JOURNAL OF TIME SERIES ANALYSIS
J. Time Ser. Anal. **39**: 502–522 (2018)
Published online 6 February 2018 in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1111/jtsa.12283

ORIGINAL ARTICLE

PRINCIPAL COMPONENTS ANALYSIS OF PERIODICALLY CORRELATED FUNCTIONAL TIME SERIES

ŁUKASZ KIDZIŃSKI,[a] PIOTR KOKOSZKA[b*] AND NEDA MOHAMMADI JOUZDANI[c]

[a] *Bioengineering Department, Stanford University, Stanford, CA, USA*
[b] *Statistics Department, Colorado State University, Fort Collins, CO, USA*
[c] *Department of Mathematical Sciences, Isfahan University of Technology, Isfahan, Iran*

Within the framework of functional data analysis, we develop principal component analysis for periodically correlated time series of functions. We define the components of the above analysis including periodic operator-valued filters, score processes, and the inversion formulas. We show that these objects are defined via a convergent series under a simple condition requiring summability of the Hilbert–Schmidt norms of the filter coefficients and that they possess optimality properties. We explain how the Hilbert space theory reduces to an approximate finite-dimensional setting which is implemented in a custom-build |R| package. A data example and a simulation study show that the new methodology is superior to existing tools if the functional time series exhibits periodic characteristics.

Received 24 November 2016; Accepted 22 December 2017

Keywords: Functional time series; periodically correlated time series; principal components; spectral analysis

MOS subject classification: 62M10; 62M15.

1. INTRODUCTION

Periodicity is one of the most important characteristics of time series, with early work going back to the very origins of the field, e.g. Walker (1914) and Fisher (1929). The class of periodically correlated time series is particularly suitable to quantify periodic behavior reflected not only in the mean structure but also in correlations. Consequently, periodically correlated (PC) time series have been used in many modeling and prediction applications, and various aspects of their theory have been studied. The book by Hurd and Miamee (2007) gives an account of the subject. It is impossible to list even a fraction of relevant references, but to indicate the many flavors of work done in this field, we cite Hurd (1989), Lund *et al.* (1995), Anderson and Meerschaert (1997), Javorskyj *et al.* (2012), and Ghanbarzadeh and Aminghafari (2016).

The last decade has seen increased interest in time series of curves, often referred to as functional time series (FTS). Examples of FTS include annual temperature or smoothed precipitation curves, e.g. Gromenko *et al.* (2017), daily pollution level curves, e.g. Aue *et al.* (2015), various daily curves derived from high-frequency asset price data, e.g. Horváth *et al.* (2014), yield curves, e.g. Hays *et al.* (2012), and daily vehicle traffic curves, e.g. Klepsch *et al.* (2017). Other examples are given in the books by Horváth and Kokoszka (2012) and Kokoszka and Reimherr (2017). The theory and methodology of FTS forms a subfield of functional data analysis (FDA). A key tool of FDA is dimension reduction via functional principal component analysis (FPCA), see, e.g. Chapter 3 of Horváth and Kokoszka (2012). FPCA has been developed for random samples of functions, i.e. for i.i.d. functional data. Recently, Hörmann *et al.* (2015) extended the theory of Brillinger (1975, Chapter 9) from linear vector-valued

* Correspondence to: Piotr Kokoszka, Statistics Department, Colorado State University, Fort Collins, CO 80523-1877, USA. E-mail: Piotr.Kokoszka@colostate.edu

Copyright © 2018 John Wiley & Sons Ltd

time series to functional weakly dependent time series. Building on earlier advances of Panaretos and Tavakoli (2013b, 2013a), they developed spectral domain PCA, which leads to a better representation of stationary FTS than the usual (static) PCA. Suitable details and definitions are given in Section 2 . The objective of this paper is to develop PCA for PC FTS. We establish the requisite theoretical framework and show that for FTS with periodic characteristics the new approach is superior to the methodology of Hörmann *et al.* (2015). We emphasize that the latter methodology was developed for stationary FTS and so is a priori not well suited for periodic functional data. Tests for periodicity in FTS have recently been developed by Hörmann *et al.* (2018) and Zamani *et al.* (2016). Zhang (2016) uses spectral methods to develop goodness-of-fit tests for FTS.

Section 2 introduces the requisite background and notation. The theory of PCA of PC FTS is presented in Section 3, with proofs postponed to Section 6. Section 4 shows how the methodology developed in the infinite-dimensional framework of function spaces is translated into an implementable setting of finite-dimensional objects. Its usefulness is illustrated in Section 5 by an application to a particulate pollution dataset and a simulation study. On-line Supporting Information contains additional information and selected proofs, referred to at appropriate locations in the paper.

2. NOTATION AND PRELIMINARIES

This section introduces the notation and background used throughout the paper. A generic separable Hilbert space is denoted by $\mathbb{H}$, and its inner product and norm are denoted, respectively, by $< .,. >_{\mathbb{H}}$ and $\|.\|_{\mathbb{H}}$. The subscript $\mathbb{H}$ is sometimes suppressed when there is no ambiguity.

The Hilbert space $\mathcal{H} = L^2([0,1])$ and its T-fold Cartesian product $\mathcal{H}^T$ are extensively used throughout this paper. They are equipped with inner products

$$\langle f, g\rangle_{\mathcal{H}} = \int_0^1 f(s)\overline{g}(s)\,\mathrm{d}s, \quad f, g \in \mathcal{H}$$

and

$$\left\langle \left(f_1 \cdots f_T\right)', \left(g_1 \cdots g_T\right)' \right\rangle_{\mathcal{H}^T} = \sum_{j=1}^{T} \left\langle f_j, g_j\right\rangle_{\mathcal{H}}, \quad f_j, g_j \in \mathcal{H},$$

respectively. An operator Ψ from a Hilbert space $\mathbb{H}$ to $\mathbb{C}^p$ is a bounded linear operator if and only if there exist (unique) elements $\Psi_1, \ldots, \Psi_p$ in $\mathbb{H}$ such that

$$\Psi(h) = \left(\langle h, \Psi_1\rangle_{\mathbb{H}}, \ldots, \left\langle h, \Psi_p\right\rangle_{\mathbb{H}}\right)', \quad \forall\, h \in \mathbb{H}. \tag{2.1}$$

An operator Υ from $\mathbb{C}^p$ to $\mathbb{H}$ is linear and bounded if and only if there exist elements $\Upsilon_1, \ldots, \Upsilon_p$ in $\mathbb{H}$ such that

$$\Upsilon(y) = \Upsilon\left(\left(y_1, \ldots, y_p\right)'\right) = \sum_{m=1}^{p} y_m \Upsilon_m, \quad \forall\, y \in \mathbb{C}^p.$$

For any two elements f and g in $\mathbb{H}$, $f \otimes g$ is a bounded linear operator defined by

$$f \otimes g : \mathbb{H} \to \mathbb{H}, \quad f \otimes g : h \longmapsto \langle h, g\rangle_{\mathbb{H}} f.$$

We use $\|.\|_{\mathcal{L}}$ to denote the operator norm, and $\|.\|_{\mathcal{N}}$ and $\|.\|_{\mathcal{S}}$ to denote, respectively, the nuclear and Hilbert–Schmidt norms, see e.g. Horváth and Kokoszka (2012), Section 13.5.

Copyright © 2018 John Wiley & Sons Ltd

DOI: 10.1111/jtsa.12283

In the following, $L^2(\mathbb{H}, (-\pi, \pi])$ denotes the space of square integrable $\mathbb{H}$-valued functions on $(-\pi, \pi]$. Similarly, for a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ in place of $(-\pi, \pi]$, we use the notation $L^2(\mathbb{H}, \Omega)$. For two random elements $X, Y \in L^2(\mathbb{H}, \Omega)$, the covariance operator, $\mathrm{Cov}(X, Y)$, is defined as

$$\mathrm{Cov}(X, Y) = E[(X - EX) \otimes (Y - EY)] : \mathbb{H} \to \mathbb{H},$$
$$\mathrm{Cov}(X, Y) : h \longmapsto E\left[\langle h, (Y - EY)\rangle_{\mathbb{H}} (X - EX)\right].$$

Definition 2.1. Let $X = \{X_t, t \in \mathbb{Z}\}$ be an $\mathbb{H}$-valued time series with finite second moment $E\|X_t\|^2 < \infty$. Then, X is said to be PC if there exists a positive integer T such that

$$EX_t = EX_{t+T}, \quad \forall\, t \in \mathbb{Z},$$
$$\mathrm{Cov}\left(X_t, X_s\right) = \mathrm{Cov}\left(X_{t+T}, X_{s+T}\right), \quad \forall\, t, s \in \mathbb{Z}.$$

The smallest such T will be called the period of the process, and X is then said to be T-PC, or T-PC, for short. When $T = 1$, the process is (weakly) stationary.

For a T-PC process $\{X_t\}$, covariance operators at lag h are defined as

$$C^X_{h,(j,j')} = \mathrm{Cov}\left(X_{Th+j}, X_{j'}\right), \quad h \in \mathbb{Z} \text{ and } j, j' = 0, 1, \ldots, T-1.$$

It is easy to verify that the condition

$$\sum_{h\in\mathbb{Z}} \left\|C^X_{h,(j,j')}\right\|_{\mathcal{S}} < \infty, \quad j, j' = 0, 1, \ldots, T-1, \tag{2.2}$$

implies that for each θ the series $\left\{\frac{1}{2\pi}\sum_{h=-n}^{n} C^X_{h,(j,j')} \mathrm{e}^{-\mathrm{i}h\theta} : n \in \mathbb{Z}_+\right\}$ is a Cauchy sequence in the Hilbert space of Hilbert–Schmidt operators on $\mathbb{H}$. Then, spectral density operators are well defined by

$$\mathcal{F}^X_{\theta,(j,j')} = \frac{1}{2\pi}\sum_{h\in\mathbb{Z}} C^X_{h,(j,j')} \mathrm{e}^{-\mathrm{i}h\theta}, \quad j, j' = 0, \ldots, T-1. \tag{2.3}$$

Definition 2.2. A sequence $\{\Psi_l, l \in \mathbb{Z}\}$ of operators from a Hilbert space $\mathbb{H}_1$ to a Hilbert space $\mathbb{H}_2$ satisfying

$$\sum_{l\in\mathbb{Z}} \|\Psi_l\|_{\mathcal{L}} < \infty, \tag{2.4}$$

is called a filter. A T-periodic filter $\left\{\left\{\Psi^t_l, l \in \mathbb{Z}\right\}, t \in \mathbb{Z}\right\}$ is a sequence of filters which is T-periodic with respect to t, i.e. $\Psi^t_l = \Psi^{t+T}_l$, for each t and l. Consequently,

$$\sum_{t=0}^{T-1}\sum_{l\in\mathbb{Z}} \|\Psi^t_l\|_{\mathcal{L}} < \infty. \tag{2.5}$$

Related to the filter $\{\Psi_l, l \in \mathbb{Z}\}$, $\Psi(B)$ is an operator from $\left(\mathbb{H}_1\right)^{\mathbb{Z}}$ to $\left(\mathbb{H}_2\right)^{\mathbb{Z}}$ of the following form:

$$\Psi(B) = \sum_{l\in\mathbb{Z}} \Psi_l B^l,$$

Copyright © 2018 John Wiley & Sons Ltd

DOI: 10.1111/jtsa.12283

where B is the backward shift operator. In other words, if $\left\{X_t, t \in \mathbb{Z}\right\}$ is a time series with values in $\mathbb{H}_1$, then $\Psi(B)$ transforms it to an $\mathbb{H}_2$ -valued time series defined by

$$(\Psi(B)(X))_t = \sum_{l \in \mathbb{Z}} \Psi_l\left(X_{t-l}\right).$$

For a $p \times p$ matrix $\mathbf{A}$, $a_{q,r}$ denotes the entry in the qth row and rth column. To indicate that $t = kT + d$ for some integer k , we write $t \overset{T}{\equiv} d$.

3. PRINCIPAL COMPONENT ANALYSIS OF PC FUNCTIONAL TIME SERIES

Before proceeding with the definitions and statements of properties of the principal component analysis for PC-FTS, we provide a brief introduction, focusing on the ideas and omitting mathematical assumptions. Suppose $\{X_t\}$ is a weakly dependent, stationary, mean-zero time series of functions in $\mathcal{H}$. It admits the Karhunen–Lo éve expansion

$$X_t(u) = \sum_{m=1}^{\infty} \xi_{tm} v_m(u), \quad \xi_{tm} = \langle X_t, v_m \rangle, \quad E\xi_{tm}^2 = \lambda_m, \tag{3.1}$$

where the v_m are the functional principal components (called *static* FPCs in Hörmann *et al.* 2015). The orthonormal functions v_m are uniquely defined up to a sign, and the random variables ξ_{tm} are called their scores. Even for stationary (rather than PC) FTS, the *dynamic* FPCs are not defined as one function for every "frequency" level m. The analog of 3.1 is

$$X_t(u) = \sum_{m=1}^{\infty} \sum_{l \in \mathbb{Z}} Y_{m,t+l} \phi_{ml}(u). \tag{3.2}$$

A single function v_m is thus replaced by an infinite sequence of functions $\{\phi_{ml}, l \in \mathbb{Z}\}$. However, one can still define the scores as single numbers for every frequency level m, using the formula $Y_{mt} = \sum_{l \in \mathbb{Z}} \langle X_{t-l}, \phi_{ml} \rangle$. The analog of λ_m is

$$\nu_m := E\left\| \sum_{l \in \mathbb{Z}} Y_{m,t+l} \phi_{ml} \right\|^2,$$

and we have the decomposition of variance $E\left\|X_t\right\|^2 = \sum_{m=1}^{\infty} \nu_m$. In this section, we will see how these results extend to the setting of PC FTS, which is necessarily more complex as it involves periodic sequences of functions. The scores, and the reconstructions obtained from them, will have certain periodic properties. *All results stated in this section are proven in Section 6.*

In order to define the dynamic functional principal components (DFPCs) in our setting, we first establish conditions for the existence of a filtered (output) process of a T-PC FTS. The periodic structure of the covariance operators of the T-PC input process $X = \left\{X_t, t \in \mathbb{Z}\right\}$ suggests applying a T-periodic functional filter $\left\{\left\{\Psi_l^t, l \in \mathbb{Z}\right\}, t \in \mathbb{Z}\right\}$ to obtain a filtered process $\mathbf{Y} = \left\{\mathbf{Y}_t, t \in \mathbb{Z}\right\}$ with values in $\mathbb{C}^p$.

Theorem 3.1. Let $X = \left\{X_t, t \in \mathbb{Z}\right\}$ be an $\mathcal{H}$-valued T-PC process and $\left\{\left\{\Psi_l^t, l \in \mathbb{Z}\right\}, t \in \mathbb{Z}\right\}$ a T-periodic filter from $\mathcal{H}$ to $\mathbb{C}^p$ with the elements $\Psi_{l,m}^t, m = 1, \dots, p$ in $\mathcal{H}$, as described in (2.1). In particular, we assume that (2.5) holds. Then, for each t, $\sum_{l \in \mathbb{Z}} \Psi_l^t\left(X_{(t-l)}\right)$ converges in mean square to a limit $\mathbf{Y}_t$.

 Copyright © 2018 John Wiley & Sons Ltd
DOI: 10.1111/jtsa.12283

If, in addition

$$\sum_{t=0}^{T-1} \sum_{l \in \mathbb{Z}} \left\| \Psi_l^t \right\|_{\mathcal{S}} < \infty, \tag{3.3}$$

then $\mathbf{Y} = \left\{ \mathbf{Y}_t, t \in \mathbb{Z} \right\}$ is a T-PC process with the following $p \times p$ spectral density matrices $\mathcal{F}^{\mathbf{Y}}_{\theta,(d,f)}$ for $d,f = 0, \dots, T-1$

$$\mathcal{F}^{\mathbf{Y}}_{\theta,(d,f)} = \left[\left\langle \begin{pmatrix} \mathcal{F}^X_{\theta,(0,0)} & \cdots & \mathcal{F}^X_{\theta,(0,T-1)} \\ \vdots & \ddots & \vdots \\ \mathcal{F}^X_{\theta,(T-1,0)} & \cdots & \mathcal{F}^X_{\theta,(T-1,T-1)} \end{pmatrix} \begin{pmatrix} \Psi^d_{\theta,d,r} \\ \vdots \\ \Psi^d_{\theta,d-T+1,r} \end{pmatrix}, \begin{pmatrix} \Psi^f_{\theta,f,q} \\ \vdots \\ \Psi^f_{\theta,f-T+1,q} \end{pmatrix} \right\rangle_{\mathcal{H}^T} \right]_{q,r=1,\dots,p},$$

where $\Psi^d_{\theta,d,q} = \sum_{l \in \mathbb{Z}} \Psi^d_{Tl+d,q} e^{il\theta}, \dots, \Psi^d_{\theta,d-T+1,q} = \sum_{l \in \mathbb{Z}} \Psi^d_{Tl+d-T+1,q} e^{il\theta}, f, d = 0, \dots, T-1$.

To illustrate the spectral density structure of the output process, we consider $T = 2$, in which case

$$\mathcal{F}^{\mathbf{Y}}_{\theta,(0,0)} = \left[\left\langle \begin{pmatrix} \mathcal{F}^X_{\theta,(0,0)} & \mathcal{F}^X_{\theta,(0,1)} \\ \mathcal{F}^X_{\theta,(1,0)} & \mathcal{F}^X_{\theta,(1,1)} \end{pmatrix} \begin{pmatrix} \Psi^0_{\theta,0,r} \\ \Psi^0_{\theta,-1,r} \end{pmatrix}, \begin{pmatrix} \Psi^0_{\theta,0,q} \\ \Psi^0_{\theta,-1,q} \end{pmatrix} \right\rangle_{\mathcal{H}^2} \right]_{q,r=1,\dots,p},$$

$$\mathcal{F}^{\mathbf{Y}}_{\theta,(1,0)} = \left[\left\langle \begin{pmatrix} \mathcal{F}^X_{\theta,(0,0)} & \mathcal{F}^X_{\theta,(0,1)} \\ \mathcal{F}^X_{\theta,(1,0)} & \mathcal{F}^X_{\theta,(1,1)} \end{pmatrix} \begin{pmatrix} \Psi^1_{\theta,1,r} \\ \Psi^1_{\theta,0,r} \end{pmatrix}, \begin{pmatrix} \Psi^0_{\theta,0,q} \\ \Psi^0_{\theta,-1,q} \end{pmatrix} \right\rangle_{\mathcal{H}^2} \right]_{q,r=1,\dots,p},$$

$$\mathcal{F}^{\mathbf{Y}}_{\theta,(0,1)} = \left[\left\langle \begin{pmatrix} \mathcal{F}^X_{\theta,(0,0)} & \mathcal{F}^X_{\theta,(0,1)} \\ \mathcal{F}^X_{\theta,(1,0)} & \mathcal{F}^X_{\theta,(1,1)} \end{pmatrix} \begin{pmatrix} \Psi^0_{\theta,0,r} \\ \Psi^0_{\theta,-1,r} \end{pmatrix}, \begin{pmatrix} \Psi^1_{\theta,1,q} \\ \Psi^1_{\theta,0,q} \end{pmatrix} \right\rangle_{\mathcal{H}^2} \right]_{q,r=1,\dots,p},$$

$$\mathcal{F}^{\mathbf{Y}}_{\theta,(1,1)} = \left[\left\langle \begin{pmatrix} \mathcal{F}^X_{\theta,(0,0)} & \mathcal{F}^X_{\theta,(0,1)} \\ \mathcal{F}^X_{\theta,(1,0)} & \mathcal{F}^X_{\theta,(1,1)} \end{pmatrix} \begin{pmatrix} \Psi^1_{\theta,1,r} \\ \Psi^1_{\theta,0,r} \end{pmatrix}, \begin{pmatrix} \Psi^1_{\theta,1,q} \\ \Psi^1_{\theta,0,q} \end{pmatrix} \right\rangle_{\mathcal{H}^2} \right]_{q,r=1,\dots,p},$$

where $\Psi^0_{\theta,0,q} := \sum_{l \in \mathbb{Z}} \Psi^0_{2l,q} e^{il\theta}$, $\Psi^0_{\theta,-1,q} := \sum_{l \in \mathbb{Z}} \Psi^0_{2l-1,q} e^{il\theta}$, $\Psi^1_{\theta,0,q} := \sum_{l \in \mathbb{Z}} \Psi^1_{2l,q} e^{il\theta}$, and $\Psi^1_{\theta,1,q} := \sum_{l \in \mathbb{Z}} \Psi^1_{2l+1,q} e^{il\theta}$.

We emphasize that (2.5) is a sufficient condition for the mean-square convergence of the series defining the filtered process $\mathbf{Y}$, and (3.3) guarantees the existence of the spectral density operator of the filtered process. Hörmann *et al.* (2015, p. 327), discuss this issue in the case of stationary input and output processes. *In the remainder of the paper, we assume (3.3) for each periodic functional filter.*

The operator matrix $(\mathcal{F}^X_{\theta,(d,f)})_{0 \le d,f \le T-1}$ in Theorem 3.1 is a non-negative, self-adjoint, compact operator from $\mathcal{H}^T$ to $\mathcal{H}^T$, and so it admits the following spectral decomposition:

$$\left(\mathcal{F}^X_{\theta,(d,f)} \right)_{0 \le d,f \le T-1} = \sum_{m \ge 1} \lambda_{\theta,m} \varphi_{\theta,m} \otimes \varphi_{\theta,m}, \tag{3.4}$$

where $\lambda_{\theta,1} \ge \lambda_{\theta,2} \ge \cdots \ge 0$, and $\left\{ \varphi_{\theta,m} \right\}_{m \ge 1}$ forms a complete orthonormal basis for $\mathcal{H}^T$. By choosing $\left(\Psi^d_{\theta,d,q} \cdots \Psi^d_{\theta,d-T+1,q} \right)'$ as the eigenfunction $\varphi_{\theta,dp+q}$, the spectral density matrices of the filtered process $\mathbf{Y} = \left\{ \mathbf{Y}_t, t \in \mathbb{Z} \right\}$ turn to diagonal matrices and an optimality property will be obtained. We are now ready to define the DFPC filter and scores of the PC process X.

 Copyright © 2018 John Wiley & Sons Ltd
DOI: 10.1111/jtsa.12283

Definition 3.1. Let $X = \{X_t, t \in \mathbb{Z}\}$ be an $\mathcal{H}$ -valued mean-zero T-PC random process satisfying condition (2.2) and $\left\{\Phi^d_{l,m}, d = 0, \dots, T-1, m = 1, \dots p, l \in \mathbb{Z}\right\}$ be elements of $\mathcal{H}$ defined by

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \varphi_{\theta,dp+m} \mathrm{e}^{-\mathrm{i}l\theta}\, \mathrm{d}\theta = \begin{pmatrix} \Phi^d_{lT+d,m} \\ \vdots \\ \Phi^d_{lT+d-T+1,m} \end{pmatrix}, \quad m = 1, \dots, p,\ d = 0, \dots, T-1 \tag{3.5}$$

for each l in $\mathbb{Z}$, or equivalently by

$$\varphi_{\theta,dp+m} = \begin{pmatrix} \Phi^d_{\theta,d,m} \\ \vdots \\ \Phi^d_{\theta,d-T+1,m} \end{pmatrix}, \quad m = 1, \dots, p,\ d = 0, \dots, T-1, \tag{3.6}$$

for each θ in $(-\pi, \pi]$. Then

$$\left\{\Phi^d_{l,m}, l \in \mathbb{Z}\right\}, \quad d = 0, \dots, T-1$$

is said to be the (d, m)th DFPC filter of the process X. Furthermore

$$\begin{aligned} Y_{t,m} &= \sum_{l\in\mathbb{Z}} \left\langle X_{(t-l)}, \Phi^d_{l,m} \right\rangle \\ &= \sum_{l\in\mathbb{Z}} \left\langle X_{(t-lT-d)}, \Phi^d_{lT+d,m} \right\rangle + \sum_{l\in\mathbb{Z}} \left\langle X_{(t-lT-d+1)}, \Phi^d_{lT+d-1,m} \right\rangle \\ &\quad + \cdots + \sum_{l\in\mathbb{Z}} \left\langle X_{(t-lT-d+T-1)}, \Phi^d_{lT+d-T+1,m} \right\rangle, \quad m = 1, \dots p,\ t \overset{T}{\equiv} d \end{aligned} \tag{3.7}$$

will be called the (t, m)th DFPC score of X.

For illustration, in the case of $T = 2$, we have for $m = 1, \dots, p$

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \varphi_{\theta,m} \mathrm{e}^{-\mathrm{i}l\theta}\, \mathrm{d}\theta = \begin{pmatrix} \Phi^0_{2l,m} \\ \Phi^0_{2l-1,m} \end{pmatrix} \text{ and } \frac{1}{2\pi}\int_{-\pi}^{\pi} \varphi_{\theta,p+m} \mathrm{e}^{-\mathrm{i}l\theta}\, \mathrm{d}\theta = \begin{pmatrix} \Phi^1_{2l+1,m} \\ \Phi^1_{2l,m} \end{pmatrix}, \tag{3.8}$$

for each l in $\mathbb{Z}$, or equivalently

$$\varphi_{\theta,m} = \begin{pmatrix} \Phi^0_{\theta,0,m} \\ \Phi^0_{\theta,-1,m} \end{pmatrix} \text{ and } \varphi_{\theta,p+m} = \begin{pmatrix} \Phi^1_{\theta,1,m} \\ \Phi^1_{\theta,0,m} \end{pmatrix}, \quad \theta \in (-\pi, \pi].$$

The filters $\left\{\Phi^d_{l,m}, l \in \mathbb{Z}\right\}$ are defined for $d = 0, 1$.

The following proposition lists some useful properties of the p-dimensional output process $\left\{\mathbf{Y}_t = (Y_{t,1}, \dots, Y_{t,p})',\ t \in \mathbb{Z}\right\}$ defined by (3.7).

Proposition 3.1. Let $X = \{X_t, t \in \mathbb{Z}\}$ be an $\mathcal{H}$-valued, mean-zero T-PC random process and assume that (2.2) holds. Then

- the eigenfunctions $\varphi_{\theta,m}$ are Hermitian i.e. $\varphi_{-\theta,m} = \overline{\varphi}_{\theta,m}$ and the DFPC scores $Y_{t,m}$ are real-valued provided that X is real-valued;

 Copyright © 2018 John Wiley & Sons Ltd
DOI: 10.1111/jtsa.12283

- for each (t, m), the series (3.7) is mean-square convergent, has mean zero:

$$EY_{t,m} = 0, \tag{3.9}$$

and satisfies for $t \overset{T}{\equiv} d$,

$$E\left\|Y_{t,m}\right\|^2 = \sum_{j_1,j_2=0}^{T-1} \sum_{k,l\in\mathbb{Z}} \left\langle C^X_{k-l,(j_1,j_2)}\left(\Phi^d_{kT+d-j_2,m}\right), \Phi^d_{lT+d-j_1,m}\right\rangle_{\mathcal{H}}; \tag{3.10}$$

- for any t and s, the DFPC scores $Y_{t,m}$ and $Y_{s,m'}$ are uncorrelated if $s-t$ is not a multiple of T or $m \neq m'$. In other words $C^{\mathbf{Y}}_{h,(j_1,j_2)} = 0$ for $j_1 \neq j_2$ and $C^{\mathbf{Y}}_{h,(j,j)}$ are diagonal matrices for all h;
- the long-run covariance matrix of the filtered process $\left\{\mathbf{Y}_t, t \in \mathbb{Z}\right\}$ satisfies the following limiting equality:

$$\lim_{n\to\infty} \frac{1}{n} \mathrm{Var}\left(\mathbf{Y}_1 + \cdots + \mathbf{Y}_n\right) = \frac{2\pi}{T} \sum_{d=0}^{T-1} \mathrm{diag}\left(\lambda_{0,dp+1}, \ldots, \lambda_{0,dp+p}\right).$$

For illustration, if $T = 2$, then

$$\begin{aligned} E\left\|Y_{t,m}\right\|^2 = & \sum_{k\in\mathbb{Z}} \sum_{l\in\mathbb{Z}} \left\langle C^X_{k-l,(0,0)}\left(\Phi^0_{2k,m}\right), \Phi^0_{2l,m}\right\rangle_{\mathcal{H}} \\ & + \sum_{k\in\mathbb{Z}} \sum_{l\in\mathbb{Z}} \left\langle C^X_{k-l,(0,1)}\left(\Phi^0_{2k-1,m}\right), \Phi^0_{2l,m}\right\rangle_{\mathcal{H}} \\ & + \sum_{k\in\mathbb{Z}} \sum_{l\in\mathbb{Z}} \left\langle C^X_{k-l,(1,0)}\left(\Phi^0_{2k,m}\right), \Phi^0_{2l-1,m}\right\rangle_{\mathcal{H}} \\ & + \sum_{k\in\mathbb{Z}} \sum_{l\in\mathbb{Z}} \left\langle C^X_{k-l,(1,1)}\left(\Phi^0_{2k-1,m}\right), \Phi^0_{2l-1,m}\right\rangle_{\mathcal{H}}, \quad t \overset{2}{\equiv} 0 \end{aligned} \tag{3.11}$$

and

$$\begin{aligned} E\left\|Y_{t,m}\right\|^2 = & \sum_{k\in\mathbb{Z}} \sum_{l\in\mathbb{Z}} \left\langle C^X_{k-l,(1,1)}\left(\Phi^1_{2k,m}\right), \Phi^1_{2l,m}\right\rangle_{\mathcal{H}} \\ & + \sum_{k\in\mathbb{Z}} \sum_{l\in\mathbb{Z}} \left\langle C^X_{k-l,(1,0)}\left(\Phi^1_{2k+1,m}\right), \Phi^1_{2l,m}\right\rangle_{\mathcal{H}} \\ & + \sum_{k\in\mathbb{Z}} \sum_{l\in\mathbb{Z}} \left\langle C^X_{k-l,(0,1)}\left(\Phi^1_{2k,m}\right), \Phi^1_{2l+1,m}\right\rangle_{\mathcal{H}} \\ & + \sum_{k\in\mathbb{Z}} \sum_{l\in\mathbb{Z}} \left\langle C^X_{k-l,(0,0)}\left(\Phi^1_{2k+1,m}\right), \Phi^1_{2l+1,m}\right\rangle_{\mathcal{H}}, \quad t \overset{2}{\equiv} 1. \end{aligned} \tag{3.12}$$

The long-run covariance matrix is given by

$$\lim_{n\to\infty} \frac{1}{n} \mathrm{Var}\left(\mathbf{Y}_1 + \cdots + \mathbf{Y}_n\right) = \frac{2\pi}{2} \left[\mathrm{diag}\left(\lambda_{0,1}, \ldots, \lambda_{0,p}\right) + \mathrm{diag}\left(\lambda_{0,p+1}, \ldots, \lambda_{0,2p}\right)\right].$$

The following theorem provides a formula for reconstructing the original $\mathcal{H}$-valued process X from its DFPC scores $\left\{Y_{t,m}, t \in \mathbb{Z}, m \geq 1\right\}$.

 Copyright © 2018 John Wiley & Sons Ltd *J. Time Ser. Anal.* **39**: 502–522 (2018)
DOI: 10.1111/jtsa.12283

Theorem 3.2. (Inversion Formula) Let $X = \left\{X_t, t \in \mathbb{Z}\right\}$ be an $\mathcal{H}$-valued, mean-zero T-PC random process, and $\left\{Y_{t,m}, t \in \mathbb{Z}, m \geq 1\right\}$ be its DFPC scores. For each time t and positive integer m, define $X_{t,m}$ to be

$$\begin{aligned} X_{t,m} \ : &= \sum_{l\in\mathbb{Z}} Y_{t+lT-d,m}\Phi^0_{lT-d,m} + \sum_{l\in\mathbb{Z}} Y_{t+lT-d+1,m}\Phi^1_{lT-d+1,m} \\ &+ \cdots + \sum_{l\in\mathbb{Z}} Y_{t+lT-d+T-1,m}\Phi^{T-1}_{lT-d+T-1,m}, \quad t \stackrel{T}{\equiv} d. \end{aligned}$$

Then

$$X_t = \sum_{m\geq 1} X_{t,m}, \quad t \stackrel{T}{\equiv} d,$$

where the convergence holds in mean square provided that

$$\sum_{d=0}^{T-1} \sum_{l\in\mathbb{Z}} \left\| \Phi^d_{l,m} \right\|_{\mathcal{H}} < \infty. \tag{3.13}$$

If $T = 2$, then

$$\begin{aligned} X_{t,m} &:= \sum_{l\in\mathbb{Z}} Y_{t+2l,m}\Phi^0_{2l,m} + \sum_{l\in\mathbb{Z}} Y_{t+2l+1,m}\Phi^1_{2l+1,m}, \quad t \stackrel{2}{\equiv} 0 \\ X_{t,m} &:= \sum_{l\in\mathbb{Z}} Y_{t+2l-1,m}\Phi^0_{2l-1,m} + \sum_{l\in\mathbb{Z}} Y_{t+2l,m}\Phi^1_{2l,m}, \quad t \stackrel{2}{\equiv} 1. \end{aligned}$$

The following theorem establishes an optimality property of the above DFPC filter based on a mean-square distance criterion:

Theorem 3.3. (Optimality) Let $X = \left\{X_t, t \in \mathbb{Z}\right\}$ be an $\mathcal{H}$-valued mean zero T-PC random process, and $\left\{X_{t,m}, t \in \mathbb{Z}, m \geq 1\right\}$ be as in Theorem 3.2.

For arbitrary $\mathcal{H}$-valued sequences

$$\left\{\Psi^t_{l,m}, t = 0, \ldots, T-1, m \geq 1, l \in \mathbb{Z}\right\} \text{ and } \left\{\Upsilon^t_{l,m}, t = 0, \ldots, T-1, m \geq 1, l \in \mathbb{Z}\right\}$$

with $\sum_{t=0}^{T-1}\sum_{l\in\mathbb{Z}} \left\|\Psi^t_{l,m}\right\| < \infty$ and $\sum_{t=0}^{T-1}\sum_{l\in\mathbb{Z}} \left\|\Upsilon^t_{l,m}\right\| < \infty$, for each m, consider

$$\begin{aligned} \check{Y}_{t,m} &= \sum_{l\in\mathbb{Z}} \left\langle X_{(t-l)}, \Psi^d_{l,m} \right\rangle \\ &= \sum_{l\in\mathbb{Z}} \left\langle X_{(t-lT-d)}, \Psi^d_{lT+d,m} \right\rangle + \sum_{l\in\mathbb{Z}} \left\langle X_{(t-lT-d+1)}, \Psi^d_{lT+d-1,m} \right\rangle \\ &+ \cdots + \sum_{l\in\mathbb{Z}} \left\langle X_{(t-lT-d+T-1)}, \Psi^d_{lT+d-T+1,m} \right\rangle, \quad t \stackrel{T}{\equiv} d \end{aligned}$$

 Copyright © 2018 John Wiley & Sons Ltd
DOI: 10.1111/jtsa.12283

and

$$\begin{aligned}\check{X}_{t,m} = &\sum_{l\in\mathbb{Z}} \check{Y}_{t+lT-d,m}\Upsilon^{0}_{lT-d,m} + \sum_{l\in\mathbb{Z}} \check{Y}_{t+lT-d+1,m}\Upsilon^{1}_{lT-d+1,m} \\ &+ \cdots + \sum_{l\in\mathbb{Z}} \check{Y}_{t+lT-d+T-1,m}\Upsilon^{T-1}_{lT-d+T-1,m}, \quad t \stackrel{T}{\equiv} d.\end{aligned}$$

Then the following inequality holds for each $t \in \mathbb{Z}$ and $p \geq 1$:

$$\begin{aligned}&E\left\| X_{Tt} - \sum_{m=1}^{p} X_{Tt,m}\right\|^2 + \cdots + E\left\| X_{Tt+T-1} - \sum_{m=1}^{p} X_{Tt+T-1,m}\right\|^2 \\ &= \sum_{m>p}\int_{-\pi}^{\pi} \lambda_{\theta,m}\,\mathrm{d}\theta \\ &\leq E\left\| X_{Tt} - \sum_{m=1}^{p} \check{X}_{Tt,m}\right\|^2 + \cdots + E\left\| X_{Tt+T-1} - \sum_{m=1}^{p} \check{X}_{Tt+T-1,m}\right\|^2.\end{aligned}$$

In practice, the scores $Y_{t,m}$ are estimated by truncated sums of the form

$$\begin{aligned}\widehat{Y}_{t,m} &= \sum_{l=-LT+d-T+1}^{LT+d} \left\langle X_{t-l}, \widehat{\Phi}^{d}_{l,m}\right\rangle \\ &= \sum_{l=-L}^{L} \left\langle X_{t-lT-d}, \widehat{\Phi}^{d}_{lT+d,m}\right\rangle + \cdots \\ &\quad + \sum_{l=-L}^{L} \left\langle X_{t-lT-d+T-1}, \widehat{\Phi}^{d}_{lT+d-T+1,m}\right\rangle,\ t \stackrel{T}{\equiv} d,\end{aligned} \tag{3.14}$$

in which $\widehat{\Phi}^{d}_{l,m}$ s are obtained from an estimator $\widehat{\mathcal{F}}^{X}_{\theta,(q,r)}$. In an asymptotic setting, the truncation level L is treated as an increasing function of n (the length of the time series). (Recommendations for the selection of L in finite samples are discussed in Sections 4 and 5.)

We conclude this section by showing in Theorem 3.4 that under mild assumptions, $\widehat{Y}_{t,m} - Y_{t,m}$ converges to zero in probability at a rate that depends on the rate of the estimation in the following condition:

Condition 3.1. The estimator $\widehat{\mathcal{F}}^{X}_{\theta,(q,r)}$ satisfies

$$\int_{-\pi}^{\pi} E\left\| \mathcal{F}^{X}_{\theta,(q,r)} - \widehat{\mathcal{F}}^{X}_{\theta,(q,r)}\right\|^2_{\mathcal{S}} \mathrm{d}\theta \to 0, \text{ as } n \to \infty; \quad q, r = 0, \ldots, T-1.$$

In Supporting Information, we show that Condition 3.1 holds for the specific estimators we recommend. The next condition is stronger than an analogous condition in Hörmann *et al.* (2015), but it is needed to establish a rate of convergence rather than merely convergence to zero.

Condition 3.2. Let $\lambda_{\theta,m}$ be as in (3.4) and define $\alpha_{\theta,1} := \lambda_{\theta,1} - \lambda_{\theta,2}$ and $\alpha_{\theta,m} := \min\left\{\lambda_{\theta,m-1} - \lambda_{\theta,m}, \lambda_{\theta,m} - \lambda_{\theta,m+1}\right\}$, $m > 1$. Assume that $\inf_{\theta\in(-\pi,\pi]} \alpha_{\theta,m} =: \alpha_{\star}(m) > 0$.

 Copyright © 2018 John Wiley & Sons Ltd
DOI: 10.1111/jtsa.12283

Condition 3.3. Set

$$\hat{c}_{\theta,m} := \frac{\langle \varphi_{\theta,m}, \hat{\varphi}_{\theta,m} \rangle_{\mathcal{H}^2}}{\left| \langle \varphi_{\theta,m}, \hat{\varphi}_{\theta,m} \rangle_{\mathcal{H}^2} \right|}$$

with $\varphi_{\theta,m}$ defined in (3.4). We assume that $\hat{c}_{\theta,m} > 0$.

Condition 3.3 merely sets a consistent orientation of the eigenfunctions.

To specify the rate of convergence, we introduce the following functions:

$$G(n) = \int_{-\pi}^{\pi} \left(\sum_{j,j'=0}^{T-1} \left\| \mathcal{F}^X_{\theta,(j,j')} - \hat{\mathcal{F}}^X_{\theta,(j,j')} \right\|_S^2 \right)^{\frac{1}{2}} \mathrm{d}\theta$$

and

$$H(L) = \sum_{d=0}^{T-1} \sum_{j=0}^{T-1} \left(\sum_{k \in \mathbb{Z}} \left\| \Phi^d_{Tk+d-j,m} \right\|^2 I\{|k| > L\} \right)^{\frac{1}{4}}.$$

Theorem 3.4. If Conditions 3.1–3.3 hold, then, for each m, as $n \to \infty$ and $L = L(n) \to \infty$ with n,

$$\hat{Y}_{t,m} - Y_{t,m} = O_P(G(n)L + H(L)).$$

In particular, $\hat{Y}_{t,m} - Y_{t,m}$ tends to zero in probability if L diverges so slowly that $G(n)L \to 0$, in probability.

The proof of Theorem 3.4 is given in Supporting Information. Using Proposition A.2, one can, under additional assumptions, easily derive a bound on $EG(n)$.

4. NUMERICAL IMPLEMENTATION

The theory presented in Section 3 is developed in the framework of infinite-dimensional Hilbert spaces in which the various functional objects live. Practically usable methodology requires a number of dimension reduction steps to create approximating finite-dimensional objects, which can be manipulated by computers. This section describes the main steps of such a reduction. Complexity and computing time are discussed in Supporting Information. We developed an `R` package, `pcdpca`, which allows us to preform all procedures described in this paper. In particular, it is used to perform the analysis and simulations in Section 5.

For a PC FTS $X = \{X_t, t \in \mathbb{Z}\}$, we operate on its projections on the first K elements of a functional basis. The main idea is to approximate a matrix of operators $[\mathcal{F}^X_{\theta,(d,f)}, \ 0 \le d,f \le T-1]$ in Theorem 3.1, by its equivalent in the projected space, which is a $TK \times TK$ matrix that can be estimated by applying methodology of Hörmann *et al.* (2015) to a series $\tilde{X}_t := (X_{Tt}, X_{Tt+1}, ..., X_{Tt+T-1})$. Next, we eigen-decompose the resulting spectral density. The eigenvector $dK + m$ corresponds to (d, m)th DFPC filter, for $1 \le m \le K$ and $0 \le d \le T-1$. By filtering $\tilde{X}_t$ with (d, m)th filters, we obtain the multivariate process $\tilde{\mathbf{Y}}_t := (\mathbf{Y}_{Tt}, \mathbf{Y}_{Tt+1}, ..., \mathbf{Y}_{Tt+T-1})$, from which we derive a process of scores $\mathbf{Y} = \{\mathbf{Y}_t, t \in \mathbb{Z}\}$. Below, we describe the details.

We use linearly independent basis functions $\{B_1, B_2, \dots, B_K\}$ to convert the data observed at discrete time points to functional objects of the form $x(u) = \sum_{j=1}^K c_j B_j(u)$. This is just the usual basis expansion step, see e.g. Chapter 3 of Ramsay *et al.* (2009) or Chapter 1 of Kokoszka and Reimherr (2017). We thus work in a finite-dimensional space $\mathcal{H}_K = \mathrm{sp}\{B_1, B_2, \dots, B_K\}$. To each bounded linear operator $A : \mathcal{H}_K \to \mathcal{H}_K$, there corresponds a complex-valued

$K \times K$ matrix $\mathfrak{A}$ defined by the relation $A(x) = \boldsymbol{B}'\mathfrak{A}\mathbf{c}$, where $\boldsymbol{B} = \left(B_1, B_2, \ldots, B_K\right)'$ and $\mathbf{c} = \left(c_1, c_2, \ldots, c_K\right)'$. Let $\mathbf{M}_B$ be the complex-valued $K \times K$ matrix $\left(\left\langle B_q, B_r \right\rangle\right)_{q,r=0,\ldots,K}$, $X_t = \boldsymbol{B}'\mathbf{c}_t$, and define

$$\mathbf{B}_T' := \left(\begin{pmatrix} B_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} B_K \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ \vdots \\ 0 \\ B_1 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ \vdots \\ 0 \\ B_K \end{pmatrix} \right) = \left(\mathbf{b}_1', \mathbf{b}_2', \ldots, \mathbf{b}_T'\right).$$

Next, define the matrix

$$\mathfrak{F}_\theta^{\tilde{X}} = \begin{pmatrix} \mathcal{F}^{\mathbf{c}}_{\theta,(0,0)} & \cdots & \mathcal{F}^{\mathbf{c}}_{\theta,(0,T-1)} \\ \vdots & \ddots & \vdots \\ \mathcal{F}^{\mathbf{c}}_{\theta,(T-1,0)} & \cdots & \mathcal{F}^{\mathbf{c}}_{\theta,(T-1,T-1)} \end{pmatrix} \begin{pmatrix} \mathbf{M}_B' & & 0 \\ & \ddots & \\ 0 & & \mathbf{M}_B' \end{pmatrix} \tag{4.1}$$

as the matrix corresponding to the operator $\mathcal{F}_\theta^{\tilde{X}}$, where all random functions are restricted to the subspace $\mathcal{H}_K^T$. Recall the definition of the spectral density operators $\mathcal{F}^{\mathbf{c}}_{\theta,(q,r)}$ corresponding to T-PC sequence $\mathbf{c} = \left\{\mathbf{c}_t, t \in \mathbb{Z}\right\}$ from (2.3)

If $\lambda_{\theta,m}$ and $\boldsymbol{\varphi}_{\theta,m} := \left(\boldsymbol{\varphi}'_{\theta,m,1}, \ldots, \boldsymbol{\varphi}'_{\theta,m,T}\right)'$ are the mth eigenvalue and eigenfunction of $TK \times TK$ complex-valued matrix $\mathfrak{F}_\theta^{\tilde{X}}$, then $\lambda_{\theta,m}$ and

$$\mathbf{B}_T'\boldsymbol{\varphi}_{\theta,m} = \left(\mathbf{b}_1', \mathbf{b}_2', \ldots, \mathbf{b}_T'\right)\left(\boldsymbol{\varphi}'_{\theta,m,1}, \ldots, \boldsymbol{\varphi}'_{\theta,m,T}\right)' = \left(\boldsymbol{B}'\,\boldsymbol{\varphi}_{\theta,m,1}, \ldots, \boldsymbol{B}'\boldsymbol{\varphi}_{\theta,m,T}\right)'$$

are the mth eigenvalue and eigenfunction of the operator $\mathcal{F}_\theta^{\tilde{X}}$. This motivates us to use the ordinary multivariate techniques to estimate $C^{\mathbf{c}}_{h,(q,r)}$, and consequently $\mathcal{F}^{\mathbf{c}}_{\theta,(q,r)}$, for $q, r = 0, \ldots, T-1$, and $\theta \in (-\pi, \pi]$, by

$$\widehat{C}^{\mathbf{c}}_{h,(q,r)} = \frac{T}{n} \sum_{j \in \mathbb{Z}} \mathbf{c}_{q+Tj}\mathbf{c}'_{r+Tj-Th} I\left\{1 \leq q + Tj \leq n\right\} I\left\{1 \leq r + Tj - Th \leq n\right\}, \quad h \geq 0$$

$$\left(\overline{\widehat{C}}^{\mathbf{c}}_{-h,(q,r)}\right)' = \widehat{C}^{\mathbf{c}}_{h,(q,r)}, \quad h < 0$$

and

$$\widehat{\mathcal{F}}^{\mathbf{c}}_{\theta,(q,r)} = \frac{1}{2\pi} \sum_{|h| \leq q(n)} w\left(\frac{h}{q(n)}\right) \widehat{C}^{\mathbf{c}}_{h,(q,r)} \mathrm{e}^{-\mathrm{i}h\theta}, \tag{4.2}$$

where w is a suitable weight function, $q(n) \to \infty$, and $q(n) = o(n)$. By substituting $\mathcal{F}^{\mathbf{c}}_{\theta,(q,r)}$ in (4.1) with its sample estimator $\widehat{\mathcal{F}}^{\mathbf{c}}_{\theta,(q,r)}$, we obtain an estimator $\widehat{\mathfrak{F}}_\theta^{\tilde{X}}$. Subsequently, we eigen-decompose $\widehat{\mathfrak{F}}_\theta^{\tilde{X}}$ to its eigenvalues $\widehat{\lambda}_{\theta,m}$ and eigenvectors $\widehat{\boldsymbol{\varphi}}_{\theta,m} := \left(\widehat{\boldsymbol{\varphi}}'_{\theta,m,1}, \ldots, \widehat{\boldsymbol{\varphi}}'_{\theta,m,T}\right)'$, $m \geq 1$. We use

$$\mathbf{B}_T'\widehat{\boldsymbol{\varphi}}_{\theta,m} = \left(\mathbf{b}_1', \ldots, \mathbf{b}_T'\right)\begin{pmatrix} \widehat{\boldsymbol{\varphi}}_{\theta,m,1} \\ \vdots \\ \widehat{\boldsymbol{\varphi}}_{\theta,m,T} \end{pmatrix} = \begin{pmatrix} \boldsymbol{B}'\widehat{\boldsymbol{\varphi}}_{\theta,m,1} \\ \vdots \\ \boldsymbol{B}'\widehat{\boldsymbol{\varphi}}_{\theta,m,T} \end{pmatrix} = \begin{pmatrix} \widehat{\varphi}_{\theta,m,1} \\ \vdots \\ \widehat{\varphi}_{\theta,m,T} \end{pmatrix} = \widehat{\varphi}_{\theta,m}$$

 Copyright © 2018 John Wiley & Sons Ltd J. Time Ser. Anal. **39**: 502–522 (2018)
DOI: 10.1111/jtsa.12283

to get estimators $\widehat{\varphi}_{\theta,m}$, and set

$$\begin{pmatrix} \boldsymbol{B}'\widehat{\boldsymbol{\varphi}}_{\theta,dp+m,1} \\ \vdots \\ \boldsymbol{B}'\widehat{\boldsymbol{\varphi}}_{\theta,dp+m,T} \end{pmatrix} = \begin{pmatrix} \widehat{\Phi}^{d}_{\theta,d,m} \\ \vdots \\ \widehat{\Phi}^{d}_{\theta,d-T+1,m} \end{pmatrix},$$

or equivalently

$$\begin{pmatrix} \boldsymbol{B}'\widehat{\boldsymbol{\Phi}}^{d}_{lT+d,m} \\ \vdots \\ \boldsymbol{B}'\widehat{\boldsymbol{\Phi}}^{d}_{lT+d-T+1,m} \end{pmatrix} := \frac{1}{2\pi}\int_{-\pi}^{\pi} \begin{pmatrix} \boldsymbol{B}'\widehat{\boldsymbol{\varphi}}_{\theta,dp+m,1} \\ \vdots \\ \boldsymbol{B}'\widehat{\boldsymbol{\varphi}}_{\theta,dp+m,T} \end{pmatrix} \mathrm{e}^{-\mathrm{i}l\theta}\,\mathrm{d}\theta = \begin{pmatrix} \widehat{\Phi}^{d}_{lT+d,m} \\ \vdots \\ \widehat{\Phi}^{d}_{lT+d-T+1,m} \end{pmatrix}, \tag{4.3}$$

for $d = 0, \dots, T-1$, $m = 1, \dots, p$. Note that one may use numerical integration to find $\widehat{\Phi}^{d}_{l,m}$. Therefore, the PC-DFPC scores can be estimated by

$$\begin{aligned} \widehat{Y}_{t,m} &= \sum_{l=-LT+d-T+1}^{LT+d} \left\langle X_{t-l}, \widehat{\Phi}^{d}_{l,m} \right\rangle \\ &= \sum_{l=-L}^{L} \left\langle X_{t-lT-d}, \widehat{\Phi}^{d}_{lT+d,m} \right\rangle + \cdots + \sum_{l=-L}^{L} \left\langle X_{t-lT-d+T-1}, \widehat{\Phi}^{d}_{lT+d-T+1,m} \right\rangle \\ &= \sum_{l=-L}^{L} \mathbf{c}'_{(t-lT-d)} \mathbf{M}_B \overline{\widehat{\boldsymbol{\Phi}}}^{d}_{lT+d,m} + \dots + \sum_{l=-L}^{L} \mathbf{c}'_{(t-lT-d+T-1)} \mathbf{M}_B \overline{\widehat{\boldsymbol{\Phi}}}^{d}_{lT+d-T+1,m}, \quad t \overset{T}{\equiv} d, \end{aligned}$$

where L satisfies

$$\sum_{l=-L}^{L} \left(\left\| \widehat{\Phi}^{d}_{lT+d,m} \right\|^2_{\mathcal{H}} + \cdots + \left\| \widehat{\Phi}^{d}_{lT+d-T+1,m} \right\|^2_{\mathcal{H}} \right) \geq 1 - \varepsilon$$

for some d and small $\varepsilon > 0$. (We use $\varepsilon = 0.01$ in our computations.) Consequently, X_t can be approximated by

$$\widehat{X}_t = \sum_{m=1}^{p} \sum_{l=-L}^{L} \widehat{Y}_{t+lT-d,m} \widehat{\Phi}^{0}_{lT-d,m} + \dots + \sum_{m=1}^{p} \sum_{l=-L}^{L} \widehat{Y}_{t+lT-d+T-1,m} \widehat{\Phi}^{T-1}_{lT-d+T-1,m}, \quad t \overset{T}{\equiv} d.$$

To facilitate understanding, Supporting Information contains explicit formulas in the case of $T = 2$.

Remark 4.1. Usually, the mean μ of the process X is not known. In this case, we first use smoothed functions $X_t = \boldsymbol{B}'\mathbf{c}_t$ to obtain estimators $\widehat{\mu}_0, \dots, \widehat{\mu}_{T-1}$ or T-periodic mean function estimator $\left\{ \widehat{\mu}_t : \widehat{\mu}_{Tk+d} = \widehat{\mu}_d; k \in \mathbb{Z} \right\}$, and then apply the above method to the centered functional observations $X^*_t = X_t - \widehat{\mu}_t$.

5. APPLICATION TO PARTICULATE POLLUTION DATA

To illustrate the advantages of PC-DFPCA relative to the (stationary) DFPCA, which may arise in certain settings, we further explore the dataset analyzed in Hörmann *et al.* (2015). The dataset contains intraday measurements of pollution in Graz, Austria, between October 1, 2010 and March 31, 2011. Observations were sampled every 30 minutes and the concentration of particulate matter of diameter of less than $10\mu m$ was measured in the ambient air. In order to facilitate the comparison with the results reported in Hörmann *et al.* (2015), we employ exactly the same preprocessing procedure, including square-root transformation and removal of the mean weekly pattern

 Copyright © 2018 John Wiley & Sons Ltd
DOI: 10.1111/jtsa.12283

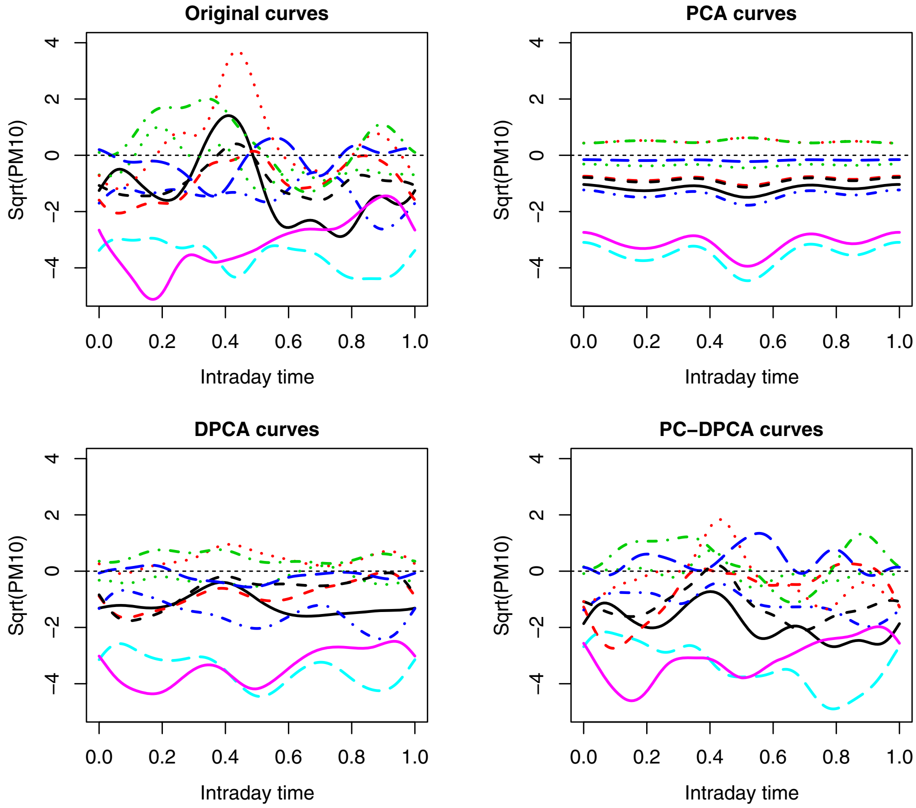

Figure 1. Ten successive intraday observations of PM10 data (top-left), the corresponding functional PCA curves reconstructed from the first principal component (top-right), dynamic functional principal component curves (bottom-left), and periodically correlated dynamic principal components (bottom-right). Colors and types of curves match the same observations among plots [Color figure can be viewed at wileyonlinelibrary.com]

and outliers. Note that the removal of the mean weekly pattern does not affect periodic covariances between weekdays, and therefore they can be exploited using the PC-DFPCA procedure applied with the period $T = 7$. The preprocessed dataset contains 175 daily curves, each converted to a functional object with 15 Fourier basis functions, yielding a FTS $\{X_t : 1 \le t \le 175\}$.

For FPCA and DFPCA, we use the same procedure as Hörmann *et al.* (2015) using the implementation and data published by those researchers as the `R` package `freqdom`. To implement the PC-DFPCA, some modifications are needed. Regarding the metaparameters q and L, we recommend $q = n^{1/3}$ and $L = 3$ for the series consisting of several hundred observations. The simulations reported in Supporting Information show that choosing a larger L does not improve the quality of prediction.

As a measure of fit, we use the normalized mean-squared error (`NMSE`), defined as

$$\texttt{NMSE} = \sum_{t=1}^{n} \|X_t - \hat{X}_t\|^2 / \sum_{t=1}^{n} \|X_t\|^2,$$

where the $\hat{X}_t$ are the observations obtained from the inverse transform. We refer to the value $\texttt{NMSE} \cdot 100$ as the *percentage of variance explained*.

 Copyright © 2018 John Wiley & Sons Ltd
DOI: 10.1111/jtsa.12283

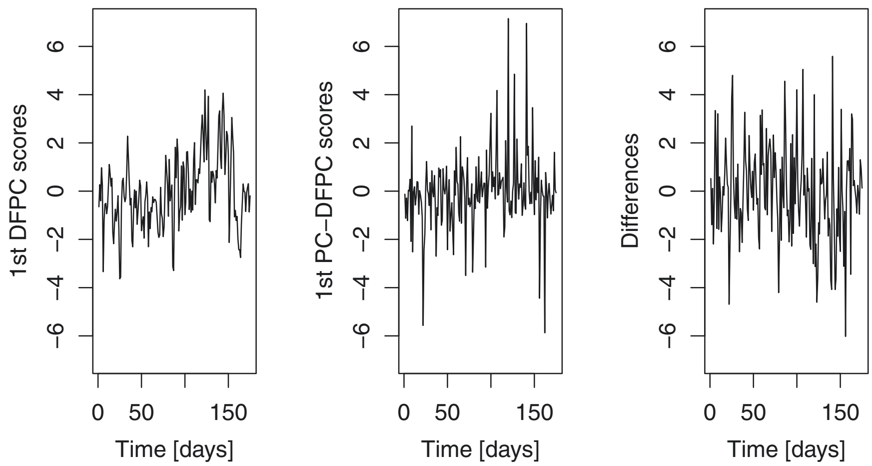

Figure 2. The first dynamic principal component scores (left), the first periodically correlated dynamic principal component scores (middle), and differences between the two series (right)

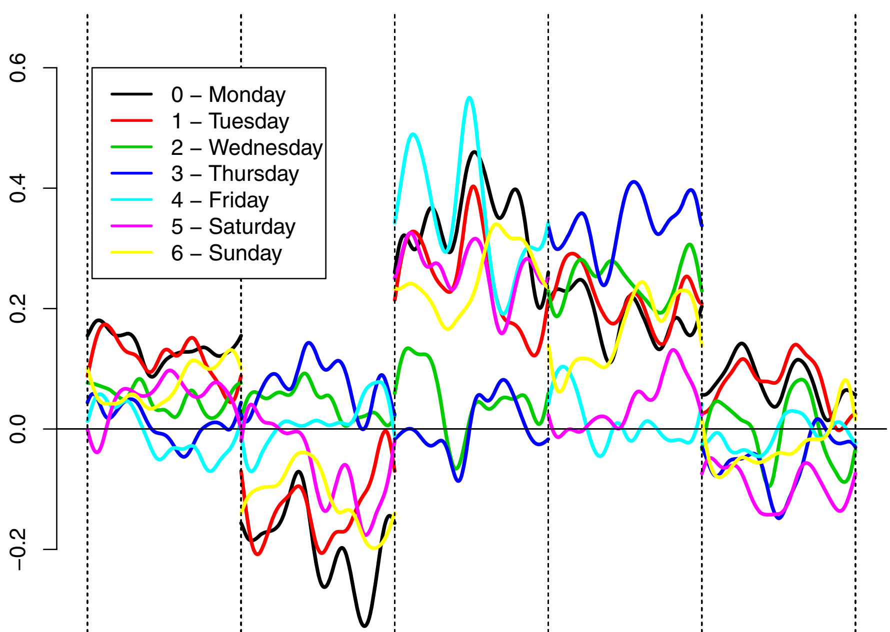

Figure 3. Filters of the first principal component, i.e. $m = 1$. For every $d = 0, 1, ..., 6$ (as in the legend), the five curves of the same color correspond to $\Phi^d_{-2,m}, \Phi^d_{-1,m}, ..., \Phi^d_{2,m}$ [Color figure can be viewed at wileyonlinelibrary.com]

For the sake of comparison and discussion, we focus only on the first principal component, which already explains 73 of variability in the static FPCA, 80 of variability in the DFPCA, and 88 of variability in the PC-DFPCA procedure. Curves corresponding to the components obtained through each of these methods are presented in Figure 1. As the percentages above suggest, there is a clear progression in the quality of the approximation using just one component. This is an important finding because the purpose of the principal component analysis of any type is to reduce the dimension of the data using the smallest possible number of projections without sacrificing the quality of approximation.

Hörmann *et al.* (2015) observed that, for this particular dataset, the sequences of scores of the DFPCs and the static FPCs were almost identical. This is no longer the case if the PC-DFPCs are used. Figure 2 compares the DFPC and the PC-DFPC scores and shows that the resulting time series are quite different. PC-DFPCA takes into account the periodic correlation structure, which neither the static nor the (stationary) DFPCA does.

DOI: 10.1111/jtsa.12283
Copyright © 2018 John Wiley & Sons Ltd

The estimated PC-DFPCA filters are very high dimensional as can be seen in Figure 3. In particular, with $L = 3$, $T = 7$, and $p = 15$, we estimated $(2L + 1)T^2p^2 = 735$ real numbers, which may raise concerns about overfitting. This, however, does not translate into problems with the finite sample performance, as the following simulation study shows.

6. PROOFS OF THE RESULTS OF SECTION 3

To explain the essence and technique of the proofs, we consider the special case of $T = 2$. The arguments for general T proceed analogously, merely with a more heavy and less explicit notation, which may obscure the essential arguments.

Proof of Theorem 3.1. To establish the mean-square convergence of the series $\sum_{l\in\mathbb{Z}} \Psi_l^t\left(X_{(t-l)}\right)$, let $S_{n,t}$ be its partial sum, i.e. $S_{n,t} = \sum_{-n\le l\le n} \Psi_l^t\left(X_{(t-l)}\right)$, for each positive integer n. Then for $m < n$, we have

$$\begin{aligned}
E\left\|S_{n,t} - S_{m,t}\right\|^2_{\mathbb{C}^p} &= \sum_{m<|l|,|k|\le n} E\left\langle \Psi_l^t\left(X_{(t-l)}\right), \Psi_k^t\left(X_{(t-k)}\right)\right\rangle_{\mathbb{C}^p} \\
&\le \sum_{m<|l|,|k|\le n} E\left(\left\|\Psi_l^t\left(X_{(t-l)}\right)\right\|_{\mathbb{C}^p}\left\|\Psi_k^t\left(X_{(t-k)}\right)\right\|_{\mathbb{C}^p}\right) \\
&\le \sum_{m<|l|,|k|\le n} \left\|\Psi_l^t\right\|_{\mathcal{L}}\left\|\Psi_k^t\right\|_{\mathcal{L}} E\left(\left\|X_{(t-l)}\right\|\left\|X_{(t-k)}\right\|\right) \\
&\le \sum_{m<|l|,|k|\le n} \left\|\Psi_l^t\right\|_{\mathcal{L}}\left\|\Psi_k^t\right\|_{\mathcal{L}} \left(E\left\|X_{(t-l)}\right\|^2 E\left\|X_{(t-k)}\right\|^2\right)^{\frac{1}{2}} \\
&\le M\sum_{|l|>m}\sum_{|k|>m} \left\|\Psi_l^t\right\|_{\mathcal{L}}\left\|\Psi_k^t\right\|_{\mathcal{L}} \quad \text{for some } M\in\mathbb{R}^+ \\
&\le M\left(\sum_{|l|>m}\left\|\Psi_l^t\right\|_{\mathcal{L}}\right)^2. \qquad (6.1)
\end{aligned}$$

Summability condition (2.5) implies that (6.1) tends to zero as n and m tend to infinity. Therefore, $\left\{S_{n,t}, n\in\mathbb{Z}^+\right\}$ forms a Cauchy sequence in $L^2(\mathbb{C}^p, \Omega)$, for each t, which implies the desired mean-square convergence. According to the representation of the filtered process $\mathbf{Y}$ at time t, i.e.

$$\begin{aligned}
\mathbf{Y}_t &= \sum_{l\in\mathbb{Z}} \Psi_l^0\left(X_{(t-l)}\right) \\
&= \sum_{l\in\mathbb{Z}} \Psi_{2l}^0\left(X_{(t-2l)}\right) + \sum_{l\in\mathbb{Z}} \Psi_{2l-1}^0\left(X_{(t-2l+1)}\right), \quad t \overset{2}{\equiv} 0 \\
\mathbf{Y}_t &= \sum_{l\in\mathbb{Z}} \Psi_l^1\left(X_{(t-l)}\right) \\
&= \sum_{l\in\mathbb{Z}} \Psi_{2l}^1\left(X_{(t-2l)}\right) + \sum_{l\in\mathbb{Z}} \Psi_{2l+1}^1\left(X_{(t-2l-1)}\right), \quad t \overset{2}{\equiv} 1,
\end{aligned}$$

for each $h \in \mathbb{Z}$ we have

$$\begin{aligned}
\operatorname{Cov}\left(\mathbf{Y}_{2h}, \mathbf{Y}_0\right) &= \lim_{n\to\infty}\sum_{|k|\le n}\sum_{|l|\le n} \operatorname{Cov}\left(\Psi_k^0\left(X_{(2h-k)}\right), \Psi_l^0\left(X_{(0-l)}\right)\right) \\
&= \sum_{k\in\mathbb{Z}}\sum_{l\in\mathbb{Z}} \Psi_{2k}^0 \operatorname{Cov}\left(X_{(2h-2k)}, X_{-2l}\right)\left(\Psi_{2l}^0\right)^*
\end{aligned}$$

Copyright © 2018 John Wiley & Sons Ltd

DOI: 10.1111/jtsa.12283

$$
\begin{aligned}
&+\sum_{k\in\mathbb{Z}}\sum_{l\in\mathbb{Z}}\Psi^0_{2k}\mathrm{Cov}\left(X_{(2h-2k)},X_{-2l+1}\right)\left(\Psi^0_{2l-1}\right)^*\\
&+\sum_{k\in\mathbb{Z}}\sum_{l\in\mathbb{Z}}\Psi^0_{2k-1}\mathrm{Cov}\left(X_{(2h-2k+1)},X_{-2l}\right)\left(\Psi^0_{2l}\right)^*\\
&+\sum_{k\in\mathbb{Z}}\sum_{l\in\mathbb{Z}}\Psi^0_{2k-1}\mathrm{Cov}\left(X_{(2h-2k+1)},X_{-2l+1}\right)\left(\Psi^0_{2l-1}\right)^*.
\end{aligned}
$$

Consequently

$$
\begin{aligned}
\mathcal{F}^{\mathbf{Y}}_{\theta,(0,0)} &= \frac{1}{2\pi}\sum_{h\in\mathbb{Z}}\mathrm{Cov}\left(\mathbf{Y}_{2h},\mathbf{Y}_0\right)\mathrm{e}^{-\mathrm{i}h\theta}\\
&= \frac{1}{2\pi}\sum_{h\in\mathbb{Z}}\sum_{k\in\mathbb{Z}}\sum_{l\in\mathbb{Z}}\Psi^0_{2k}\mathrm{Cov}\left(X_{(2h-2k)},X_{-2l}\right)\left(\Psi^0_{2l}\right)^*\mathrm{e}^{-\mathrm{i}h\theta}\\
&\quad+\frac{1}{2\pi}\sum_{h\in\mathbb{Z}}\sum_{k\in\mathbb{Z}}\sum_{l\in\mathbb{Z}}\Psi^0_{2k}\mathrm{Cov}\left(X_{(2h-2k)},X_{-2l+1}\right)\left(\Psi^0_{2l-1}\right)^*\mathrm{e}^{-\mathrm{i}h\theta}\\
&\quad+\frac{1}{2\pi}\sum_{h\in\mathbb{Z}}\sum_{k\in\mathbb{Z}}\sum_{l\in\mathbb{Z}}\Psi^0_{2k-1}\mathrm{Cov}\left(X_{(2h-2k+1)},X_{-2l}\right)\left(\Psi^0_{2l}\right)^*\mathrm{e}^{-\mathrm{i}h\theta}\\
&\quad+\frac{1}{2\pi}\sum_{h\in\mathbb{Z}}\sum_{k\in\mathbb{Z}}\sum_{l\in\mathbb{Z}}\Psi^0_{2k-1}\mathrm{Cov}\left(X_{(2h-2k+1)},X_{-2l+1}\right)\left(\Psi^0_{2l-1}\right)^*\mathrm{e}^{-\mathrm{i}h\theta},
\end{aligned}
$$

which leads to

$$
\begin{aligned}
\mathcal{F}^{\mathbf{Y}}_{\theta,(0,0)} &= \frac{1}{2\pi}\sum_{k\in\mathbb{Z}}\sum_{l\in\mathbb{Z}}\sum_{h\in\mathbb{Z}}\Psi^0_{2k}\mathrm{Cov}\left(X_{(2h-2k+2l)},X_0\right)\left(\Psi^0_{2l}\right)^*\mathrm{e}^{-\mathrm{i}(h-k+l)\theta}\mathrm{e}^{il\theta}\mathrm{e}^{-\mathrm{i}k\theta}\\
&\quad+\frac{1}{2\pi}\sum_{k\in\mathbb{Z}}\sum_{l\in\mathbb{Z}}\sum_{h\in\mathbb{Z}}\Psi^0_{2k}\mathrm{Cov}\left(X_{(2h-2k+2l)},X_1\right)\left(\Psi^0_{2l-1}\right)^*\mathrm{e}^{-\mathrm{i}(h-k+l)\theta}\mathrm{e}^{il\theta}\mathrm{e}^{-\mathrm{i}k\theta}\\
&\quad+\frac{1}{2\pi}\sum_{k\in\mathbb{Z}}\sum_{l\in\mathbb{Z}}\sum_{h\in\mathbb{Z}}\Psi^0_{2k-1}\mathrm{Cov}\left(X_{(2h-2k+2l+1)},X_0\right)\left(\Psi^0_{2l}\right)^*\mathrm{e}^{-\mathrm{i}(h-k+l)\theta}\mathrm{e}^{il\theta}\mathrm{e}^{-\mathrm{i}k\theta}\\
&\quad+\frac{1}{2\pi}\sum_{k\in\mathbb{Z}}\sum_{l\in\mathbb{Z}}\sum_{h\in\mathbb{Z}}\Psi^0_{2k-1}\mathrm{Cov}\left(X_{(2h-2k+2l+1)},X_1\right)\left(\Psi^0_{2l-1}\right)^*\mathrm{e}^{-\mathrm{i}(h-k+l)\theta}\mathrm{e}^{il\theta}\mathrm{e}^{-\mathrm{i}k\theta},\\
&= \sum_{k\in\mathbb{Z}}\sum_{l\in\mathbb{Z}}\Psi^0_{2k}\mathcal{F}^X_{\theta,(0,0)}\left(\Psi^0_{2l}\right)^*\mathrm{e}^{il\theta}\mathrm{e}^{-\mathrm{i}k\theta}\\
&\quad+\sum_{k\in\mathbb{Z}}\sum_{l\in\mathbb{Z}}\Psi^0_{2k}\mathcal{F}^X_{\theta,(0,1)}\left(\Psi^0_{2l-1}\right)^*\mathrm{e}^{il\theta}\mathrm{e}^{-\mathrm{i}k\theta}\\
&\quad+\sum_{k\in\mathbb{Z}}\sum_{l\in\mathbb{Z}}\Psi^0_{2k-1}\mathcal{F}^X_{\theta,(1,0)}\left(\Psi^0_{2l}\right)^*\mathrm{e}^{il\theta}\mathrm{e}^{-\mathrm{i}k\theta}\\
&\quad+\sum_{k\in\mathbb{Z}}\sum_{l\in\mathbb{Z}}\Psi^0_{2k-1}\mathcal{F}^X_{\theta,(1,1)}\left(\Psi^0_{2l-1}\right)^*\mathrm{e}^{il\theta}\mathrm{e}^{-\mathrm{i}k\theta}\\
&=:\ \Psi^0_{\theta,0}\mathcal{F}^X_{\theta,(0,0)}\left(\Psi^0_{\theta,0}\right)^*\\
&\quad+\Psi^0_{\theta,0}\mathcal{F}^X_{\theta,(0,1)}\left(\Psi^0_{\theta,-1}\right)^*\\
&\quad+\Psi^0_{\theta,-1}\mathcal{F}^X_{\theta,(1,0)}\left(\Psi^0_{\theta,0}\right)^*\\
&\quad+\Psi^0_{\theta,-1}\mathcal{F}^X_{\theta,(1,1)}\left(\Psi^0_{\theta,-1}\right)^*.
\end{aligned}
$$

 Copyright © 2018 John Wiley & Sons Ltd
DOI: 10.1111/jtsa.12283

The operator $\mathcal{F}^{\mathbf{Y}}_{\theta,(0,0)}$ from $\mathbb{C}^p$ to $\mathbb{C}^p$ has the following matrix form:

$$\begin{aligned}
&\left(\left\langle \left(\mathcal{F}^{X}_{\theta,(0,0)}\right)^{*}\left(\Psi^{0}_{\theta,0,r}\right), \Psi^{0}_{\theta,0,q}\right\rangle_{\mathcal{H}}\right)_{p\times p} \\
&+\left(\left\langle \left(\mathcal{F}^{X}_{\theta,(0,1)}\right)^{*}\left(\Psi^{0}_{\theta,0,r}\right), \Psi^{0}_{\theta,-1,q}\right\rangle_{\mathcal{H}}\right)_{p\times p} \\
&+\left(\left\langle \left(\mathcal{F}^{X}_{\theta,(1,0)}\right)^{*}\left(\Psi^{0}_{\theta,-1,r}\right), \Psi^{0}_{\theta,0,q}\right\rangle_{\mathcal{H}}\right)_{p\times p} \\
&+\left(\left\langle \left(\mathcal{F}^{X}_{\theta,(1,1)}\right)^{*}\left(\Psi^{0}_{\theta,-1,r}\right), \Psi^{0}_{\theta,-1,q}\right\rangle_{\mathcal{H}}\right)_{p\times p}.
\end{aligned}$$

Finally

$$\begin{aligned}
\mathcal{F}^{\mathbf{Y}}_{\theta,(0,0)} &= \left\langle \begin{pmatrix} \left(\mathcal{F}^{X}_{\theta,(0,0)}\right)^{*} & \left(\mathcal{F}^{X}_{\theta,(1,0)}\right)^{*} \\ \left(\mathcal{F}^{X}_{\theta,(0,1)}\right)^{*} & \left(\mathcal{F}^{X}_{\theta,(1,1)}\right)^{*} \end{pmatrix} \begin{pmatrix} \Psi^{0}_{\theta,0,r} \\ \Psi^{0}_{\theta,-1,r} \end{pmatrix}, \begin{pmatrix} \Psi^{0}_{\theta,0,q} \\ \Psi^{0}_{\theta,-1,q} \end{pmatrix} \right\rangle_{\mathcal{H}^2} \\
&= \left\langle \begin{pmatrix} \mathcal{F}^{X}_{\theta,(0,0)} & \mathcal{F}^{X}_{\theta,(0,1)} \\ \mathcal{F}^{X}_{\theta,(1,0)} & \mathcal{F}^{X}_{\theta,(1,1)} \end{pmatrix} \begin{pmatrix} \Psi^{0}_{\theta,0,r} \\ \Psi^{0}_{\theta,-1,r} \end{pmatrix}, \begin{pmatrix} \Psi^{0}_{\theta,0,q} \\ \Psi^{0}_{\theta,-1,q} \end{pmatrix} \right\rangle_{\mathcal{H}^2}.
\end{aligned}$$

Using similar arguments leads to the following representations for $\mathcal{F}^{\mathbf{Y}}_{\theta,(1,0)}$, $\mathcal{F}^{\mathbf{Y}}_{\theta,(0,1)}$, and $\mathcal{F}^{\mathbf{Y}}_{\theta,(1,1)}$:

$$\mathcal{F}^{\mathbf{Y}}_{\theta,(1,0)} = \left\langle \begin{pmatrix} \mathcal{F}^{X}_{\theta,(0,0)} & \mathcal{F}^{X}_{\theta,(0,1)} \\ \mathcal{F}^{X}_{\theta,(1,0)} & \mathcal{F}^{X}_{\theta,(1,1)} \end{pmatrix} \begin{pmatrix} \Psi^{1}_{\theta,1,r} \\ \Psi^{1}_{\theta,0,r} \end{pmatrix}, \begin{pmatrix} \Psi^{0}_{\theta,0,q} \\ \Psi^{0}_{\theta,-1,q} \end{pmatrix} \right\rangle_{\mathcal{H}^2}$$

$$\mathcal{F}^{\mathbf{Y}}_{\theta,(0,1)} = \left\langle \begin{pmatrix} \mathcal{F}^{X}_{\theta,(0,0)} & \mathcal{F}^{X}_{\theta,(0,1)} \\ \mathcal{F}^{X}_{\theta,(1,0)} & \mathcal{F}^{X}_{\theta,(1,1)} \end{pmatrix} \begin{pmatrix} \Psi^{0}_{\theta,0,r} \\ \Psi^{0}_{\theta,-1,r} \end{pmatrix}, \begin{pmatrix} \Psi^{1}_{\theta,1,q} \\ \Psi^{1}_{\theta,0,q} \end{pmatrix} \right\rangle_{\mathcal{H}^2}$$

$$\mathcal{F}^{\mathbf{Y}}_{\theta,(1,1)} = \left\langle \begin{pmatrix} \mathcal{F}^{X}_{\theta,(0,0)} & \mathcal{F}^{X}_{\theta,(0,1)} \\ \mathcal{F}^{X}_{\theta,(1,0)} & \mathcal{F}^{X}_{\theta,(1,1)} \end{pmatrix} \begin{pmatrix} \Psi^{1}_{\theta,1,r} \\ \Psi^{1}_{\theta,0,r} \end{pmatrix}, \begin{pmatrix} \Psi^{1}_{\theta,1,q} \\ \Psi^{1}_{\theta,0,q} \end{pmatrix} \right\rangle_{\mathcal{H}^2}.$$

Note that the 2-periodic behavior of the covariance operators of the filtered process $\mathbf{Y}$ is an implicit result of the above argument, which completes the proof of Theorem 3.1 for the special case $T = 2$. The general case is similar.

Proof of Proposition 3.1. To establish part (a), consider the eigenvalue decomposition (3.4). We then have

$$\begin{aligned}
\lambda_{\theta,m}\varphi_{\theta,m} &= \begin{pmatrix} \mathcal{F}^{X}_{\theta,(0,0)} & \mathcal{F}^{X}_{\theta,(0,1)} \\ \mathcal{F}^{X}_{\theta,(1,0)} & \mathcal{F}^{X}_{\theta,(1,1)} \end{pmatrix} \left(\varphi_{\theta,m}\right) \\
&= \begin{pmatrix} \mathcal{F}^{X}_{\theta,(0,0)} & \mathcal{F}^{X}_{\theta,(0,1)} \\ \mathcal{F}^{X}_{\theta,(1,0)} & \mathcal{F}^{X}_{\theta,(1,1)} \end{pmatrix} \begin{pmatrix} \varphi_{\theta,m,1} \\ \varphi_{\theta,m,2} \end{pmatrix} \\
&= \begin{pmatrix} \mathcal{F}^{X}_{\theta,(0,0)}\left(\varphi_{\theta,m,1}\right) + \mathcal{F}^{X}_{\theta,(0,1)}\left(\varphi_{\theta,m,2}\right) \\ \mathcal{F}^{X}_{\theta,(1,0)}\left(\varphi_{\theta,m,1}\right) + \mathcal{F}^{X}_{\theta,(1,1)}\left(\varphi_{\theta,m,2}\right) \end{pmatrix} \\
&= \frac{1}{2\pi}\sum_{h\in\mathbb{Z}} \begin{pmatrix} E\left[\left\langle \varphi_{\theta,m,1}, X_0\right\rangle_{\mathcal{H}} + \left\langle \varphi_{\theta,m,2}, X_1\right\rangle_{\mathcal{H}}\right] X_{2h} \\ E\left[\left\langle \varphi_{\theta,m,1}, X_0\right\rangle_{\mathcal{H}} + \left\langle \varphi_{\theta,m,2}, X_1\right\rangle_{\mathcal{H}}\right] X_{2h+1} \end{pmatrix} \mathrm{e}^{-\mathrm{i}h\theta}.
\end{aligned}$$

 Copyright © 2018 John Wiley & Sons Ltd
DOI: 10.1111/jtsa.12283

Consequently

$$\lambda_{\theta,m}\overline{\varphi}_{\theta,m} = \frac{1}{2\pi}\sum_{h\in\mathbb{Z}}\begin{pmatrix} E\left[\overline{\langle\varphi_{\theta,m,1},X_0\rangle}_{\mathcal{H}}+\overline{\langle\varphi_{\theta,m,2},X_1\rangle}_{\mathcal{H}}\right]\overline{X}_{2h} \\ E\left[\overline{\langle\varphi_{\theta,m,1},X_0\rangle}_{\mathcal{H}}+\overline{\langle\varphi_{\theta,m,2},X_1\rangle}_{\mathcal{H}}\right]\overline{X}_{2h+1}\end{pmatrix}e^{+ih\theta}$$
$$\frac{1}{2\pi}\sum_{h\in\mathbb{Z}}\begin{pmatrix} E\left[\langle\overline{\varphi}_{\theta,m,1},X_0\rangle_{\mathcal{H}}+\langle\overline{\varphi}_{\theta,m,2},X_1\rangle_{\mathcal{H}}\right]X_{2h} \\ E\left[\langle\overline{\varphi}_{\theta,m,1},X_0\rangle_{\mathcal{H}}+\langle\overline{\varphi}_{\theta,m,2},X_1\rangle_{\mathcal{H}}\right]X_{2h+1}\end{pmatrix}e^{+ih\theta}$$
$$= \begin{pmatrix}\mathcal{F}^X_{-\theta,(0,0)} & \mathcal{F}^X_{-\theta,(0,1)} \\ \mathcal{F}^X_{-\theta,(1,0)} & \mathcal{F}^X_{-\theta,(1,1)}\end{pmatrix}\left(\overline{\varphi}_{\theta,m}\right).$$

Hence, $\lambda_{\theta,m}$ and $\overline{\varphi}_{\theta,m}$ are an eigenvalue/eigenfunction pair of $\begin{pmatrix}\mathcal{F}^X_{-\theta,(0,0)} & \mathcal{F}^X_{-\theta,(0,1)} \\ \mathcal{F}^X_{-\theta,(1,0)} & \mathcal{F}^X_{-\theta,(1,1)}\end{pmatrix}$. Now, use (3.8) to obtain $\Phi^t_{l,m} = \overline{\Phi}^t_{l,m}$, which implies that the DFPC scores $Y_{t,m}$ satisfy

$$Y_{t,m} = \sum_{l\in\mathbb{Z}}\left\langle X_{t-l},\Phi^t_{l,m}\right\rangle_{\mathcal{H}} = \sum_{l\in\mathbb{Z}}\left\langle \overline{X}_{t-l},\overline{\Phi}^t_{l,m}\right\rangle_{\mathcal{H}}$$
$$= \sum_{l\in\mathbb{Z}}\overline{\left\langle X_{t-l},\Phi^t_{l,m}\right\rangle_{\mathcal{H}}} = \overline{Y}_{t,m},$$

and so are real for each t and m.

For part (b), first we define $Y_{t,m,n} := \sum_{l=-n}^{n}\left\langle X_{t-l},\Phi^t_{l,m}\right\rangle$. Then we use a similar argument as in the proof of Theorem 3.1 to show that $Y_{t,m,n}$ is converges in mean square to $Y_{t,m} = \sum_{l\in\mathbb{Z}}\left\langle X_{t-l},\Phi^t_{l,m}\right\rangle$. Hence

$$\left\|E\left(Y_{t,m,n}\otimes Y_{t,m,n}\right) - E\left(Y_{t,m}\otimes Y_{t,m}\right)\right\|_{\mathcal{S}} \to 0, \quad \text{as } n\to\infty$$

or equivalently

$$\left|E\left\|Y_{t,m,n}\right\|^2_{\mathbb{C}} - E\left\|Y_{t,m}\right\|^2_{\mathbb{C}}\right| \to 0, \quad \text{as } n\to\infty.$$

Consequently, for $t \overset{2}{\equiv} 0$, we have

$$E\left\|Y_{t,m}\right\|^2_{\mathbb{C}} = \lim_{n\to\infty} EY_{t,m,n}\overline{Y}_{t,m,n} = \lim_{n\to\infty}\sum_{|k|\le n}\sum_{|l|\le n} E\left\langle X_{t-l},\Phi^0_{l,m}\right\rangle\left\langle\Phi^0_{k,m},X_{t-k}\right\rangle$$
$$= \sum_{k\in\mathbb{Z}}\sum_{l\in\mathbb{Z}} E\left\langle X_{t-2l},\Phi^0_{2l,m}\right\rangle\left\langle\Phi^0_{2k,m},X_{t-2k}\right\rangle$$
$$+ \sum_{k\in\mathbb{Z}}\sum_{l\in\mathbb{Z}} E\left\langle X_{t-2l},\Phi^0_{2l,m}\right\rangle\left\langle\Phi^0_{2k-1,m},X_{t-2k+1}\right\rangle$$
$$+ \sum_{k\in\mathbb{Z}}\sum_{l\in\mathbb{Z}} E\left\langle X_{t-2l+1},\Phi^0_{2l-1,m}\right\rangle\left\langle\Phi^0_{2k,m},X_{t-2k}\right\rangle$$
$$+ \sum_{k\in\mathbb{Z}}\sum_{l\in\mathbb{Z}} E\left\langle X_{t-2l+1},\Phi^0_{2l-1,m}\right\rangle\left\langle\Phi^0_{2k-1,m},X_{t-2k+1}\right\rangle$$

 Copyright © 2018 John Wiley & Sons Ltd
DOI: 10.1111/jtsa.12283

$$
\begin{aligned}
&= \sum_{k\in\mathbb{Z}}\sum_{l\in\mathbb{Z}} \left\langle C^X_{k-l,(0,0)}\left(\Phi^0_{2k,m}\right), \Phi^0_{2l,m}\right\rangle_{\mathcal{H}} \\
&\quad + \sum_{k\in\mathbb{Z}}\sum_{l\in\mathbb{Z}} \left\langle C^X_{k-l,(0,1)}\left(\Phi^0_{2k-1,m}\right), \Phi^0_{2l,m}\right\rangle_{\mathcal{H}} \\
&\quad + \sum_{k\in\mathbb{Z}}\sum_{l\in\mathbb{Z}} \left\langle C^X_{k-l,(1,0)}\left(\Phi^0_{2k,m}\right), \Phi^0_{2l-1,m}\right\rangle_{\mathcal{H}} \\
&\quad + \sum_{k\in\mathbb{Z}}\sum_{l\in\mathbb{Z}} \left\langle C^X_{k-l,(1,1)}\left(\Phi^0_{2k-1,m}\right), \Phi^0_{2l-1,m}\right\rangle_{\mathcal{H}} .
\end{aligned}
$$

That is the desired result for the case $t \overset{2}{\equiv} 0$. The case $t \overset{2}{\equiv} 1$ is handled in a similar way.

Part (c) is a direct result of Theorem 3.1, so we can proceed to the proof of part (d). Considering part (c) and using the results of Proposition 3 of Hörmann *et al.* (2015) lead to the desired result for $2n$ in place of n.

$$
\begin{aligned}
&\lim_{n\to\infty} \frac{1}{2n} \operatorname{Var}\left(\mathbf{Y}_1 + \cdots + \mathbf{Y}_{2n}\right) \\
&\quad = \lim_{n\to\infty} \frac{1}{2n} \left[\operatorname{Var}\left(\mathbf{Y}_1 + \mathbf{Y}_3 + \cdots + \mathbf{Y}_{2n-1}\right) + \operatorname{Var}\left(\mathbf{Y}_2 + \mathbf{Y}_4 + \cdots + \mathbf{Y}_{2n}\right)\right] \\
&\quad = \frac{2\pi}{2}\left[\operatorname{diag}\left(\lambda_{0,1}, \ldots, \lambda_{0,p}\right) + \operatorname{diag}\left(\lambda_{0,p+1}, \ldots, \lambda_{0,2p}\right)\right]
\end{aligned}
$$

and similarly for $2n+1$ in place of $2n$. This completes the proof.

Proofs of Theorems 3.2 and 3.3. Consider the H^2-valued, mean-zero stationary process $\tilde{X} = \left\{\tilde{X}_t = \left(X_{2t}, X_{2t+1}\right)', t \in \mathbb{Z}\right\}$ and the filter $\left\{\boldsymbol{\Psi}_l, l \in \mathbb{Z}\right\}$ with the following matrix form:

$$
\boldsymbol{\Psi}_l = \begin{pmatrix} \Psi^0_{2l} & \Psi^0_{2l-1} \\ \Psi^1_{2l+1} & \Psi^1_{2l} \end{pmatrix} : \mathcal{H}^2 \to \left((\mathbb{C}^p)^2\right)\mathbb{C}^{2p},
$$

where

$$
\begin{aligned}
\Psi^t_l &: \mathcal{H} \to \mathbb{C}^p \\
\Psi^t_l &: h \longmapsto \left(\left\langle h, \Psi^t_{l,1}\right\rangle, \ldots, \left\langle h, \Psi^t_{l,p}\right\rangle\right)', \quad t = 0, 1,\ l \in \mathbb{Z}.
\end{aligned}
$$

Similarly, define the sequence of operators $\left\{\boldsymbol{\Upsilon}_l, l \in \mathbb{Z}\right\}$ with

$$
\boldsymbol{\Upsilon}_{-l} = \begin{pmatrix} \Upsilon^0_{2l} & \Upsilon^0_{2l+1} \\ \Upsilon^1_{2l-1} & \Upsilon^1_{2l} \end{pmatrix} : \left((\mathbb{C}^p)^2\right)\mathbb{C}^{2p} \to \mathcal{H}^2,
$$

where

$$
\begin{aligned}
\Upsilon^t_l &: \mathbb{C}^p \to \mathcal{H} \\
\Upsilon^t_l &: y \longmapsto \sum_{m=1}^{p} y_m \Upsilon^t_{l,m}, \quad t = 0, 1,\ l \in \mathbb{Z}.
\end{aligned}
$$

 Copyright © 2018 John Wiley & Sons Ltd
DOI: 10.1111/jtsa.12283

Therefore

$$\boldsymbol{\Upsilon}(B)\boldsymbol{\Psi}(B)\tilde{X}_t = \sum_{m=1}^{p} \begin{pmatrix} \check{X}_{2t,m} \\ \check{X}_{2t+1,m} \end{pmatrix}.$$

On the other hand, there exist elements $\psi_q^l = \begin{pmatrix} \psi_{q,1}^l & \psi_{q,2}^l \end{pmatrix}'$ and $\upsilon_q^l = \begin{pmatrix} \upsilon_{q,1}^l & \upsilon_{q,2}^l \end{pmatrix}'$, $q = 1, \dots, 2p$, in $\mathcal{H}^2$, such that

$$\begin{aligned}\boldsymbol{\Psi}_l(h) &= \boldsymbol{\Psi}_l\left(\left(h_1, h_2\right)'\right) = \left(\left\langle h, \psi_1^l\right\rangle_{\mathcal{H}^2}, \dots, \left\langle h, \psi_{2p}^l\right\rangle_{\mathcal{H}^2}\right)' \\ &= \left(\left\langle h_1, \psi_{1,1}^l\right\rangle + \left\langle h_2, \psi_{1,2}^l\right\rangle, \dots, \left\langle h_1, \psi_{2p,1}^l\right\rangle + \left\langle h_2, \psi_{2p,2}^l\right\rangle\right)', \quad \forall h_1, h_2 \in \mathcal{H}\end{aligned}$$

and

$$\begin{aligned}\boldsymbol{\Upsilon}_{-l}(y) &= \boldsymbol{\Upsilon}_{-l}\left(\left(y_1, y_2\right)'\right) \\ &= \sum_{m=1}^{p} y_{1,m}\upsilon_m^l + \sum_{m=1}^{p} y_{2,m}\upsilon_{m+p}^l, \quad \forall y_1, y_2 \in \mathbb{C}^p.\end{aligned}$$

Simple calculations lead to the following relations, valid for $m = 1, \dots, p$, which play a crucial role in the remainder of the proof:

$$\psi_m^l = \begin{pmatrix} \psi_{m,1}^l \\ \psi_{m,2}^l \end{pmatrix} = \begin{pmatrix} \Psi_{2l,m}^0 \\ \Psi_{2l-1,m}^0 \end{pmatrix}, \quad \psi_{m+p}^l = \begin{pmatrix} \psi_{m+p,1}^l \\ \psi_{m+p,2}^l \end{pmatrix} = \begin{pmatrix} \Psi_{2l+1,m}^1 \\ \Psi_{2l,m}^1 \end{pmatrix},$$

$$\upsilon_m^l = \begin{pmatrix} \upsilon_{m,1}^l \\ \upsilon_{m,2}^l \end{pmatrix} = \begin{pmatrix} \Upsilon_{2l,m}^0 \\ \Upsilon_{2l-1,m}^1 \end{pmatrix}, \quad \upsilon_{m+p}^l = \begin{pmatrix} \upsilon_{m+p,1}^l \\ \upsilon_{m+p,2}^l \end{pmatrix} = \begin{pmatrix} \Upsilon_{2l+1,m}^0 \\ \Upsilon_{2l,m}^1 \end{pmatrix}.$$

According to Hörmann *et al.* (2015), we can minimize

$$E\left\|\tilde{X}_t - \boldsymbol{\Upsilon}(B)\boldsymbol{\Psi}(B)\left(\tilde{X}_t\right)\right\|_{\mathcal{H}^2}^2$$

by choosing $\upsilon_{\theta,m} = \sum_{l\in\mathbb{Z}} \upsilon_m^l \mathrm{e}^{il\theta} = \psi_{\theta,m} = \sum_{l\in\mathbb{Z}} \psi_m^l \mathrm{e}^{il\theta}$ as the m-th eigenfunctions of the spectral density operator $\mathcal{F}_\theta^{\tilde{X}}$ of the process $\tilde{X}$. Note that $\mathcal{F}_\theta^{\tilde{X}}$ is nothing other than

$$\mathcal{F}_\theta^{\tilde{X}}(h) = \mathcal{F}_\theta^{\tilde{X}}\left(\begin{pmatrix} h_1 \\ h_2 \end{pmatrix}\right) = \begin{pmatrix} \mathcal{F}_{\theta,(0,0)}^X & \mathcal{F}_{\theta,(0,1)}^X \\ \mathcal{F}_{\theta,(1,0)}^X & \mathcal{F}_{\theta,(1,1)}^X \end{pmatrix}\begin{pmatrix} h_1 \\ h_2 \end{pmatrix}, \quad \forall h_1, h_2 \in \mathcal{H}.$$

This completes the proof.

ACKNOWLEDGMENTS

Piotr Kokoszka is supported by NSF grants DMS-1462067 and DMS-1737795. Neda Mohammadi Jouzdani is supported by a travel grant from the Iranian Ministry of Science Research and Technology.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

Copyright © 2018 John Wiley & Sons Ltd
DOI: 10.1111/jtsa.12283

REFERENCES

Anderson PL, Meerschaert MM. 1997. Periodic moving averages of random variables with regularly varying tails. *Annals of Statistics* **24**: 771–785.

Aue A, Dubart-Norinho D, Hörmann S. 2015. On the prediction of stationary functional time series. *Journal of the American Statistical Association* **110**: 378–392.

Brillinger DR. 1975. *Time Series: Data Analysis and Theory*. New York: Holt.

Fisher RA. 1929. Tests of significance in harmonic analysis. *Proceedings of the Royal Society (A)* **125**: 54–59.

Ghanbarzadeh M, Aminghafari M. 2016. A wavelet characterization of continuous-time periodically correlated processes with application to simulation. *Journal of Time Series Analysis* **37**: 741–762.

Gromenko O, Kokoszka P, Reimherr M. 2017. Detection of change in the spatiotemporal mean function. *Journal of the Royal Statistical Society (B)* **79**: 29–50.

Hays S, Shen H, Huang JZ. 2012. Functional dynamic factor models with application to yield curve forecasting. *The Annals of Applied Statistics* **6**: 870–894.

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

Hörmann S, Kokoszka P, Nisol G. 2018. Testing for periodicity in functional time series. *The Annals of Statistics*. Forthcoming.

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

Horváth L, Kokoszka P, Rice G. 2014. Testing stationarity of functional time series. *Journal of Econometrics* **179**: 66–82.

Hurd HL. 1989. Nonparametric time series analysis for periodically correlated processes. *IEEE Transactions on Information Theory* **35**: 350–359.

Hurd HL, Miamee A. 2007. *Periodically Correlated Random Sequences*. Hoboken: Wiley.

Javorskyj I, Leśkow J, Kravets I, Isayev I, Gajecka E. 2012. Linear filtration methods for statistical analysis of periodically correlated random processes—part I: coherent and component methods and their generalization. *Signal Processing* **92**: 1559–1566.

Klepsch J, Klüppelberg C, Wei T. 2017. Prediction of functional ARMA processes with an application to traffic data. *Econometrics and Statistics* **1**: 128–149.

Kokoszka P, Reimherr M. 2017. *Introduction to Functional Data Analysis*. Boca Raton: CRC Press.

Lund R, Hurd H, Bloomfield P, Smith R. 1995. Climatological time series with periodic correlation. *Journal of Climate* **8**: 410–427.

Panaretos VM, Tavakoli S. 2013a. Cramér–Karhunen–Loève representation and harmonic principal component analysis of functional time series. *Stochastic Processes and their Applications* **123**: 2779–2807.

Panaretos VM, Tavakoli S. 2013b. Fourier analysis of stationary time series in function space. *The Annals of Statistics* **41**: 568–603.

Ramsay J, Hooker G, Graves S. 2009. *Functional Data Analysis with R and MATLAB*. New York: Springer.

Walker G. 1914. On the criteria for the reality of relationships or periodicities. Memoirs of the India Meteorological Department, Vol. 21, 13–15.

Zamani A, Haghbin H, Shishebor Z. 2016. Some tests for detecting cyclic behavior in functional time series. Technical report. Shiraz University.

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

Copyright © 2018 John Wiley & Sons Ltd
J. Time Ser. Anal. **39**: 502–522 (2018)
DOI: 10.1111/jtsa.12283