ is related to the one of $X$ through the formal relation $dZ^Y_\omega = \widetilde{A}_\omega dZ^X_\omega$. The spectral density operator of $Y_t$ is given by the following Proposition:

**Proposition 3.5.** *Under the assumptions of Proposition 3.4, the spectral density operator of $Y$ is given by*

$$\mathscr{F}^Y_\omega = \widetilde{A}_\omega \mathscr{F}^X_\omega \widetilde{A}^\dagger_\omega,$$

*where $\mathscr{F}^X$ denotes the spectral density operator of $X_t$.*

**Proof.** We first show that the spectral density operator of $Y$ is well defined. Using the same techniques as in the proof of Proposition 7.15 of Panaretos and Tavakoli [32], we obtain

$$\|\mathscr{R}^Y_t\|_2 \le \sum_{s,u\in\mathbb{Z}} \|A_s\|_2 \|A_u\|_2 \|\mathscr{R}^X_{t+u-s}\|_2,$$

where $\mathscr{R}^X_t$ denotes the autocovariance operator of $X$ at lag $t$ (and similarly for $Y$). Hence

$$\sum_t \|\mathscr{R}^Y_t\|_2 \le \sum_{s,u\in\mathbb{Z}} \|A_s\|_2 \|A_u\|_2 \sum_t \|\mathscr{R}^X_{t+u-s}\|_2 \le \kappa \sum_{s,u\in\mathbb{Z}} \|A_s\|_2 \|A_u\|_2 < \infty$$

where $\kappa = \sum_t \|\mathscr{R}^X_t\|_2$. The spectral density operator of $Y$ is therefore well defined in $\mathcal{S}_2(H)$. Using Proposition 3.2, we obtain

$$\mathscr{R}^Y_t = \int_{-\pi}^{\pi} e^{\mathrm{i}\omega t} \widetilde{A}_\omega \mathscr{F}^X_\omega \widetilde{A}^\dagger_\omega d\omega, \quad \text{in } H,$$

which corresponds to the inversion formula for the spectral density operator. Hence $\mathscr{F}^Y_\omega = \widetilde{A}_\omega \mathscr{F}^X_\omega \widetilde{A}^\dagger_\omega$. □

Consider now the problem of reducing the functional time series $X_t$ to a finite dimensional vector series (say of dimension $q$), by filtering $X_t$:

$$Y_t = \sum_s A_s X_{t-s} \in \mathbb{R}^q, \quad A_s \in \mathcal{S}_2(H, \mathbb{R}^q), \tag{28}$$

where $\mathcal{S}_2(H, \mathbb{R}^q)$ denotes the space of Hilbert–Schmidt operators from $H$ to $\mathbb{R}^q$. Though the series $Y_t$ is no longer interpretable in a functional sense, it may be filtered anew to yield a functional process

$$X^*_t = \sum_s B_s Y_{t-s}, \quad B_s \in \mathcal{S}_2(\mathbb{R}^q, H), \tag{29}$$

which is interpretable in a functional sense, and is in fact a rank $q$ approximation of $X_t$. The corresponding Cramér representation of $X^*_t$ will be:

**Lemma 3.6.** *Let $\{A_s\}_{s\in\mathbb{Z}} \subset \mathcal{S}_2(H, \mathbb{R}^q)$ and $\{B_s\}_{s\in\mathbb{Z}} \subset \mathcal{S}_2(\mathbb{R}^q, H)$ such that*

$$\sum_{s\in\mathbb{Z}} (|||A_s|||_2 + |||B_s|||_2) < \infty.$$

*Then, the Cramér representation of $X^*_t$, defined in (29), is*

$$X^*_t = \int_{-\pi}^{\pi} e^{\mathrm{i}\omega t} \widetilde{B}_\omega \widetilde{A}_\omega dZ^X_\omega, \quad \text{a.s. a.e.},$$

*where $\widetilde{A}_\omega = \sum_{s\in\mathbb{Z}} e^{-\mathrm{i}\omega s} A_s$ and $\widetilde{B}_\omega = \sum_{s\in\mathbb{Z}} e^{-\mathrm{i}\omega s} B_s$.*

**Proof.** The proof is similar to the proof of Proposition 3.4 and is thus omitted. □

Therefore, the Cramér representation of $X^*_t$ is given by

$$X^*_t = \int_{-\pi}^{\pi} e^{\mathrm{i}\omega t} \widetilde{C}_\omega dZ^X_\omega,$$

where $\widetilde{C}_\omega = \widetilde{B}_\omega \widetilde{A}_\omega$, and is hence of rank at most $q$. We are now in a position to show that the truncated Cramér–Karhunen–Loève expansion (21) provides a *Harmonic Principal Component Analysis* of $X_t$: under the mean square error approximation criterion

$$\mathbb{E} \left\| X_t - X^*_t \right\|^2,$$

which is independent of $t$ by stationarity, the optimal choice of $\widetilde{C}_\omega$ is given by $\sum_{n=1}^{q(\omega)} \varphi^\omega_n \otimes \varphi^\omega_n$, where we recall that $\{\varphi^\omega_n\}$ are the eigenfunctions of $\mathscr{F}_\omega$.

**Theorem 3.7** (*Harmonic Principal Component Analysis*)**.** *Assume Conditions 1.1 hold. Let $X_t = \int_{-\pi}^{\pi} e^{\mathrm{i}\omega t} dZ_\omega$ be a stationary time series in $L^2([0,1],\mathbb{R})$, and let $X^*_t = \int_{-\pi}^{\pi} e^{\mathrm{i}\omega t} \widetilde{C}_\omega dZ_\omega$, with $\widetilde{C} \in M$. Let $q : [-\pi, \pi] \to \mathbb{N}$ be a cadlag function. Then, the solution to*

$$\min_{\widetilde{C}\in M} \mathbb{E} \left\| X_t - X^*_t \right\|^2$$
$$\text{subject to rank}\left(\widetilde{C}_\omega\right) \le q(\omega),$$

*is given by*

$$\widetilde{C}_{\omega} = \sum_{j=1}^{q(\omega)} \varphi_j^{\omega} \otimes \varphi_j^{\omega},$$

*where* $\mathscr{F}_{\omega} = \sum_{j=1}^{\infty} \mu_j(\omega)\varphi_j^{\omega} \otimes \varphi_j^{\omega}$ *is the spectral decomposition of* $\mathscr{F}_{\omega}$. *The approximation error is given by*

$$\mathbb{E}\left\|X_t - X_t^*\right\|^2 = \int_{-\pi}^{\pi}\left\{\sum_{j>q(\omega)} \mu_j(\omega)\right\} d\omega.$$

**Proof.** The proof is an adaptation of Brillinger [8, Theorem 9.3.1] to our case. Since $X_t - X_t^* = \int_{-\pi}^{\pi} e^{\mathrm{i}\omega t}(I - \widetilde{C}_{\omega})dZ_{\omega}$, Corollary 3.3 yields

$$\begin{aligned}\mathbb{E}\left\|X_t - X_t^*\right\|^2 &= \int_0^1 \int_{-\pi}^{\pi} (I - \widetilde{C}_{\omega})\mathscr{F}_{\omega}(I - \widetilde{C}_{\omega})^{\dagger}(\tau,\tau)d\omega d\tau \\ &= \int_{-\pi}^{\pi}\int_0^1 (I - \widetilde{C}_{\omega})\mathscr{F}_{\omega}^{1/2}\left[(I - \widetilde{C}_{\omega})\mathscr{F}_{\omega}^{1/2}\right]^{\dagger}(\tau,\tau)d\tau d\omega \\ &= \int_{-\pi}^{\pi} |||(I - \widetilde{C}_{\omega})\mathscr{F}_{\omega}^{1/2}|||_2^2 d\omega, \qquad (30)\end{aligned}$$

where $\mathscr{F}_{\omega}^{1/2} = \sum_{j=1}^{\infty}\sqrt{\mu_j(\omega)}\varphi_j^{\omega} \otimes \varphi_j^{\omega}$ and the permutation of integrals is justified by Fubini's Theorem. The term (30) is minimized by minimizing $|||(I - \widetilde{C}_{\omega})\mathscr{F}_{\omega}^{1/2}|||_2$ for each $\omega$. This is achieved, under our constraints, by $\widetilde{C}_{\omega} = \sum_{j=1}^{q(\omega)} \varphi_j^{\omega} \otimes \varphi_j^{\omega}$. The expression for the error term follows directly. □

**Remark 3.8.** Contrary to the classical finite-dimensional results (e.g. Brillinger [8]), we do not restrict $q(\omega)$ to be constant over $\omega$. In this sense, when restricted to finite dimensional Hilbert spaces, our results are more general than analogous results for vector-valued time series.

**Remark 3.9.** Restricting $q(\omega) = q \in \mathbb{N}$ yields a rank $q$ version $X_t^*$ of $X_t$. This can be represented in a 1–1 fashion by the filtered vector series $Y_t$ (an element of $\mathbb{R}^q$) in (28), whose important characteristic is the lack of correlation between its coordinates—just as one expects with the scores obtained in a traditional principal component analysis. The $Y_t$ can therefore serve as the harmonic principal component scores.

**Remark 3.10** (*Cramér–Karhunen–Loève Decomposition*)**.** The theorem makes precise the way in which a Cramér–Karhunen–Loève representation of the form (20) holds, in that we may now explicitly and rigorously state

$$\mathbb{E}\left\|X_t - \int_{-\pi}^{\pi} e^{\mathrm{i}\omega t}\left(\sum_{n=1}^{K}\varphi_n^{\omega}\otimes\varphi_n^{\omega}\right)dZ_{\omega}\right\|^2 = \int_{-\pi}^{\pi}\left\{\sum_{j>K}\mu_j(\omega)\right\}d\omega \qquad (31)$$

as an immediate corollary, under the same conditions as in Theorem 3.7.

We now provide the expressions of the filters $A$, $B$ involved in (28) and (29). Write $q(\omega) = \sum_{l=1}^{L} \mathbf{1}_{[\omega_l,\omega_{l+1})} q_l$, where $-\pi = \omega_1 < \cdots < \omega_{L+1} = \pi$ and the $q_l$s are non-negative integers.

Note that allowing $q_l = 0$ for some $l$ corresponds to filtering out the frequencies in the range $[\omega_l, \omega_{l+1})$. If we let $q = \max_\omega q(\omega)$, and $v_1, \ldots, v_q$ be an orthonormal basis of $\mathbb{R}^q$, then

$$\widetilde{A}_\omega = \sum_{j=1}^{q(\omega)} v_j \otimes \varphi_j^\omega,$$

and $\widetilde{B}_\omega = \widetilde{A}_\omega^\dagger$. The corresponding filters are given by

$$\begin{aligned} A_s &= (2\pi)^{-1} \int_{-\pi}^{\pi} e^{\mathrm{i}\alpha s} \widetilde{A}_\alpha d\alpha = (2\pi)^{-1} \int_{-\pi}^{\pi} e^{\mathrm{i}\alpha s} \left[ \sum_{j=1}^{q(\alpha)} v_j \otimes \varphi_j^\alpha \right] d\alpha \\ &= (2\pi)^{-1} \sum_{l=1}^{L} \int_{\omega_l}^{\omega_{l+1}} e^{\mathrm{i}\alpha s} \left[ \sum_{j=1}^{q_l} v_j \otimes \varphi_j^\alpha \right] d\alpha \\ &= (2\pi)^{-1} \sum_{l=1}^{L} \sum_{j=1}^{q_l} v_j \otimes \left[ \int_{\omega_l}^{\omega_{l+1}} e^{-\mathrm{i}\alpha s} \varphi_j^\alpha \right], \end{aligned}$$

and $B_s = A_{-s}^\dagger$.

## 4. Asymptotics for the empirical harmonic PCA

It is clear from the results presented that, to carry out a harmonic principal component analysis in practice, when the spectral density operator may be unknown, one needs estimators of the eigenstructure of $\mathscr{F}_\omega$. The purpose of this section is to provide such empirical versions of the eigenstructure of $\mathscr{F}_\omega$, and to describe their large sample properties. Our approach will be to estimate the eigenstructure via the plug in method: given a finite stretch of the stationary time series $\{X_t\}_{t=0}^T$, we will use an estimator $\mathscr{F}_\omega^{(T)}$ of $\mathscr{F}_\omega$ proposed in Panaretos and Tavakoli [32], and use its eigenstructure as an estimator of the eigenstructure of $\mathscr{F}_\omega$. In brief, Panaretos and Tavakoli [32] define their estimator $\mathscr{F}_\omega^{(T)}$ as follows. Let $W(x)$ be a real function defined on $\mathbb{R}$ such that: (1) $W$ is positive, even, and bounded in variation; (2) $W(x) = 0$ if $|x| \geq 1$; (3) $\int_{-\infty}^{\infty} W(x)dx = 1$; (4) $\int_{-\infty}^{\infty} W(x)^2 dx < \infty$. For a bandwidth $B_T > 0$, write

$$W^{(T)}(x) = \sum_{j\in\mathbb{Z}} \frac{1}{B_T} W\left(\frac{x + 2\pi j}{B_T}\right). \tag{32}$$

Define the *estimator of the spectral density kernel* by

$$f_\omega^{(T)}(\tau, \sigma) = \frac{2\pi}{T} \sum_{s=1}^{T-1} W^{(T)}\left(\omega - \frac{2\pi s}{T}\right) \widetilde{X}_{\frac{2\pi s}{T}}^{(T)}(\tau) \widetilde{X}_{-\frac{2\pi s}{T}}^{(T)}(\sigma),$$

where

$$\widetilde{X}_\omega^{(T)}(\tau) = (2\pi T)^{-1/2} \sum_{t=0}^{T-1} X_t(\tau) \exp(-\mathrm{i}\omega t)$$

is the discrete Fourier transform of the segment $\{X_t\}_{t=1}^T$. Denote by $\mathscr{F}_\omega^{(T)}$ the operator with kernel $f_\omega^{(T)}$.

To quantify the strength of dependence among the observations $\{X_t\}$ we will use the *cumulant kernels* of the series, as introduced in Panaretos and Tavakoli [32]; the pointwise definition of a $k$-th *order cumulant kernel* is,

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

where the sum extends over all unordered partitions of $\{1, \dots, k\}$. Especially the cumulant kernel of order 4 gives rise to a corresponding 4-th *order cumulant operator* $\mathscr{R}_{t_1,t_2,t_3} : L^2([0,1]^2, \mathbb{R}) \to L^2([0,1]^2, \mathbb{R})$, defined by

$$\mathscr{R}_{t_1,t_2,t_3}(u \otimes v) = \operatorname{cum}\left(X_{t_1} \otimes X_{t_2} \left\langle u, X_{t_3}\right\rangle \langle v, X_0\rangle\right), \quad u, v \in L^2([0,1], \mathbb{R}).$$

Weak dependence will be quantified in terms of the following mixing conditions, which are summability conditions on the cumulant kernels extending the classical mixing conditions of Brillinger [8]. Throughout this Section, we will assume the following conditions hold:

**Conditions 4.1.** *$X_t$ is a stationary times series in $L^2([0,1], \mathbb{R})$, satisfying:*

(0) $\mathbb{E}\,\|X_0\|^k < \infty$ *for all* $k \geq 1$
(i) $\sum_{t_1,\dots,t_{k-1}=-\infty}^{\infty} \left\|\operatorname{cum}\left(X_{t_1}, \dots, X_{t_{k-1}}, X_0\right)\right\|_2 < \infty$, *for all* $k \geq 2$.
(i$'$) $\sum_{t_1,\dots,t_{k-1}=-\infty}^{\infty} (1+|t_j|) \left\|\operatorname{cum}\left(X_{t_1}, \dots, X_{t_{k-1}}, X_0\right)\right\|_2 < \infty$, *for* $k \in \{2,4\}$ *and* $j < k$.
(ii) $\sum_{t\in\mathbb{Z}} (1+|t|) |||\mathscr{R}_t|||_1 < \infty$.
(iii) $\sum_{t_1,t_2,t_3\in\mathbb{Z}} |||\mathscr{R}_{t_1,t_2,t_3}|||_1 < \infty$.
(iv) $(\tau,\sigma) \mapsto r_t(\tau,\sigma)$ *is continuous, and* $\sum_{t\in\mathbb{Z}} \|r_t\|_\infty < \infty$.

We note that these conditions are not the weakest possible for the results that we will state—however, they considerably simplify the technical aspects of the proofs. For a detailed discussion of the interpretation and role of these conditions, and a comparative discussion in relation with the finite-dimensional versions thereof (as given in Brillinger [8]), the reader is referred to Panaretos and Tavakoli [32].

The following Theorem gives the asymptotic distribution of the spectral density estimators, and follows directly from results in Panaretos and Tavakoli [32]:

**Theorem 4.2.** *Assume that Conditions 4.1 hold. If $B_T \to 0$ such that $TB_T \to \infty$ and $TB_T^3 \to 0$. Then, for any distinct frequencies $\omega_1, \dots, \omega_J \in [0,\pi]$, with $J < \infty$,*

$$\sqrt{TB_T}\left(\mathscr{F}_{\omega_j}^{(T)} - \mathscr{F}_{\omega_j}\right) \xrightarrow{d} \breve{\mathscr{F}}_{\omega_j}, \quad j = 1, \dots, J,$$

*where $\breve{\mathscr{F}}_{\omega_j}$, $j = 1, \dots, J$ are independent mean zero complex Gaussian elements in $L^2([0,1]^2, \mathbb{C})$, with covariance kernel*

$$\operatorname{cov}(\breve{\mathscr{F}}_\omega, \breve{\mathscr{F}}_\omega) = 2\pi \int_{\mathbb{R}} W(\alpha)^2 d\alpha \cdot \mathscr{F}_\omega \widetilde{\otimes} \mathscr{F}_\omega, \quad \omega \in (0,\pi),$$

*and*

$$\operatorname{cov}(\breve{\mathscr{F}}_\omega, \breve{\mathscr{F}}_\omega) = 2\pi \int_{\mathbb{R}} W(\alpha)^2 d\alpha \cdot \left[\mathscr{F}_\omega \widetilde{\otimes} \mathscr{F}_\omega + \mathscr{F}_\omega \widetilde{\otimes}_{\mathsf{T}} \mathscr{F}_\omega\right], \quad \omega = 0, \pi.$$

*In particular, $\breve{\mathscr{F}}_\omega$ is real Gaussian if $\omega = 0, \pi$.*

Before stating our results concerning the asymptotic distribution of the estimators of the eigenvalues/eigenfunctions, we need to introduce some necessary notation. For any $\omega \in [0, \pi]$, let

$$\mathscr{F}_\omega^{(T)} = \sum_{i=1}^{\infty} \mu_{i,T}(\omega) \varphi_{i,T}^\omega \otimes \varphi_{i,T}^\omega$$

be the spectral decomposition of $\mathscr{F}_\omega^{(T)}$, and recall that

$$\mathscr{F}_\omega = \sum_{i=1}^{\infty} \mu_i(\omega) \varphi_i^\omega \otimes \varphi_i^\omega$$

is the spectral decomposition of $\mathscr{F}_\omega$. For any fixed $\omega$, $\{\mu_{i,T}(\omega)\}_{i\ge1}$ and $\{\mu_i(\omega)\}_{i\ge1}$ are non-increasing positive sequences tending to zero. We denote by $\{\lambda_i(\omega)\}_{i\ge1}$ the decreasing sequence of distinct elements of $\{\mu_i(\omega)\}_{i\ge1}$, define the set $I_k(\omega) = \{i \ge 1 : \mu_i(\omega) = \lambda_k(\omega)\}$ and we denote its cardinality by $m_k(\omega) = |I_k(\omega)|$. We can now define

$$\Pi_k(\omega) = \sum_{i \in I_k(\omega)} \varphi_i^\omega \otimes \varphi_i^\omega,$$

which is the projection onto the $k$th eigenspace of $\mathscr{F}_\omega$. This way,

$$\mathscr{F}_\omega = \sum_{i=1}^{\infty} \lambda_i(\omega) \Pi_i(\omega).$$

The estimator of $\Pi_k(\omega)$ is defined by

$$\Pi_{k,T}(\omega) = \sum_{i \in I_k(\omega)} \varphi_{i,T}^\omega \otimes \varphi_{i,T}^\omega.$$

We also define $S_k(\omega) = \sum_{j \ne k} (\lambda_k(\omega) - \lambda_j(\omega))^{-1} \Pi_j(\omega)$, where the sum is over all $j \ne k$ such that $\lambda_j(\omega) \ne 0$. We define the operator

$$\eta_k^\omega = S_k(\omega) \widetilde{\otimes} \Pi_k(\omega) + \Pi_k(\omega) \widetilde{\otimes} S_k(\omega) \in \mathcal{S}_\infty(\mathcal{S}_2(H)),$$

and the bounded operator $p_k^\omega : \mathcal{S}_2(H) \to \mathbb{C}$ by $p_k(A) = \langle A, \Pi_k(\omega) \rangle_{\mathcal{S}_2}$.

The following Theorem gives the asymptotic distribution of the estimators of the eigenvalues and eigenvectors of the spectral density matrix.

**Theorem 4.3.** *Let $\omega_1, \ldots, \omega_K \in [0, \pi]$ be distinct, and $L \subset \mathbb{N}^*$ be a set of cardinality $|L| < \infty$. Provided Conditions 4.1 hold, and $B_T \to 0$ such that $T B_T \to \infty$ and $T B_T^3 \to 0$, then*

$$\sqrt{T B_T} \{\Pi_{j,T}(\omega_i) - \Pi_j(\omega_i) : j \in L\} \xrightarrow{d} \eta_L^{\omega_i}(\breve{\mathscr{F}}_{\omega_i}), \quad i = 1, \ldots, K$$

*and*

$$\sqrt{T B_T} \left\{ \sum_{s \in I_j(\omega_i)} [\mu_{s,T}(\omega_i) - \lambda_j(\omega_i)] : j \in L \right\} \xrightarrow{d} p_L^{\omega_i}(\breve{\mathscr{F}}_{\omega_i}), \quad i = 1, \ldots, K.$$

*The limiting random elements $\{\eta_L^{\omega_i}(\breve{\mathscr{F}}_{\omega_i})\}_{i=1,\ldots,K}$ and $\{p_L^{\omega_i}(\breve{\mathscr{F}}_{\omega_i})\}_{i=1,\ldots,K}$ are all independent complex Gaussian random elements. Their covariances are given by the following formulae (in which we have written $\lambda_k$ instead of $\lambda_k(\omega)$ for clarity, and similarly for $\Pi_k, \varphi_k$)*

$$\mathbb{E}\left[ \eta_k^\omega(\breve{\mathscr{F}}_\omega) \otimes \eta_l^\omega(\breve{\mathscr{F}}_\omega) \right]$$

$$= \begin{cases} -\kappa\lambda_k\lambda_l(\lambda_k-\lambda_l)^{-2}\left[\Pi_k\widetilde{\otimes}\Pi_l + \Pi_l\widetilde{\otimes}\Pi_k + A^{\omega}_{kl} + (A^{\omega}_{kl})^{\dagger}\right] & \text{if } k\neq l \\ \kappa\sum_{s\neq k}\lambda_k\lambda_s(\lambda_k-\lambda_s)^{-2}\left[\Pi_k\widetilde{\otimes}\Pi_s + \Pi_s\widetilde{\otimes}\Pi_k + A^{\omega}_{ks} + (A^{\omega}_{ks})^{\dagger}\right] & \text{if } k=l, \end{cases}$$

*where* $A^{\omega}_{ks} = \mathbf{1}_{\{0,\pi\}}(\omega)(\varphi^{\omega}_k\otimes\varphi^{\omega}_s)\widetilde{\otimes}(\varphi^{\omega}_s\otimes\varphi^{\omega}_k)$, *and*

$$\operatorname{cov}(p^{\omega}_l(\breve{\mathscr{F}}_{\omega}), p^{\omega}_k(\breve{\mathscr{F}}_{\omega})) = \left(1+\mathbf{1}_{\{0,\pi\}}(\omega)\right)\kappa\delta_{lk}\lambda_l^2 m_l,$$

*with* $\kappa = 2\pi\int_{\mathbb{R}} W^2(x)dx$.

**Proof.** The proof rests on the adaptation of Theorem 1.3 of Mas and Menneteau [30] to our case, and we therefore give only a sketch. For $l\geq 1$, we denote by $\mathcal{S}^l$ the $l$-fold product space $\mathcal{S}_2(H)\times\cdots\times\mathcal{S}_2(H)$, equipped with the norm

$$\|(A_1,\ldots,A_l)\|_{\mathcal{S}^l} = \max_{j=1,\ldots,l} |||A_j|||_2,$$

and equip $\mathbb{C}^l$ with the norm

$$|(\alpha_1,\ldots,\alpha_l)|_{\infty} = \max_{j=1,\ldots,l}|\alpha_j|.$$

We endow the space $\mathcal{S}^l\times\mathbb{C}^l$ with the norm

$$\|(A_1,\ldots,A_l,\alpha_1,\ldots,\alpha_l)\|_{*} = \max_{j=1,\ldots,l}\{|||A_j|||_2, |\alpha_j|\}.$$

Defining the bounded linear operator $J:\mathcal{S}^K\to\mathcal{S}^{Kl}\times\mathbb{C}^{Kl}$ by

$$J(A_1,\ldots,A_K) = (\eta_L(A_1),\ldots,\eta_L(A_K), p_L(A_1),\ldots,p_L(A_K)),$$

we show that

$$\sqrt{TB_T}\left[\left\{\Pi_{j,T}(\omega_i)-\Pi_j(\omega_i)\right\}_{j\in L; i=1,\ldots,K}, \left\{\sum_{s\in I_j(\omega_i)}[\mu_{s,T}(\omega_i)-\lambda_j(\omega_i)]\right\}_{j\in L; i=1,\ldots,K}\right]$$
$$= J(\sqrt{TB_T}\{\mathscr{F}^{(T)}_{\omega_i}-\mathscr{F}_{\omega_i}\}_{i=1,\ldots,K}) + \sqrt{TB_T}\mathcal{R},$$

where $\mathcal{R} = \left(\{R_{L,T}(\omega_i)\}_{i=1,\ldots,K}, \{r_{L,T}(\omega_i)\}_{i=1,\ldots,K}\right)$, where $R_{L,T}(\omega_i), r_{L,T}(\omega_i)$ are given by Mas and Menneteau [30, Proposition 2.3]. The proof is completed by showing that $\sqrt{TB_T}\mathcal{R}\xrightarrow{p} 0$ and applying the continuous mapping Theorem. The determination of the covariance structure of the limiting random elements is given separately in Section 4.1. □

Notice that the estimators of the eigenspaces are not asymptotically independent, which is expected since they are constrained to be mutually orthogonal.

### 4.1. Computation of covariances

In this section, we determine the asymptotic covariances of estimators of the eigenprojections and eigenvalues of $\mathscr{F}_{\omega}$, as stipulated in Theorem 4.3. Notice that the covariance operator of $\breve{\mathscr{F}}_{\omega}$ is given by either

$$\mathcal{C} = \kappa\cdot C\widetilde{\otimes}C, \quad \omega\in(0,\pi), \tag{33}$$

or

$$\mathcal{C} = \kappa \cdot [C \widetilde{\otimes} C + C \widetilde{\otimes}_{\mathsf{T}} C], \quad \omega \in \{0, \pi\} \tag{34}$$

where $\kappa = 2\pi \int_{\mathbb{R}} W(\omega)^2 d\omega$, $C$ is a nuclear operator on $L^2([0,1],\mathbb{C})$ in the first case, and on $L^2([0,1],\mathbb{R})$ in the second case. We therefore restrict ourselves to the computation of the covariances between elements $\eta_k(Y)$ and $p_k(Y)$, where $Y$ is either a self-adjoint random element in $\mathcal{S}_2(L^2([0,1],\mathbb{C}))$ with covariance operator given by (33), or $Y$ is a self-adjoint random element in $\mathcal{S}_2(L^2([0,1],\mathbb{R}))$ with covariance operator given by (34). To simplify the presentation, we first state and prove some useful lemmas. In this context we let $C = \sum_i \mu_i \varphi_i \otimes \varphi_i$ be the singular value decomposition of $C$.

**Lemma 4.4.** *Let $H$ be a complex Hilbert space, and $C$ be a nuclear and self-adjoint operator on $H$, with spectral decomposition $C = \sum_i \mu_i \varphi_{ii}$, where $\varphi_{ij} = \varphi_i \otimes \varphi_j$. If $Y$ is a complex Gaussian random element on $\mathcal{S}_2(H)$ that takes self-adjoint values, with mean 0 and covariance operator*

$$\mathcal{C} = C \widetilde{\otimes} C, \tag{35}$$

*then*

$$Y = \sum_i \xi_i \varphi_{ii} + \sum_{i<j} \xi_{ij} e_{ij} + \mathring{\imath} \zeta_{ij} \tilde{e}_{ij}, \tag{36}$$

*where the convergence holds in expected mean square (in $\mathcal{S}_2(H)$), $e_{ij} = 2^{-1/2}(\varphi_{ij} + \varphi_{ji})$, $\tilde{e}_{ij} = 2^{-1/2}(\varphi_{ij} - \varphi_{ji})$ and $\{\xi_i\}_i \cup \{\xi_{ij}\}_{i<j} \cup \{\zeta_{ij}\}_{i<j}$ are independent real Gaussian random variables, defined by*

$$\xi_i = \langle Y, \varphi_{ii} \rangle_{\mathcal{S}_2}, \qquad \xi_{ij} = \left\langle Y, e_{ij} \right\rangle_{\mathcal{S}_2}, \qquad \zeta_{ij} = -\mathring{\imath} \left\langle Y, \tilde{e}_{ij} \right\rangle_{\mathcal{S}_2}. \tag{37}$$

*They have mean zero and variance*

$$\operatorname{var} \xi_i = \mu_i^2, \qquad \operatorname{var} \xi_{ij} = \mu_i \mu_j, \qquad \operatorname{var} \zeta_{ij} = \mu_i \mu_j.$$

**Proof.** First notice that using Proposition 5.2,

$$\mathcal{C} = \sum_i \mu_i^2 \varphi_{ii} \bigotimes \varphi_{ii} + \sum_{i \neq j} \mu_i \mu_j \varphi_{ij} \bigotimes \varphi_{ij}. \tag{38}$$

Using this, we see that

$$\mathcal{C} = \sum_i \mu_i^2 \varphi_{ii} \bigotimes \varphi_{ii} + \sum_{i<j} \mu_i \mu_j \left( e_{ij} \bigotimes e_{ij} + \tilde{e}_{ij} \bigotimes \tilde{e}_{ij} \right).$$

The elements $\varphi_{ii}$, $e_{ij} (i < j)$ and $\tilde{e}_{ij} (i < j)$ are orthonormal, and are eigenvectors of $\mathcal{C}$. Thus (36) follows directly. Since $Y$ is Gaussian with mean zero, the variables defined in (37) are jointly Gaussian and have mean zero. Using the fact that $Y = Y^\dagger$ and Proposition 5.1, it is straightforward to see that they are real, uncorrelated (thus independent) random variables, with variance as given in the statement of the Lemma. □

**Lemma 4.5.** *Let $H$ be a real Hilbert space, and $C$ be a nuclear and self-adjoint operator on $H$, with spectral decomposition*

$$C = \sum_i \mu_i \varphi_{ii},$$

*where $\varphi_{ij} = \varphi_i \otimes \varphi_j$.*

*If $Y$ is a Gaussian random element on $\mathcal{S}_2(H)$ that takes self-adjoint values, with mean* 0 *and covariance operator*

$$\mathcal{C} = C\widetilde{\otimes}C + C\widetilde{\otimes}_{\mathsf{T}}C, \tag{39}$$

*then*

$$Y = \sum_i \xi_i \varphi_{ii} + \sum_{i<j} \xi_{ij} e_{ij}, \tag{40}$$

*where the convergence holds in expected mean square (in $\mathcal{S}_2(H)$), $e_{ij} = 2^{-1/2}(\varphi_{ij} + \varphi_{ji})$, and*

$$(\xi_i)_i \cup (\xi_{ij})_{i<j}$$

*are independent real Gaussian random variables, defined by*

$$\xi_i = \langle Y, \varphi_{ii} \rangle_{\mathcal{S}_2}, \qquad \xi_{ij} = \left\langle Y, e_{ij} \right\rangle_{\mathcal{S}_2}. \tag{41}$$

*They have mean zero and variance*

$$\operatorname{var} \xi_i = 2\mu_i^2, \quad \operatorname{var} \xi_{ij} = 2\mu_i \mu_j.$$

**Proof.** First we notice that

$$\mathcal{C} = \sum_i 2\mu_i^2 \varphi_{ii} \bigotimes \varphi_{ii} + \sum_{i<j} 2\mu_i \mu_j e_{ij} \bigotimes e_{ij},$$

by evaluating this expression and (39) on the elements $\varphi_{ij}$. The elements $\varphi_{ii}$ and $e_{ij}$ are orthonormal eigenelements of $\mathcal{C}$. The rest of the proof follows by arguments similar to those used to prove Lemma 4.4. □

We will use the same notation as in Section 4, but will suppress dependence on the frequency $\omega$, for tidiness. Simple calculations yield

$$\eta_k(\varphi_i \otimes \varphi_j) = E_k(i, j)\varphi_i \otimes \varphi_j,$$

where $E_k$ is defined by

$$E_k(i, j) = (\lambda_k - \lambda_s)^{-1} \quad \text{if for some } s \neq k, i \in I_k \text{ and } j \in I_s \quad \text{or} \quad j \in I_k \text{ and } i \in I_s,$$

and $E_k(i, j) = 0$ otherwise. Notice that $E_k(i, i) = 0$, and $E_k(i, j) = E_k(j, i)$. Thus $\eta_k(e_{ij}) = E_k(i, j)e_{ij}$ and $\eta_k(\tilde{e}_{ij}) = E_k(i, j)\tilde{e}_{ij}$. If $Y$ is a random element of $\mathcal{S}_2(H)$ of the form (36), we have

$$\eta_k(Y) = \sum_{i<j} E_k(i, j) \left[\xi_{ij} e_{ij} + \mathbf{i}\zeta_{ij} \tilde{e}_{ij}\right],$$

and thus

$$\eta_k(Y) \otimes \eta_l(Y) = \sum_{i<j} \sum_{s<t} E_k(i, j) E_l(s, t) \times \left[\xi_{ij}\xi_{st} e_{ij} \otimes e_{st} - \mathbf{i}\xi_{ij}\zeta_{st} e_{ij} \otimes \tilde{e}_{st} \right.$$
$$\left. + \mathbf{i}\zeta_{ij}\xi_{st}\tilde{e}_{ij} \otimes e_{st} + \zeta_{ij}\zeta_{st}\tilde{e}_{ij} \otimes \tilde{e}_{st}\right].$$

Therefore, if $Y$ is a random element of $\mathcal{S}_2(H)$ of the form (36), the covariance operator between $\eta_k(Y)$ and $\eta_l(Y)$ is given by

$$\begin{aligned}\mathbb{E}\left[\eta_k(Y)\otimes\eta_l(Y)\right] &= \sum_{i<j} E_k(i,j)E_l(i,j)\left[\mathrm{var}(\xi_{ij})e_{ij}\otimes e_{ij}+\mathrm{var}(\zeta_{ij})\tilde{e}_{ij}\otimes\tilde{e}_{ij}\right]\\ &= \sum_{i<j} E_k(i,j)E_l(i,j)\mu_i\mu_j\left[\varphi_{ii}\widetilde{\otimes}\varphi_{jj}+\varphi_{jj}\widetilde{\otimes}\varphi_{ii}\right]\\ &= \begin{cases}-\lambda_k\lambda_l(\lambda_k-\lambda_l)^{-2}\left[\Pi_k\widetilde{\otimes}\Pi_l+\Pi_l\widetilde{\otimes}\Pi_k\right] & \text{if } k\neq l\\ \sum_{s\neq k}\lambda_k\lambda_s(\lambda_k-\lambda_s)^{-2}\left[\Pi_k\widetilde{\otimes}\Pi_s+\Pi_s\widetilde{\otimes}\Pi_k\right] & \text{if } k=l.\end{cases}\end{aligned}$$

If $Y$ is of the form (40), we obtain

$$\mathbb{E}\left[\eta_k(Y)\otimes\eta_l(Y)\right] = \sum_{i<j} E_k(i,j)E_l(i,j)\left[\mathrm{var}(\xi_{ij})e_{ij}\otimes e_{ij}\right]$$

$$= \begin{cases}-\lambda_k\lambda_l(\lambda_k-\lambda_l)^{-2}\left[\Pi_k\widetilde{\otimes}\Pi_l+\Pi_l\widetilde{\otimes}\Pi_k+A_{kl}+A_{kl}^{\dagger}\right] & \text{if } k\neq l\\ \sum_{s\neq k}\lambda_k\lambda_s(\lambda_k-\lambda_s)^{-2}\left[\Pi_k\widetilde{\otimes}\Pi_s+\Pi_s\widetilde{\otimes}\Pi_k+A_{ks}+A_{ks}^{\dagger}\right] & \text{if } k=l,\end{cases}$$

where $A_{kl}=\varphi_{kl}\widetilde{\otimes}\varphi_{lk}$.

We now turn our attention to the covariance operator between $p_l(Y)$ and $\eta_k(Y)$,

$$\mathbb{E}\left[p_l(Y)\otimes\eta_k(Y)\right]: S_2(H)\to\mathbb{C}.$$

For any $f, g\in H$, we have

$$\begin{aligned}\mathbb{E}\left[p_l(Y)\otimes\eta_k(Y)\right](f\otimes g) &= \mathbb{E}p_l(Y)\left\langle f\otimes g,\eta_k(Y)\right\rangle_{\mathcal{S}_2}\\ &= \mathbb{E}p_l(Y)\left\langle f, S_kY\Pi_kg+\Pi_kYS_kg\right\rangle\\ &= \left\langle S_kf,\mathbb{E}\left[\overline{p_l(Y)}Y\right]\Pi_kg\right\rangle+\left\langle \Pi_kf,\mathbb{E}\left[\overline{p_l(Y)}Y\right]S_kg\right\rangle\end{aligned}$$

which, using the fact that $\mathbb{E}\left[\overline{p_l(Y)}Y\right]=\mathbb{E}\left[\langle\Pi_l,Y\rangle\,Y\right]=\mathcal{C}(\Pi_l)$, reduces to

$$\left\langle f\otimes g,\eta_k(\mathcal{C}\Pi_l)\right\rangle_{\mathcal{S}_2}=\left\langle f\otimes g,\eta_k(\lambda_l^2\Pi_l)\right\rangle_{\mathcal{S}_2}=0,$$

for all $l, k$.

For the covariance of the estimated eigenvalues, we use the KL expansion of $Y$: in both cases (36) and (40), it is given by

$$p_l(Y)=\sum_{i\in I_k}\xi_i. \tag{42}$$

This automatically gives us $\mathrm{cov}(p_l(Y),p_k(Y))=\delta_{lk}\lambda_l^2\,\mathrm{trace}\,(\Pi_l)$, if $Y$ is of the form (36), and $\mathrm{cov}(p_l(Y),p_k(Y))=\delta_{lk}2\lambda_l^2\,\mathrm{trace}\,(\Pi_l)$, if $Y$ is of the form (40).

## 5. Technical lemmas

For completeness, we present in this section some technical results used in the paper.

*5.1. Tensor and Kronecker products*

**Proposition 5.1.** *For any $u, v, f, g \in H$, $A, B \in \mathcal{S}_2(H)$*

1. $\cdot \otimes \cdot$ *is linear on the left, and sesquilinear on the right.*
2. $\langle u \otimes v, f \otimes g \rangle_{\mathcal{S}_2} = \langle u, f \rangle \langle g, v \rangle = \langle (u \otimes v)g, f \rangle$.
3. $\langle A, u \otimes v \rangle_{\mathcal{S}_2} = \langle Av, u \rangle = \langle v \otimes u, A^\dagger \rangle_{\mathcal{S}_2}$.
4. $|||u \otimes v|||_2 = \|u\| \, \|v\|$.
5. $(u \otimes v)(f \otimes g) = \langle f, v \rangle \, u \otimes g$.
6. $(u \otimes v)^\dagger = v \otimes u$, $(A \bigotimes B)^\dagger = B \bigotimes A$.

For two operators $A, B \in \mathcal{S}_2(H)$, we define their *Kronecker product* $A \widetilde{\otimes} B \in \mathcal{S}_2(\mathcal{S}_2(H))$, by $A \widetilde{\otimes} B(C) = ACB^\dagger$, for $C \in \mathcal{S}_2(H)$. It has the following properties:

**Proposition 5.2.** *For any $A, B, C, D \in \mathcal{S}_2(H)$, $u, v, f, g \in H$,*

1. $\cdot \widetilde{\otimes} \cdot$ *is linear on the left, and sesquilinear on the right*
2. $(A \widetilde{\otimes} B)(u \otimes v) = Au \otimes Bv$
3. $\langle A \widetilde{\otimes} B, C \widetilde{\otimes} D \rangle_{\mathcal{S}_2} = \langle A, C \rangle_{\mathcal{S}_2} \langle D, B \rangle_{\mathcal{S}_2}$
4. $|||A \widetilde{\otimes} B|||_2 = |||A|||_2 |||B|||_2$
5. $(A \widetilde{\otimes} B)(C \widetilde{\otimes} D) = AC \widetilde{\otimes} BD$
6. $(u \otimes v) \widetilde{\otimes} (f \otimes g) = (u \otimes f) \bigotimes (v \otimes g)$
7. $(A \widetilde{\otimes} B)^\dagger = A^\dagger \widetilde{\otimes} B^\dagger$.

In the case $H = L^2([0, 1], \mathbb{C})$, $A, B \in \mathcal{S}_2(H)$ are Hilbert–Schmidt operators, hence they are also kernel operators, with kernels $a(\tau, \sigma), b(\tau, \sigma)$, respectively. The operator $A \widetilde{\otimes} B$ is then also a Hilbert–Schmidt operator on $\mathcal{S}_2(H)$, with kernel $k(\tau, \sigma, x, y) = a(\tau, x)\overline{b(\sigma, y)}$. If $H_\mathbb{R}$ is a real Hilbert space, then for two operators $A, B \in \mathcal{S}_2(H_\mathbb{R})$, we also define their *transpose Kronecker product* $A \widetilde{\otimes}_\top B \in \mathcal{S}_2(\mathcal{S}_2(H_\mathbb{R}))$, by $A \widetilde{\otimes}_\top B(C) = (A \widetilde{\otimes} B)(C^\top) = AC^\top B^\top$, for $C \in \mathcal{S}_2(H_\mathbb{R})$.

**Proposition 5.3.** *For any $A, B, C, D \in \mathcal{S}_2(H_\mathbb{R})$, $u, v, f, g \in H_\mathbb{R}$,*

1. $\cdot \widetilde{\otimes}_\top \cdot$ *is bilinear*
2. $(A \widetilde{\otimes}_\top B)(u \otimes v) = Av \otimes Bu$
3. $\langle A \widetilde{\otimes}_\top B, C \widetilde{\otimes}_\top D \rangle_{\mathcal{S}_2} = \langle A, C \rangle_{\mathcal{S}_2} \langle D, B \rangle_{\mathcal{S}_2}$
4. $|||A \widetilde{\otimes}_\top B|||_2 = |||A|||_2 |||B|||_2$.

In the case $H_\mathbb{R} = L^2([0, 1], \mathbb{R})$, if $A, B \in \mathcal{S}_2(H_\mathbb{R})$ are Hilbert–Schmidt operators, they are also kernel operators, with kernels $a(\tau, \sigma), b(\tau, \sigma)$, respectively. The operator $A \widetilde{\otimes}_\top B$ is then also a Hilbert–Schmidt operator on $\mathcal{S}_2(H_\mathbb{R})$, with kernel

$$k(\tau, \sigma, x, y) = a(\tau, y) b(\sigma, x).$$

**Lemma 5.4.** *If $\sum_t |||\mathscr{R}_t|||_1 < \infty$, and $\alpha < \beta$, the operator $\int_\alpha^\beta \mathscr{F}_\omega d\omega$ is non-negative, and*

$$\left|\left|\left| \int_\alpha^\beta \mathscr{F}_\omega d\omega \right|\right|\right|_1 = \int_\alpha^\beta |||\mathscr{F}_\omega|||_1 d\omega.$$

**Lemma 5.5.** *For $A \in \mathcal{S}_2(H)$, $H = L^2([0, 1], \mathbb{C})$, we have*

$$A \int_{-\pi}^{\pi} e^{\mathbf{i} t\omega} dZ_\omega = \int_{-\pi}^{\pi} e^{\mathbf{i} t\omega} A dZ_\omega.$$

## Acknowledgements

This research was supported by a European Research Council (ERC) Starting Grant Award. We wish to thank the editors and reviewers for their thoughtful and constructive comments.

## References

[1] A.A. Amini, M.J. Wainwright, Sampled forms of functional PCA in reproducing kernel Hilbert spaces, Annals of Statistics 40 (5) (2012) 2452–2482.
[2] P. Besse, J.O. Ramsay, Principal components-analysis of sampled functions, Psychometrika 51 (1986) 285–311.
[3] D. Bosq, Propriétés des opérateurs de covariance empiriques d'un processus stationnaire Hilbertien, Comptes Rendus de l'Académie des Sciences. Série I 309 (14) (1989) 873–875.
[4] D. Bosq, Modèle autorégressif Hilbertien, application à la prévision du comportement d'un processus á temps continu sur un intervalle de temps donné, Comptes Rendus de l'Académie des Sciences. Série I 310 (11) (1990) 787–790.
[5] D. Bosq, Estimation of mean and covariance operator of autoregressive processes in Banach spaces, Statistical Inference for Stochastic Processes 5 (2002) 287–306. 10.1023/A:1021279131053.
[6] D. Bosq, Linear Processes in Function Spaces, Springer, 2000.
[7] D. Bosq, D. Blanke, Inference and Prediction in Large Dimensions, in: Wiley Series in Probability and Statistics, John Wiley & Sons, 2008.
[8] D.R. Brillinger, Time Series: Data Analysis and Theory ((classics edition) ed.), in: Classics in Applied Mathematics, SIAM, 2001.
[9] C. Brislawn, Kernels of trace class operators, in: Proc. Am. Math. Soc, vol. 104, 1988, pp. 1181–1190.
[10] H. Cardot, Nonparametric estimation of smoothed principal components analysis of sampled noisy functions, Journal of Nonparametric Statistics 12 (2000) 503–538.
[11] H. Cardot, Conditional functional principal components analysis, Scandinavian Journal of Statistics 34 (2) (2007) 317–335.
[12] H. Cramér, On harmonic analysis in certain functional spaces, Arkiv för Matematik, Astronomi och Fysik 28B (1942) 1–7.
[13] J. Dauxois, A. Pousse, Y. Romain, Asymptotic theory for the principal component analysis of a vector random function: some applications to statistical inference, Journal of Multivariate Analysis 12 (1982) 136–154.
[14] N. Dinculeanu, Vector Integration and Stochastic Integration in Banach Spaces, in: Pure and Applied Mathematics, Wiley, New York, 2000.
[15] F. Ferraty, Y. Romain, The Oxford Handbook of Functional Data Analysis, Oxford University Press, 2011.
[16] U. Grenander, Stochastic processes and statistical inference, Arkiv för Matematik 1 (1950) 195–277.
[17] U. Grenander, Abstract Inference, in: Wiley Series in Probability and Mathematical Statistics, vol. IX, John Wiley & Sons, New York etc., 1981, p. 526.
[18] P. Hall, M. Hosseini-Nasab, On properties of functional principal components analysis, Journal of the Royal Statistical Society, Series B, Statistical Methodology 68 (1) (2006) 109–126.
[19] P. Hall, M. Hosseini-Nasab, Theory for high-order bounds in functional principal components analysis, Mathematical Proceedings of the Cambridge Philosophical Society 146 (1) (2009) 225–256.
[20] P. Hall, H.G. Müller, J.L. Wang, Properties of principal component methods for functional and longitudinal data analysis, Annals of Statistics 34 (3) (2006) 1493–1517.
[21] S. Hörmann, P. Kokoszka, Weakly dependent functional data, Annals of Statistics 38 (3) (2010) 1845–1884.
[22] L. Horváth, P. Kokoszka, Inference for Functional Data with Applications, in: Springer Series in Statistics, Springer, New York, NY, 2012.
[23] L. Horváth, P. Kokoszka, R. Reeder, Estimation of the mean of functional time series and a two-sample problem, Journal of the Royal Statistical Society: Series B (Statistical Methodology) 75 (1) (2013) 103–122.
[24] K. Karhunen, Über lineare Methoden in der Wahrscheinlichkeitsrechnung, Annales Academiae Scientiarum Fennicae. Series A I (37) (1947) 79.
[25] J. Kleffe, Principal components of random variables with values in a seperable Hilbert space, Mathematische Operationsforschung Statistik 4 (5) (1973) 391–406.
[26] P. Kokoszka, Dependent functional data, ISRN Probability and Statistics 2012 (2012).
[27] P. Lévy, Fonctions aléatoires du seconde ordre, in: Processus stochastiques et mouvement Brownien, Suivi d'une note de M. Loève. Levy, P., Gauthier-Villars, Éditeur, Paris, 1948, p. 365.

[28] A. Mas, Estimation d'opérateurs de corrélation de processus linéaires fonctionnels: lois limites, déviations modérées, Ph.D. Thesis, Université Paris VI, 2000.
[29] A. Mas, Weak convergence for the covariance operators of a Hilbertian linear process, Stochastic Processes and their Applications 99 (2002) 117–135.
[30] A. Mas, L. Menneteau, Perturbation approach applied to the asymptotic study of random operators, in: High Dimensional Probability III, vol. 55, Birkhäuser, 2003, pp. 127–133.
[31] A. Mas, B. Pumo, Linear processes for functional data, Arxiv preprint arXiv:0901.2503, 2009.
[32] V.M. Panaretos, S. Tavakoli, Fourier analysis of stationary processes in function space, Annals of Statistics, 2013, (in press).
[33] V.M. Panaretos, S. Tavakoli, Cramér–Karhunen–Loève representation and harmonic principal component analysis of functional time series, Technical Report #03–12, Chair of Mathematical Statistics (EPFL), 2012 November.
[34] J. Ramsay, B. Silverman, Functional Data Analysis, second ed., Springer, 2005.
[35] J.A. Rice, B. Silverman, Estimating the mean and covariance structure nonparametrically when the data are curves, Journal of the Royal Statistical Society. Series B 53 (1) (1991) 233–243.
[36] B.W. Silverman, Smoothed functional principal components analysis by choice of norm, Annals of Statistics 24 (1) (1996) 1–24.
[37] J.M. Steele, Stochastic Calculus and Financial Applications (Corr. 3rd printing ed.), in: 3rd corr. pr of Stochastic Modelling and Applied Probability, vol. 45, Springer, New York, 2003.
[38] F. Yao, H.G. Müller, J.L. Wang, Functional data analysis of sparse longitudinal data, Journal of the American Statistical Association 100 (2005) 577–590.