



**HAL**  
open science

# Hilbert valued fractionally integrated autoregressive moving average processes with long memory operators

Amaury Durand, François Roueff

► **To cite this version:**

Amaury Durand, François Roueff. Hilbert valued fractionally integrated autoregressive moving average processes with long memory operators. 2022. hal-02961227v2

**HAL Id: hal-02961227**

**<https://hal.science/hal-02961227v2>**

Preprint submitted on 25 Sep 2022 (v2), last revised 6 Feb 2024 (v5)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Hilbert space-valued fractionally integrated autoregressive moving average processes with long memory operators

Amaury Durand<sup>\*†</sup>      François Roueff<sup>\*</sup>

September 25, 2022

## Abstract

Fractionally integrated autoregressive moving average (FIARMA) processes have been widely and successfully used to model and predict univariate time series exhibiting long range dependence. Vector and functional extensions of these processes have also been considered more recently. Here we study these processes by relying on a spectral domain approach in the case where the processes are valued in a separable Hilbert space  $\mathcal{H}_0$ . In this framework, the usual univariate long memory parameter  $d$  is replaced by a long memory *operator*  $D$  acting on  $\mathcal{H}_0$ , leading to a class of  $\mathcal{H}_0$ -valued FIARMA( $D, p, q$ ) processes, where  $p$  and  $q$  are the degrees of the AR and MA polynomials. When  $D$  is a normal operator, we provide a necessary and sufficient condition for the  $D$ -fractional integration of an  $\mathcal{H}_0$ -valued ARMA( $p, q$ ) process to be well defined. Then, we derive the best predictor for a class of causal FIARMA processes and study how this best predictor can be consistently estimated from a finite sample of the process. To this end, we provide a general result on quadratic functionals of the periodogram, which incidentally yields a result of independent interest. Namely, for any ergodic stationary process valued in  $\mathcal{H}_0$  with finite second moment, the empirical autocovariance operator converges, in trace-norm, to the true autocovariance operator almost surely at each lag.

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Preliminaries and useful notation</b>	<b>3</b>
2.1	Operators, measurability and integrals . . . . .	3
2.2	Linear filtering of Hilbert space-valued time series in the spectral domain . . . .	4
<b>3</b>	<b>Hilbert space-valued FIARMA processes</b>	<b>5</b>
3.1	Definition of FIARMA processes . . . . .	5
3.2	Existence of FIARMA processes . . . . .	7
3.3	Other long memory processes . . . . .	9
<b>4</b>	<b>Prediction and estimation</b>	<b>10</b>
4.1	Main assumptions and preliminary result . . . . .	10
4.2	FIARMA prediction and estimation . . . . .	12

---

<sup>\*</sup>LTCI, Telecom Paris, Institut Polytechnique de Paris.

<sup>†</sup>EDF R&D, TREE, E36, Lab Les Renardieres, Ecuelles, 77818 Moret sur Loing, France.

Math Subject Classification. Primary: 60G22, 60G12; Secondary: 47A56, 46G10

Keywords. ARFIMA processes. Long memory. Spectral representation of random processes. Functional time series. Hilbert space

<b>5</b>	<b>Postponed proofs</b>	<b>15</b>
5.1	Proofs of Section 3.2 . . . . .	15
5.1.1	Proofs of Lemma 3.2 and Theorem 3.3 . . . . .	15
5.1.2	Proof of Theorem 3.4 . . . . .	16
5.2	Proofs of Section 3.3 . . . . .	19
5.3	Proofs of Section 4.1 . . . . .	21
5.3.1	Preliminary results . . . . .	21
5.3.2	Proof of Theorem 4.2 . . . . .	26
5.3.3	Proof of Theorem 4.3 . . . . .	27
5.4	Proof of Theorem 1.1 . . . . .	29
5.5	Proofs of Section 4.2 . . . . .	30
5.5.1	Preliminary results . . . . .	30
5.5.2	Proof of main results . . . . .	33
<b>A</b>	<b>Technical lemmas</b>	<b>38</b>
<b>B</b>	<b><math>L^2(\mathbb{V}, \mathcal{V}, \xi)</math>-valued weakly stationary time series</b>	<b>39</b>

# 1 Introduction

Over the past several decades, the study of weakly stationary time series valued in a separable Hilbert space has been an active field of research. For example, functional ARMA processes were discussed in [2, 28, 19], a spectral theory was detailed in [24, 23, 29] and several estimation methods were studied in [16, 17, 15, 18, 20, 1, 21, 30]. However, these references mainly focus on short memory processes. The study of long memory processes valued in a separable Hilbert space is a more recent topic as seen in [26, 4, 5, 12, 22]. More specifically, in [22, Section 4], the fractionally integrated autoregressive moving average (often abbreviated as ARFIMA, but we prefer to use FIARMA for reasons that will be made explicit in Remark 3.1) processes are generalized to the case of *curve*, or *functional*, time series. In short, the authors consider the functional case in which the Hilbert space is an  $L^2$  space of real valued functions defined on a compact subset  $\mathcal{C}$  of  $\mathbb{R}$ , and they introduce the time series  $(X_t)_{t \in \mathbb{Z}}$  valued in this Hilbert space defined by

$$X_t(v) = Y_t + \sum_{k=1}^{\infty} \frac{\prod_{j=0}^{k-1} (d+j)}{k!} Y_{t-k}(v), \quad t \in \mathbb{Z}, v \in \mathcal{C}, \quad (1.1)$$

where  $-1/2 < d < 1/2$  and  $Y_t$  is a functional ARMA process. As pointed out in [22, Remark 9], taking the same  $d$  for all  $v \in \mathcal{C}$  in (1.1) is highly restrictive compared to other long memory processes recently introduced. For instance in [4, 5], they consider long memory processes of the form

$$X_t(v) = \sum_{k=0}^{\infty} (1+k)^{-n(v)} \epsilon_{t-k}(v), \quad t \in \mathbb{Z}, v \in \mathbb{V},$$

where  $(\mathbb{V}, \mathcal{V}, \xi)$  is a  $\sigma$ -finite measure space, and  $(\epsilon_t)_{t \in \mathbb{Z}}$  is a white noise valued in  $L^2(\mathbb{V}, \mathcal{V}, \xi)$ . Since the ratio in (1.1) is asymptotically equivalent to  $(1+k)^{-1+d}$  as  $k \rightarrow \infty$ , this new process is, in fact, close to the previous one in the case where  $n(v) = d-1$  for all  $v \in \mathbb{V}$ . A formulation that is not restricted to an  $L^2$  space was proposed in [12] where the author considers long memory processes of the form

$$X_t = \sum_{k=0}^{\infty} (1+k)^{-N} \epsilon_{t-k}, \quad t \in \mathbb{Z}. \quad (1.2)$$

Here,  $(\epsilon_t)_{t \in \mathbb{Z}}$  is a white noise valued in a separable Hilbert space  $\mathcal{H}_0$  and  $N$  is a bounded normal operator on  $\mathcal{H}_0$ . This suggests defining FIARMA processes in (1.1) with  $d$  replaced by a function  $d(v)$ , or in the case where it is valued in an arbitrary separable Hilbert space  $\mathcal{H}_0$ , by a bounded normal operator  $D$  acting on this space.

Therefore, in this paper, we fill this gap by providing a definition of FIARMA processes valued in a separable Hilbert space  $\mathcal{H}_0$  with a long memory operator  $D$ , taken as a bounded linear operator on  $\mathcal{H}_0$ . If  $D$  is normal, then we can rely on its singular value decomposition

and find necessary and sufficient conditions to ensure that the  $\mathcal{H}_0$ -valued FIARMA process with long memory operator  $D$  is well defined. This allows us to compare FIARMA processes with the processes defined by (1.2) as in [12]. Our definition relies on linear filtering in the spectral domain. It is a well known fact that linear filtering of real valued time series in the time domain is equivalent to pointwise multiplication by a transfer function in the frequency domain. This duality also applies to Hilbert space valued time series using a proper spectral representation for them. In this context, pointwise multiplication becomes a pointwise application of an operator-valued transfer function defined on the set of frequencies. A complete account is provided in [11]. Here, we rely on the spectral approach to define a  $D$ -fractional integration filter acting on a weakly stationary process  $X$  valued in  $\mathcal{H}_0$ . We provide necessary and sufficient conditions for this filter to be well defined on a  $X$ , when  $X$  is a  $\mathcal{H}_0$ -valued ARMA process and  $D$  a normal operator. When the ARMA process is a causal, we derive the best predictor of  $X_t$  given its past  $(X_s)_{s < t}$ . It is thus of interest to investigate whether this best predictor can be consistently estimated from a finite sample  $X_1, \dots, X_n$ . We provide a positive answer to this question when the long memory parameter operator  $D$  has a non-negative real part, under mild additional conditions. To this end, we study quadratic functionals based on the periodogram of  $X_1, \dots, X_n$ . A result of this study, which is of independent interest, and appears to be novel based on our up-to-date-knowledge, is the following:

**Theorem 1.1.** *Let  $\mathcal{H}_0$  be a separable Hilbert space and let  $(X_t)_t$  be an  $\mathcal{H}_0$ -valued ergodic stationary process defined on  $(\Omega, \mathcal{F}, \mathbb{P})$  satisfying  $\mathbb{E} \left[ \|X_0\|_{\mathcal{H}_0}^2 \right] < \infty$ . Let us define, for all  $n \geq 1$  and  $1 \leq k \leq n$ ,*

$$X_{n,k}^c = X_k - \frac{1}{n} \sum_{j=1}^n X_j. \quad (1.3)$$

Then, we have, for all  $h \in \mathbb{Z}$ ,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{\substack{1 \leq k, k' \leq n \\ k-k'=h}} (X_{k,n}^c) \otimes (X_{k',n}^c) = \text{Cov}(X_h, X_0) \quad \text{in } \mathcal{S}_1(\mathcal{H}_0), \quad \mathbb{P}\text{-a.s.}, \quad (1.4)$$

where  $\mathcal{S}_1(\mathcal{H}_0)$  is the space of trace-class operators endowed with the trace-norm.

This paper is organized as follows. We first recall in Section 2 the necessary definitions and facts on operator theory and linear filtering needed for our purpose. Then, the construction of FIARMA processes is introduced and discussed in Section 3 with a focus on the case where the long memory operator is normal. In Section 4, the prediction of FIARMA processes is studied. To this end, in Section 4.1, we provide general results for parametric contrast estimation in the spectral domain, based on a finite sample. Then, in Section 4.2, we show how to apply such results for FIARMA prediction. Finally, proofs are provided in Section 5. In particular, Theorem 1.1 is proven in Section 5.4.

## 2 Preliminaries and useful notation

### 2.1 Operators, measurability and integrals

Throughout this paper, we denote by  $\mathcal{L}_b(\mathcal{H}_0, \mathcal{G}_0)$  the set of continuous linear operators defined on the separable (complex) Hilbert space  $\mathcal{H}_0$  onto the separable (complex) Hilbert space  $\mathcal{G}_0$ . The operator norm on  $\mathcal{L}_b(\mathcal{H}_0, \mathcal{G}_0)$  is denoted by  $\|\cdot\|_\infty$ . We denote by  $\mathcal{S}_\infty(\mathcal{H}_0, \mathcal{G}_0)$  its subset of compact operators, and by  $\mathcal{S}_1(\mathcal{H}_0, \mathcal{G}_0)$  and  $\mathcal{S}_2(\mathcal{H}_0, \mathcal{G}_0)$ , its subsets of trace-class and Hilbert-Schmidt operators, respectively, with their respective norms denoted by  $\|\cdot\|_1$  and  $\|\cdot\|_2$ . We follow the usual convention of omitting  $\mathcal{G}_0$  in the notation of operator spaces when  $\mathcal{G}_0 = \mathcal{H}_0$ . We use the notation  $P^H$  for the Hermitian adjoint of an operator  $P \in \mathcal{L}_b(\mathcal{H}_0, \mathcal{G}_0)$ . An operator  $P \in \mathcal{L}_b(\mathcal{H}_0)$  is said to be normal if  $PP^H = P^HP$  and we denote by  $\mathcal{N}(\mathcal{H}_0)$  the set of normal bounded operators. We further denote by  $\mathcal{L}_b^+(\mathcal{H}_0)$ ,  $\mathcal{S}_1^+(\mathcal{H}_0)$  and  $\mathcal{S}_2^+(\mathcal{H}_0)$  the sets of positive, positive trace-class and positive Hilbert-Schmidt operators. Here *positive* refers to *positive-semidefinite*, that is,  $P$  is positive if  $\langle Px, x \rangle_{\mathcal{H}_0} \geq 0$  for all  $x$ . For a positive operator  $P$ , the

operator  $P^{1/2}$  is the unique positive operator satisfying  $(P^{1/2})^2 = P$ . A general and detailed presentation of operator theory can be found in [31].

We will make extensive use of integrals of functions valued in a Banach space (see [9, Chapter 1] for details). Given a measure space  $(\Lambda, \mathcal{A}, \mu)$ , a Banach space  $(E, \|\cdot\|_E)$  and  $p \in [1, \infty]$ , we denote by  $\mathcal{L}^p(\Lambda, \mathcal{A}, E, \mu)$  the space of functions  $f : \Lambda \rightarrow E$  which are Borel-measurable such that  $\int \|f\|_E^p d\mu$  (or  $\mu$ -essup  $\|f\|_E$  for  $p = \infty$ ) is finite. Its quotient space for the  $\mu$ -a.e. equality is denoted by  $L^p(\Lambda, \mathcal{A}, E, \mu)$ . We use the same notation for  $E = \mathcal{L}_b^+(\mathcal{H}_0)$ ,  $\mathcal{S}_1^+(\mathcal{H}_0)$  or  $\mathcal{S}_2^+(\mathcal{H}_0)$ , in which case  $L^p(\Lambda, \mathcal{A}, E, \mu)$  is a cone subset of the corresponding  $L^p$  space.

In the particular case where  $E = \mathcal{L}_b(\mathcal{H}_0, \mathcal{G}_0)$  for two separable Hilbert spaces  $\mathcal{H}_0, \mathcal{G}_0$ , we also use a weaker notion of measurability. Namely, we say that a function  $\Phi : \Lambda \rightarrow \mathcal{L}_b(\mathcal{H}_0, \mathcal{G}_0)$  is *simply measurable* if for all  $x \in \mathcal{H}_0$ ,  $\lambda \mapsto \Phi(\lambda)x$  is measurable as a  $\mathcal{G}_0$ -valued function. The set of simple measurable functions from  $(\Lambda, \mathcal{A})$  to  $\mathcal{L}_b(\mathcal{H}_0, \mathcal{G}_0)$  is denoted by  $\mathbb{F}_s(\Lambda, \mathcal{A}, \mathcal{H}_0, \mathcal{G}_0)$  where, again, we omit  $\mathcal{G}_0$  if  $\mathcal{H}_0 = \mathcal{G}_0$ . A mapping  $\Phi : \Lambda \rightarrow E$  with  $E = \mathcal{S}_1(\mathcal{H}_0, \mathcal{G}_0)$  or  $E = \mathcal{S}_2(\mathcal{H}_0, \mathcal{G}_0)$  is simply measurable if and only if it is Borel measurable (see Lemma 5.1 in [11]). A useful consequence is that, if  $\Phi \in L^1(\Lambda, \mathcal{A}, \mathcal{S}_1^+(\mathcal{H}_0), \mu)$ , then the function  $\Phi^{1/2} : \lambda \mapsto \Phi(\lambda)^{1/2}$  is in  $L^2(\Lambda, \mathcal{A}, \mathcal{S}_2^+(\mathcal{H}_0), \mu)$ .

## 2.2 Linear filtering of Hilbert space-valued time series in the spectral domain

This section gathers the spectral theory used for linear filtering of times series valued in a separable Hilbert space. We refer the reader to [11, Section 4] for details. In the following, we denote by  $\mathbb{T}$  the set  $\mathbb{R}/2\pi\mathbb{Z}$ , which can be represented by an interval such as  $[-\pi, \pi)$ . Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space and  $\mathcal{H}_0$  a separable Hilbert space. We recall that the expectation of  $X \in L^2(\Omega, \mathcal{F}, \mathcal{H}_0, \mathbb{P})$  is the unique vector  $\mathbb{E}[X] \in \mathcal{H}_0$  satisfying

$$\langle \mathbb{E}[X], x \rangle_{\mathcal{H}_0} = \mathbb{E}[\langle X, x \rangle_{\mathcal{H}_0}], \quad \text{for all } x \in \mathcal{H}_0.$$

The covariance operator between  $X, Y \in L^2(\Omega, \mathcal{F}, \mathcal{H}_0, \mathbb{P})$  is the unique linear operator  $\text{Cov}(X, Y) \in \mathcal{L}_b(\mathcal{H}_0)$  satisfying

$$\langle \text{Cov}(X, Y)y, x \rangle_{\mathcal{H}_0} = \text{Cov}(\langle X, x \rangle_{\mathcal{H}_0}, \langle Y, y \rangle_{\mathcal{H}_0}), \quad \text{for all } x, y \in \mathcal{H}_0.$$

A process  $X := (X_t)_{t \in \mathbb{Z}}$  is said to be an  $\mathcal{H}_0$ -valued, weakly stationary process if

- (i) For all  $t \in \mathbb{Z}$ ,  $X_t \in L^2(\Omega, \mathcal{F}, \mathcal{H}_0, \mathbb{P})$ .
- (ii) For all  $t \in \mathbb{Z}$ ,  $\mathbb{E}[X_t] = \mathbb{E}[X_0]$ . We say that  $X$  is centered if  $\mathbb{E}[X_0] = 0$ .
- (iii) For all  $t, h \in \mathbb{Z}$ ,  $\text{Cov}(X_{t+h}, X_t) = \text{Cov}(X_h, X_0)$ .

We denote by  $\mathcal{M}(\Omega, \mathcal{F}, \mathcal{H}_0, \mathbb{P})$  the space of all centered random variables in  $L^2(\Omega, \mathcal{F}, \mathcal{H}_0, \mathbb{P})$ . Let  $\mathcal{H} = \mathcal{M}(\Omega, \mathcal{F}, \mathcal{H}_0, \mathbb{P})$  and  $X = (X_t)_{t \in \mathbb{Z}} \in \mathcal{H}^{\mathbb{Z}}$  be a centered, weakly stationary,  $\mathcal{H}_0$ -valued time series. As explained in [11, Section 4], a spectral representation for  $X$  amounts to define a random Gramian-orthogonally scattered measure  $\hat{X}$  on  $(\mathbb{T}, \mathcal{B}(\mathbb{T}))$  such that

$$X_t = \int e^{i\lambda t} \hat{X}(d\lambda) \quad \text{for all } t \in \mathbb{Z}. \quad (2.1)$$

The intensity measure  $\nu_X : \mathcal{B}(\mathbb{T}) \rightarrow \mathcal{S}_1^+(\mathcal{H}_0)$  of  $\hat{X}$  is called the *spectral operator measure* and is characterized by the identity

$$\text{Cov}(X_h, X_0) = \int e^{ih\lambda} \nu_X(d\lambda), \quad \text{for all } h \in \mathbb{Z}.$$

The spectral operator measure is a trace-class Positive Operator-Valued Measure (p.o.v.m.) in the sense that it is a mapping from  $\mathcal{B}(\mathbb{T})$  to  $\mathcal{S}_1^+(\mathcal{H}_0)$  which is  $\sigma$ -additive for the  $\|\cdot\|_1$ -norm. Note that, in this case, the mapping  $\|\nu_X\|_1 : A \mapsto \|\nu_X(A)\|_1$  is a finite non-negative measure. Throughout this paper, we use the Radon-Nikodym property of the trace-class p.o.v.m.  $\nu_X$  which is a consequence of Theorem 1 in [8, Chapter III, Section 3]. Namely, for any  $\sigma$ -finite non-negative measure  $\mu$  on  $(\mathbb{T}, \mathcal{B}(\mathbb{T}))$ , which dominates  $\|\nu_X\|_1$ , there exists a unique

$g \in L^1(\mathbb{T}, \mathcal{B}(\mathbb{T}), \mathcal{S}_1^+(\mathcal{H}_0), \mu)$  such that, for all  $A \in \mathcal{B}(\mathbb{T})$ ,  $\nu_X(A) = \int_A g d\mu$ . In this case, we say that  $g$  is the *spectral operator density* of  $X$  with respect to  $\mu$ , and we write  $d\nu_X = g d\mu$ . In the following, when we say that  $g$  is the spectral operator density of  $X$  with respect to a  $\sigma$ -finite non-negative measure  $\mu$ , it is implicitly assumed that  $\mu$  dominates  $\|\nu_X\|_1$ .

Let us now briefly introduce how linear filtering in the spectral domain. We only state the facts that will be useful in the following and refer the reader to [11] for further details. The spectral representation (2.1) can be extended to define a Gramian-isometric mapping from the modular *spectral domain*  $\widehat{\mathcal{H}}^X$  to the modular *time domain*  $\mathcal{H}^X$ , also denoted as an integral with respect to  $\widehat{X}$ , namely,

$$Y = \int \Phi(\lambda) \widehat{X}(d\lambda), \quad Y \in \mathcal{H}^X, \quad \Phi \in \widehat{\mathcal{H}}^X.$$

Here,  $\mathcal{H}^X$  is the smallest closed linear subspace of  $\mathcal{M}(\Omega, \mathcal{F}, \mathcal{H}_0, \mathbb{P})$ , which contains  $\{X_t : t \in \mathbb{Z}\}$  and is stable through the left multiplication by any operator of  $\mathcal{L}_b(\mathcal{H}_0)$ . The space  $\widehat{\mathcal{H}}^X$  is its spectral counterpart, a space of operator-valued functions defined on  $(\mathbb{T}, \mathcal{B}(\mathbb{T}))$  which only depends on  $\nu_X$  and is stable through the same module action, namely, through left multiplication by an operator of  $\mathcal{L}_b(\mathcal{H}_0)$ . Conversely, given an operator-valued function  $\Phi$  defined on  $(\mathbb{T}, \mathcal{B}(\mathbb{T}))$ , we denote by  $\mathcal{S}_\Phi(\Omega, \mathcal{F}, \mathbb{P})$  the class of all centered weakly stationary processes  $X$  such that  $\Phi \in \widehat{\mathcal{H}}^X$ . Then, the time-shift invariant linear filter with *transfer operator function*  $\Phi$  is the mapping defined on  $\mathcal{S}_\Phi(\Omega, \mathcal{F}, \mathbb{P})$  by mapping a centered weakly stationary process  $X$  (the input) to the centered weakly stationary process  $Y$  (the output) defined by

$$Y_t = \int e^{it\lambda} \Phi(\lambda) \widehat{X}(d\lambda), \quad t \in \mathbb{Z},$$

which we also write

$$Y = F_\Phi(X) \quad \text{or} \quad \widehat{Y}(d\lambda) = \Phi(\lambda) \widehat{X}(d\lambda). \quad (2.2)$$

We will use the following result, where we characterize the domain of definition of a filter  $F_\Phi$  in the case where  $\Phi$  is valued in  $\mathcal{L}_b(\mathcal{H}_0)$ . It follows by applying [11, Proposition 4.8] with  $\mathbb{G} = \mathbb{Z}$  and  $\mathcal{G}_0 = \mathcal{H}_0$ .

**Proposition 2.1.** *Let  $\mathcal{H}_0$  be a separable Hilbert space,  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space, and  $\Phi \in \mathbb{F}_s(\mathbb{T}, \mathcal{B}(\mathbb{T}), \mathcal{H}_0)$ . Let  $X$  be an  $\mathcal{H}_0$ -valued centered weakly stationary process admitting  $g_X$  as spectral operator density with respect to a  $\sigma$ -finite non-negative measure  $\mu$  on  $(\mathbb{T}, \mathcal{B}(\mathbb{T}))$ . Then, the mapping  $\|\Phi g_X \Phi^H\|_1$  is measurable from  $(\mathbb{T}, \mathcal{B}(\mathbb{T}))$  to  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ , and we have  $X \in \mathcal{S}_\Phi(\Omega, \mathcal{F}, \mathbb{P})$  if and only if  $\int \|\Phi g_X \Phi^H\|_1 d\mu < \infty$ .*

### 3 Hilbert space-valued FIARMA processes

In this section, we propose a definition of FIARMA processes valued in a separable Hilbert space thus extending the definition of [22, Section 4] to an operator long memory parameter. This definition is introduced in Section 3.1 where we also recall known results on the existence of ARMA processes. We then state the main results, namely 1) Theorem 3.3 where necessary and sufficient conditions are given for a weakly stationary  $\mathcal{H}_0$ -valued process  $X$  to belong the domain of definition of the fractional integration operator filter, 2) Theorem 3.4 where we specify these conditions to the case where  $X$  is an ARMA process, thus ensuring the existence of FIARMA processes, and 3) Proposition 3.6 where we compare the obtained FIARMA processes to the processes introduced in [12]. The first two points are found in Section 3.2 and the third in Section 3.3.

#### 3.1 Definition of FIARMA processes

Let  $\mathcal{H}_0$  be a separable Hilbert space. In the following, for all  $D \in \mathcal{L}_b(\mathcal{H}_0)$  and  $z \in \mathbb{C} \setminus [1, \infty)$ , we will use

$$(1 - z)^D = \exp(D \ln(1 - z)) = \sum_{k=0}^{\infty} \frac{1}{k!} (D \ln(1 - z))^k,$$

where  $\ln$  denotes the principal complex logarithm, so that  $z \mapsto \ln(1 - z)$  is holomorphic on  $\mathbb{C} \setminus [1, \infty)$ , and so is  $z \mapsto (1 - z)^D$ , as a  $\mathcal{L}_b(\mathcal{H}_0)$ -valued function (see [13, Chapter 1] for an

overview on the subject). Let us now introduce the fractional integration operator transfer function.

**Definition 3.1** (Fractional integration operator transfer function). *Let  $\mathcal{H}_0$  be a separable Hilbert space and  $D \in \mathcal{L}_b(\mathcal{H}_0)$ . We define the  $D$ -order fractional integration operator transfer function  $\text{FI}_D$  by*

$$\text{FI}_D(\lambda) = \begin{cases} (1 - e^{-i\lambda})^{-D} & \text{if } \lambda \neq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Using the properties of  $z \mapsto (1 - z)^D$  recalled previously, we see that  $\text{FI}_D$  is a mapping from  $\mathbb{T}$  to  $\mathcal{L}_b(\mathcal{H}_0)$ , continuous on  $\mathbb{T} \setminus \{0\}$ . Then, we have  $\text{FI}_D \in \mathbb{F}_s(\mathbb{T}, \mathcal{B}(\mathbb{T}), \mathcal{H}_0)$  and we can define the filter  $F_{\text{FI}_D}$  as in (2.2) of which the domain of definition are the centered weakly stationary  $\mathcal{H}_0$ -valued processes  $X \in \mathcal{S}_{\text{FI}_D}(\Omega, \mathcal{F}, \mathbb{P})$ . Since  $\text{FI}_D$  has a singularity at the null frequency, the domain  $\mathcal{S}_{\text{FI}_D}(\Omega, \mathcal{F}, \mathbb{P})$  is not obvious. For instance, in the scalar case, it is well known that if  $X$  has a positive and continuous spectral density at the null frequency, then  $F_{\text{FI}_d}(X)$  is well defined if and only if  $d < 1/2$ . We provide a complete description of  $\mathcal{S}_{\text{FI}_D}(\Omega, \mathcal{F}, \mathbb{P})$  in Section 3.2 when  $D$  is a normal operator.

A *fractionally integrated autoregressive moving average* (FIARMA) process is simply the output of the filter in the case where  $X$  is an  $\mathcal{H}_0$ -valued autoregressive moving average (ARMA) process. Let us first recall a basic result on the existence of weakly stationary ARMA processes (see [28, Corollary 2.2]).

**Theorem 3.1.** *Let  $\mathcal{H}_0$  be a separable Hilbert space and  $p, q$  be two positive integers. Let  $A_1, \dots, A_p \in \mathcal{L}_b(\mathcal{H}_0)$ ,  $B_1, \dots, B_q \in \mathcal{L}_b(\mathcal{H}_0)$  and  $Z = (Z_t)_{t \in \mathbb{Z}}$  be a centered  $\mathcal{H}_0$ -valued white noise (i.e. a weakly stationary  $\mathcal{H}_0$ -valued process with constant spectral density operator). Suppose that*

$$\Phi(z) := \text{Id}_{\mathcal{H}_0} - \sum_{k=1}^p A_k z^k \quad \text{is invertible for all } z \in \mathbb{U}, \quad (3.1)$$

where  $\mathbb{U} = \{z \in \mathbb{C} : |z| = 1\}$  is the complex unit circle. Then,

$$X_t - \sum_{k=1}^p A_k X_{t-k} = Z_t + \sum_{k=1}^q B_k Z_{t-k}, \quad t \in \mathbb{Z}, \quad (3.2)$$

admits a unique weakly stationary solution. This solution is called an  $\mathcal{H}_0$ -valued ARMA( $p, q$ ) process.

Explicit constructions of the solution in the time domain can be found in [2, 28, 19]. Using a spectral approach, with  $\Phi$  as in (3.1) and  $\Theta(z) := \text{Id}_{\mathcal{H}_0} + \sum_{k=1}^q B_k z^k$ , the solution is more directly given by

$$\hat{X}(d\lambda) = \left[ \Phi(e^{-i\lambda}) \right]^{-1} \Theta(e^{-i\lambda}) \hat{Z}(d\lambda),$$

using the notation introduced in (2.2). In the following, for any integer  $d \in \mathbb{N}$ ,  $\mathcal{P}_d(\mathcal{H}_0)$  denotes the set of polynomials  $\mathfrak{p}$  of degree  $d$  with coefficients in  $\mathcal{L}_b(\mathcal{H}_0)$ , such that  $\mathfrak{p}(0) = \text{Id}_{\mathcal{H}_0}$  and  $\mathcal{P}_d^*(\mathcal{H}_0)$  denotes the subset of all  $\mathfrak{p} \in \mathcal{P}_d(\mathcal{H}_0)$ , which are invertible on  $\mathbb{U}$ . In particular, (3.1) is equivalent to saying that  $\Phi \in \mathcal{P}_p^*(\mathcal{H}_0)$ . Time domain approaches for defining ARMA processes are easier to derive when Condition (3.1) is extended on the closed unit disk, that is,

$$\Phi(z) = \text{Id}_{\mathcal{H}_0} - \sum_{k=1}^p A_k z^k \quad \text{is invertible for all } z \in \overline{\mathbb{D}}, \quad (3.3)$$

where  $\mathbb{D} := \{z \in \mathbb{C} : |z| < 1\}$  and  $\overline{\mathbb{D}} = \{z \in \mathbb{C} : |z| \leq 1\}$  denote the open and closed complex unit discs of  $\mathbb{C}$ , respectively. We do not need Condition (3.3) for defining FIARMA processes. However, we will assume that  $\Phi$  and  $\Theta$  satisfy such a condition to derive predictors (see Section 4), as in the well known case of univariate FIARMA processes.

We can now define FIARMA processes as follows.

**Definition 3.2** (Hilbert space-valued FIARMA processes). *Let  $\mathcal{H}_0$  be a separable Hilbert space and  $p, q$  be two non-negative integers. Let  $D \in \mathcal{L}_b(\mathcal{H}_0)$ ,  $\Theta \in \mathcal{P}_q(\mathcal{H}_0)$ ,  $\Phi \in \mathcal{P}_p^*(\mathcal{H}_0)$  and  $Z$  be an  $\mathcal{H}_0$ -valued centered white noise. Let  $X$  be the ARMA( $p, q$ ) process defined by*

$\hat{X}(d\lambda) = [\mathbb{H}(e^{-i\lambda})]^{-1}\theta(e^{-i\lambda})\hat{Z}(d\lambda)$  and suppose that  $X \in \mathcal{S}_{\text{FI}_D}(\Omega, \mathcal{F}, \mathbb{P})$ . Then, the process defined by  $Y = \text{FI}_D(X)$ , or, in the spectral domain, by

$$\hat{Y}(d\lambda) = \text{FI}_D(\lambda)[\mathbb{H}(e^{-i\lambda})]^{-1}\theta(e^{-i\lambda})\hat{Z}(d\lambda), \quad (3.4)$$

is called a FIARMA process of order  $(p, q)$  with long memory operator  $D$ , abbreviated as FIARMA( $D, p, q$ ).

**Remark 3.1.** Definition 3.2 extends the usual definition of univariate ( $\mathbb{C}$  or  $\mathbb{R}$ -valued) ARFIMA( $p, d, q$ ) processes to the Hilbert space-valued case. In the general case, we use the acronym FIARMA to indicate the order of the operators in the definition (3.4), where the fractional integration operator appears on the left of the autoregressive operator, which then is on the left of the moving average operator. We also respected this order in the list of parameters  $(D, p, q)$ . Following this convention, an ARFIMA( $p, D, q$ ) process is, in turn, defined as the solution of (3.2), with  $Z$  defined as a FIARMA( $0, D, q$ ) process. Having this convention in mind is important since the ARFIMA( $p, D, q$ ) process does not coincide with the FIARMA( $D, p, q$ ) process, except in highly unique instances such as the univariate case where all operators commute.

Definition 3.2 extends the definition of ARFIMA curve time series proposed in [22] where it is restricted to the case where  $D$  is a scalar operator, that is  $D : f \mapsto d \times f$  for a constant  $-1/2 < d < 1/2$ . In this particular case, it is rather straightforward to show that  $X \in \mathcal{S}_{\text{FI}_D}(\Omega, \mathcal{F}, \mathbb{P})$  for any ARMA process  $X$  by directly making use of Proposition 2.1. However, in Remark 3.2(5), it will be obtained as a special case of Theorem 3.4.

## 3.2 Existence of FIARMA processes

In this section, we provide a necessary and sufficient condition for the existence of FIARMA processes as defined in Definition 3.2 in the case where  $D$  is a *normal* operator. In this case, we can rely on the singular value decomposition of  $D$  (see [6, Theorem 9.4.6, Proposition 9.4.7]). Namely, if  $D \in \mathcal{N}(\mathcal{H}_0)$ , then there exists a  $\sigma$ -finite measure space  $(\mathcal{V}, \mathcal{V}, \xi)$ , a unitary operator  $U : \mathcal{H}_0 \rightarrow L^2(\mathcal{V}, \mathcal{V}, \xi)$  and  $d \in L^\infty(\mathcal{V}, \mathcal{V}, \xi)$ , such that

$$UDU^H = M_d, \quad (3.5)$$

where  $M_d$  denotes the pointwise multiplicative operator on  $L^2(\mathcal{V}, \mathcal{V}, \xi)$  associated to  $d$ , that is  $M_d : f \mapsto d \times f$ . We say that  $D$  has a *singular value function*  $d$  on  $L^2(\mathcal{V}, \mathcal{V}, \xi)$  with a *decomposition operator*  $U$ . Using the decomposition operator, we can rely on the process  $UX = (UX_t)_{t \in \mathbb{Z}}$  valued in  $\mathcal{G}_0 := L^2(\mathcal{V}, \mathcal{V}, \xi)$ . Note that  $\mathcal{G}_0$  is separable because it is isometrically isomorphic to  $\mathcal{H}_0$  through the unitary operator  $U$ . It is straightforward to check that, if  $g_X$  is the spectral operator density of  $X$  with respect to a non-negative measure  $\mu$  on  $(\mathbb{T}, \mathcal{B}(\mathbb{T}))$ , then the function  $g_{UX} \in L^2(\mathbb{T}, \mathcal{B}(\mathbb{T}), \mathcal{S}_1^+(\mathcal{G}_0), \mu)$  defined by  $g_{UX}(\lambda) = Ug_X(\lambda)U^H$ , for all  $\lambda \in \mathbb{T}$ , is the spectral operator density of  $UX$  with respect to  $\mu$ . Note that we can always find a function  $h \in L^2(\mathbb{T}, \mathcal{B}(\mathbb{T}), \mathcal{S}_2(\mathcal{G}_0), \mu)$  such that  $g_{UX}(\lambda) = h(\lambda)[h(\lambda)]^H$  for  $\mu$ -a.e.  $\lambda \in \mathbb{T}$  (take e.g.  $h = g_{UX}^{1/2}$ ). Then, [31, Theorem 6.11] gives that, for all  $\lambda \in \mathbb{T}$ , the operator  $h(\lambda)$  can be written as an integral operator with a kernel  $\mathcal{H}(\cdot, \cdot; \lambda)$  in  $L^2(\mathcal{V}^2, \mathcal{V}^{\otimes 2}, \xi^{\otimes 2})$ . In the following, we need the measurability of  $\mathcal{H}$  on  $(\mathcal{V}^2 \times \mathbb{T}, \mathcal{V}^{\otimes 2} \otimes \mathcal{B}(\mathbb{T}))$  as given by the following lemma.

**Lemma 3.2.** *Let  $(\mathcal{V}, \mathcal{V}, \xi)$  be a  $\sigma$ -finite measure space and suppose that the Hilbert space  $\mathcal{G}_0 = L^2(\mathcal{V}, \mathcal{V}, \xi)$  is separable. Let  $K$  be a measurable function from  $(\Lambda, \mathcal{A})$  to  $(\mathcal{S}_2(\mathcal{G}_0), \mathcal{B}(\mathcal{S}_2(\mathcal{G}_0)))$ . Then, there exists a function  $\mathcal{K} : (v, v', \lambda) \mapsto \mathcal{K}(v, v'; \lambda)$  measurable from  $(\mathcal{V}^2 \times \Lambda, \mathcal{V}^{\otimes 2} \otimes \mathcal{A})$  to  $(\mathbb{C}, \mathcal{B}(\mathbb{C}))$  such that, for all  $\lambda \in \Lambda$ ,  $f \in \mathcal{H}_0$  and  $v \in \mathcal{V}$ ,*

$$[K(\lambda)f](v) = \int \mathcal{K}(v, v'; \lambda) f(v') \xi(dv'). \quad (3.6)$$

Moreover, if  $K \in \mathcal{L}^2(\Lambda, \mathcal{A}, \mathcal{S}_2(\mathcal{G}_0), \mu)$  for some non-negative measure  $\mu$  on  $(\Lambda, \mathcal{A})$ , then  $\mathcal{K} \in \mathcal{L}^2(\mathcal{V}^2 \times \Lambda, \mathcal{V}^{\otimes 2} \otimes \mathcal{A}, \xi^{\otimes 2} \otimes \mu)$ .

Based on this lemma, for all  $\lambda \in \Lambda$ , the identity (3.6) defines  $\mathcal{K}(\cdot; \lambda)$  uniquely over  $\mathcal{V}^2$  up to a  $\xi^{\otimes 2}$ -null set. This allows us to introduce the following definition.



**Definition 3.3** (Joint kernel of  $\mathcal{S}_2$ -valued functions). *Under the assumptions of Lemma 3.2, we call  $\mathcal{K}$  the  $\Lambda$ -joint kernel of  $K$ .*

Assuming that  $D$  is normal allows us to characterize the domain of definition of the  $D$ -order fractional integration operator filter, as shown in the following result, which may be of independent interest.

**Theorem 3.3.** *Let  $\mathcal{H}_0$  be a separable Hilbert space and  $X = (X_t)_{t \in \mathbb{Z}}$  be an  $\mathcal{H}_0$ -valued weakly stationary time series defined on  $(\Omega, \mathcal{F}, \mathbb{P})$  with spectral operator density  $g_X$  with respect to a non-negative measure  $\mu$  on  $(\mathbb{T}, \mathcal{B}(\mathbb{T}))$ . Let  $D$  be in  $\mathcal{N}(\mathcal{H}_0)$  with singular value function  $d$  on  $\mathcal{G}_0 := L^2(\mathcal{V}, \mathcal{V}, \xi)$  and decomposition operator  $U$ . Let  $h \in L^2(\mathbb{T}, \mathcal{B}(\mathbb{T}), \mathcal{S}_2(\mathcal{G}_0), \mu)$  be such that  $\lambda \mapsto h(\lambda)[h(\lambda)]^H$  is the spectral operator density of  $UX = (UX_t)_{t \in \mathbb{Z}}$  with respect to  $\mu$ , that is,  $h(\lambda)[h(\lambda)]^H = U g_X(\lambda) U^H$  for  $\mu$ -a.e.  $\lambda \in \mathbb{T}$ . Let  $\mathcal{K}$  denote the  $\mathbb{T}$ -joint kernel function of  $h$ . Then, the following assertions are equivalent.*

(i) *We have  $X \in \mathcal{S}_{\text{FID}}(\Omega, \mathcal{F}, \mathbb{P})$ .*

(ii) *There exists  $\eta \in (0, \pi)$  such that*

$$\int_{\mathbb{V}^2 \times ((-\eta, \eta) \setminus \{0\})} |\lambda|^{-2\Re(d(v))} |\mathcal{K}(v, v'; \lambda)|^2 \xi(dv) \xi(dv') \mu(d\lambda) < \infty .$$

(iii) *We have*

$$\int_{\mathbb{V}^2 \times ((-\pi, \pi) \setminus \{0\})} |\lambda|^{-2\Re_+(d(v))} |\mathcal{K}(v, v'; \lambda)|^2 \xi(dv) \xi(dv') \mu(d\lambda) < \infty , \quad (3.7)$$

where, for all  $z \in \mathbb{C}$ ,  $\Re_+(z) = \max(0, (z + \bar{z})/2)$  denotes the non-negative part of  $z$ .

In the following theorem, we specify Theorem 3.3 in the case where  $X$  is an ARMA process, as in [22], but we let  $D$  being any normal operator and not necessarily a scalar one. Our necessary and sufficient condition relies on the following definition:

$$P_n(\phi, \theta) = \{ [\phi]^{-1} \theta \} \circ \exp^{(n)}(0) , \quad n \in \mathbb{N}, \phi \in \mathcal{P}_p^*, \theta \in \mathcal{P}_q , \quad (3.8)$$

where the exponent  $^{(n)}$  here denotes the  $n$ -th derivative of the mapping  $z \mapsto [\phi(e^z)]^{-1} \theta(e^z)$ , which is infinitely differentiable in a neighborhood of  $z = 0$  as a  $\mathcal{L}_b(\mathcal{H}_0)$ -valued function, since  $\phi \in \mathcal{P}_p^*(\mathcal{H}_0)$ . In fact we have

$$P_0(\phi, \theta) = [\phi(1)]^{-1} \theta(1) \quad \text{and} \quad P_n(\phi, \theta) = \sum_{k=1}^n b_{n,k} [\phi]^{-1} \theta^{(k)}(1) , \quad n \geq 1 ,$$

where  $b_{n,k}$  are some positive rational coefficients (namely the exponential Bell  $n$ -order Bell polynomial taken at  $(1, \dots, 1)$ ).

**Theorem 3.4.** *Let  $\mathcal{H}_0$  be a separable Hilbert space. Let  $X$  be an  $\mathcal{H}_0$ -valued ARMA( $p, q$ ) process defined by  $\dot{X}(d\lambda) = [\phi(e^{-i\lambda})]^{-1} \theta(e^{-i\lambda}) \dot{Z}(d\lambda)$  with  $\theta \in \mathcal{P}_q(\mathcal{H}_0)$ ,  $\phi \in \mathcal{P}_p^*(\mathcal{H}_0)$  and  $Z$  a white noise with covariance operator  $\Sigma$  defined on  $(\Omega, \mathcal{F}, \mathbb{P})$ . Let  $D \in \mathcal{N}(\mathcal{H}_0)$ , with singular value function  $d$  on  $\mathcal{G}_0 := L^2(\mathcal{V}, \mathcal{V}, \xi)$  and decomposition operator  $U$ . Let  $\sigma_n : v \mapsto (\mathbb{E} [ |W_n(v, \cdot)|^2 ])^{1/2}$  where  $W_n = U P_n(\phi, \theta) Z_0$  is seen as a  $\mathbb{C}$ -valued function defined on  $\mathbb{V} \times \Omega$ . Then, the two following assertions are equivalent:*

(i) *We have  $X \in \mathcal{S}_{\text{FID}}(\Omega, \mathcal{F}, \mathbb{P})$ .*

(ii) *For all  $n \in \mathbb{N}$ , we have  $\Re(d) < n + 1/2$ ,  $\xi$ -a.e. on  $\{\sigma_n > 0\}$  and*

$$\int_{\{\Re(d) < n+1/2\}} \frac{\sigma_n^2(v)}{1 + 2n - 2\Re(d(v))} \xi(dv) < \infty . \quad (3.9)$$

Note that there is a slight abuse of notation in the definition of  $\sigma_n$  since the definition of measurability for a  $L^2(\mathcal{V}, \mathcal{V}, \xi)$ -valued random variable does not necessarily ensure measurability of  $W_n(v, \cdot)$  as a  $\mathbb{C}$ -valued random variable. This abuse of notation is common in the literature on functional data analysis and is harmless because we can always find a version of  $W$  which is jointly measurable on  $(\mathbb{V} \times \Omega, \mathcal{V} \otimes \mathcal{F})$ , see Proposition B.1 in Appendix B.

**Remark 3.2.** *Let us briefly comment on Theorem 3.4.*

- (1) First, by definition of  $W_n$  and  $\sigma_n$ , we have  $\sigma_n \in L^2(\mathcal{V}, \mathcal{V}, \xi)$  and thus Assertion (ii) always holds for  $n$  large enough, namely, for all  $n > \sup(\Re(d)) - 1/2$  (recall that the singular value function of a normal operator is bounded).
- (2) If  $\theta = \phi$  is the unit polynomial, the ARMA process  $X$  equals the white noise process  $Z$ ,  $P_n(\phi, \theta) = \text{Id}_{\mathcal{H}_0}$ , and  $\sigma_n = \sigma_0$  for all  $n \in \mathbb{N}$ , so that Condition (ii) only needs to be verified for  $n = 0$ .
- (3) In the  $N$ -dimensional case with  $n$  finite, we have  $\mathcal{V} = \{1, \dots, N\}$ ,  $\xi$  is the counting measure on  $\mathcal{V}$ , and  $U$  can be interpreted as a  $n \times n$  unitary matrix, and  $d$  and  $\sigma_n$  as  $N$ -dimensional vectors. Condition (ii) then says that  $\Re(d(k)) < n + 1/2$  for all  $n \in \mathbb{N}$  and  $k \in \{1, \dots, N\}$  such that  $\sigma_n(k) > 0$ .
- (4) For the real univariate case ( $N = 1$ ,  $D = d \in \mathbb{R}$  in (3)), Condition (ii) says that  $d < n_0 + 1/2$ , where  $n_0$  is the smallest  $n$  such that  $\sigma_n > 0$ . Ruling out the case where  $Z$  is the null process (in which case  $\Sigma = 0$  and  $\sigma_n = 0$  for all  $n \in \mathbb{N}$ ), one can see that  $n_0$  equals 0 if  $\theta(1) \neq 0$  and  $n_0$  equals the order of multiplicity of 1 as a root of  $\theta$  otherwise (in other words, it corresponds to the difference operator largest order contained in the MA operator). In particular, we find the usual  $d < 1/2$  condition for the existence of a weakly stationary ARFIMA( $p, d, q$ ) model in the case where the underlying ARMA( $p, q$ ) process is canonical ( $\phi$  and  $\theta$  do not vanish on the unit disk). If  $n_0 \geq 1$ , the usual convention is to include the difference operator as a negative exponent of the fractional integration operator hence leading to an ARFIMA( $p, d - n_0, q - n_0$ ) with  $d - n_0 < 1/2$ .
- (5) We already mentioned in (1.1) the case treated in [22, Section 4]. In the setting of Theorem 3.4, it corresponds to the case where  $D = d \times \text{Id}_{\mathcal{H}_0}$  is a scalar operator on  $\mathcal{H}_0 = L^2(\mathcal{C}, \mathcal{B}(\mathcal{C}), \xi)$  for a compact subset  $\mathcal{C}$  of  $\mathbb{R}$ ,  $\xi$  being the Lebesgue measure on  $\mathcal{C}$  and  $-1/2 < d < 1/2$  (thus  $d(v) \equiv d$  and  $U = \text{Id}_{\mathcal{H}_0}$ ). Under this assumption, Condition (3) trivially holds since  $1 + 2n - 2d > 0$  and  $\sigma_n \in L^2(\mathcal{C}, \mathcal{B}(\mathcal{C}), \xi)$  for all  $n \in \mathbb{N}$ .

### 3.3 Other long memory processes

Several non-equivalent definitions of long range dependence, or long memory, are available in the literature for time series. Some approaches focus on the behavior of the auto-covariance function at large lags, others on the spectral density at low frequencies (see [25, Section 2.1] and the references therein). Separating short range from long range dependence is often made more natural within a particular class of models. For instance, for a Hilbert space-valued process  $Y = (Y_t)_{t \in \mathbb{Z}}$ , one may rely on a causal linear representation, namely

$$Y_t = \sum_{k=0}^{\infty} P_k \epsilon_{t-k}, \quad t \in \mathbb{Z}, \quad (3.10)$$

where  $\epsilon = (\epsilon_t)_{t \in \mathbb{Z}}$  is a centered white noise valued in the separable Hilbert space  $\mathcal{H}_0$  and  $(P_k)_{k \in \mathbb{Z}}$  is a sequence of  $\mathcal{L}_b(\mathcal{H}_0)$  operators. A sufficient condition for convergence of this series in  $\mathcal{M}(\Omega, \mathcal{F}, \mathcal{H}_0, \mathbb{P})$  is that  $\sum_{k=0}^{\infty} \|P_k\|_{\infty} < \infty$ , and this assumption is referred to as the *short range dependence* (or short memory) case, in contrast to *long range dependence* (long memory) case, for which  $\sum_{k=0}^{\infty} \|P_k\|_{\infty} = \infty$ , under which the convergence in (3.10) is no longer granted. The case where  $P_k = (k+1)^{-N}$  for some  $N \in \mathcal{N}(\mathcal{H}_0)$  is investigated in [12]. More precisely, let  $n$  and  $U$  be the singular value function and decomposition operator of  $N$  on  $\mathcal{G}_0 := L^2(\mathcal{V}, \mathcal{V}, \xi)$ . Assume that

$$h > \frac{1}{2} \quad \xi\text{-a.e.} \quad \text{and} \quad \int_{\mathcal{V}} \frac{\sigma_W^2(v)}{2h(v) - 1} \xi(dv) < \infty, \quad (3.11)$$

where  $h : v \mapsto \Re(n(v))$  and  $\sigma_W^2 : s \mapsto \mathbb{E}[|W(v, \cdot)|^2]$  with  $W = U\epsilon_0$ . Then, using the arguments of the proof of [12, Lemma A.1], one can show that, for all  $t \in \mathbb{Z}$ ,

$$Y_t = \sum_{k=0}^{\infty} (k+1)^{-N} \epsilon_{t-k} \quad (3.12)$$

converges in  $\mathcal{M}(\Omega, \mathcal{F}, \mathcal{H}_0, \mathbb{P})$ .

In [12, Theorem 2.1], the author also studies the partial sums of the process (3.12) and exhibits asymptotic properties which naturally extend the usual behavior observed for univariate long memory processes. In the following, we explain how the process (3.12) can be related to a FIARMA(D,0,0) process. First, we prove that Condition (3.11) also implies the existence of this FIARMA process.

**Lemma 3.5.** *Condition (3.11) implies  $\epsilon \in \mathcal{S}_{\text{FI}_D}(\Omega, \mathcal{F}, \mathbb{P})$  with  $D = \text{Id}_{\mathcal{H}_0} - N$ .*

We can now state a result which shows that the process defined by (3.12) is closely related to a FIARMA process up to a bounded operator  $C$  and to an additive short memory process.

**Proposition 3.6.** *Assume that (3.11) holds and define  $Y = (Y_t)_{t \in \mathbb{Z}}$  by (3.12). Then, there exists  $C \in \mathcal{L}_b(\mathcal{H}_0)$  and  $(\Delta_k)_{k \in \mathbb{N}} \in \mathcal{L}_b(\mathcal{H}_0)^{\mathbb{N}}$  with  $\sum_{k=0}^{\infty} \|\Delta_k\|_{\infty} < \infty$  such that*

$$F_{\text{FI}_D}(\epsilon) = CY + Z,$$

where  $Z$  is the short memory process defined, for all  $t \in \mathbb{Z}$ , by  $Z_t = \sum_{k=0}^{\infty} \Delta_k \epsilon_{t-k}$ .

## 4 Prediction and estimation

### 4.1 Main assumptions and preliminary result

We denote by  $\text{Leb}_{\mathbb{T}}$  the Lebesgue measure on  $(\mathbb{T}, \mathcal{B}(\mathbb{T}))$  divided by  $2\pi$ , so that for any locally integrable  $2\pi$ -periodic function  $g$ ,

$$\int g \, d\text{Leb}_{\mathbb{T}} = (2\pi)^{-1} \int_{\mathbb{T}} g(x) \, dx = (2\pi)^{-1} \int_0^{2\pi} g(x) \, dx.$$

Let  $\mathcal{H}_0$  be a separable Hilbert space,  $X = (X_t)_{t \in \mathbb{Z}}$  be a process defined on  $(\Omega, \mathcal{F}, \mathbb{P})$ , and valued in  $\mathcal{H}_0$ , and consider the following assumptions.

- (A-1) The process  $X$  is stationary and ergodic.
- (A-2) The process  $X$  is weakly stationary, and we denote by  $\nu_X$  its spectral operator measure.

Denote the discrete Fourier coefficients of  $X_1, \dots, X_n$  by

$$d_n^X(\lambda) = \frac{1}{\sqrt{n}} \sum_{k=1}^n X_k e^{-i\lambda k}, \quad \lambda \in \mathbb{T}, \quad (4.1)$$

and the periodogram by

$$I_n^X(\lambda) = d_n^X(\lambda) \otimes d_n^X(\lambda), \quad \lambda \in \mathbb{T}.$$

If  $X$  is not a centered process, one can use the empirical mean to center it, that is, in (4.1), replace  $X_k$  by  $X_{n,k}^c$  as defined in (1.3), in which case we denote the corresponding discrete Fourier coefficients and the corresponding periodogram by  $d_n^{X_n^c}$  and  $I_n^{X_n^c}$ , respectively.

The periodogram is related to the empirical covariance estimators through the following identity, for all  $s, t \in \mathbb{Z}$ ,

$$\hat{\Gamma}_n(s-t) = \int I_n^{X_n^c}(\lambda) e^{i(s-t)\lambda} \, \text{Leb}_{\mathbb{T}}(d\lambda) = \frac{1}{n} \sum_{\substack{1 \leq k, k' \leq n \\ k-k'=(s-t)}} X_{n,k}^c \otimes X_{n,k'}^c. \quad (4.2)$$

The integral in this equation can be interpreted as a sesquilinear functional  $\mathcal{Q}_{I_n^{X_n^c}}$  applied to exponential functions  $\lambda \mapsto e^{i s \lambda}$  on the left and  $\lambda \mapsto e^{i t \lambda}$  on the right, where, for any operator functions  $L, g$  and  $R$  defined on  $\mathbb{T}$ , we set

$$\mathcal{Q}_g(L, R) = \int L g R^{\text{H}} \, d\text{Leb}_{\mathbb{T}}. \quad (4.3)$$

Similarly, if  $\nu$  is a trace-class p.o.v.m. defined on  $(\mathbb{T}, \mathcal{B}(\mathbb{T}))$  and valued in  $\mathcal{S}_1^+(\mathcal{H}_0)$ , we set

$$\mathcal{Q}_{\nu}(L, R) = \int L \, d\nu \, R^{\text{H}}. \quad (4.4)$$

To ensure that these integrals are well defined, we assume that  $L$  and  $R$  are measurable bounded functions valued in  $\mathcal{L}_b(\mathcal{H}_0, \mathcal{G}_0)$ , with  $\mathcal{G}_0$  an additional separable Hilbert spaces. Namely, for any Banach space  $(\mathcal{E}, \|\cdot\|_{\mathcal{E}})$ , we further denote, by  $\mathbb{F}_b(\mathbb{T}, \mathcal{B}(\mathbb{T}), \mathcal{E})$  the set of bounded measurable functions from  $(\mathbb{T}, \mathcal{B}(\mathbb{T}))$  to  $(\mathcal{E}, \mathcal{B}(\mathcal{E}))$ , and we endow  $\mathbb{F}_b(\mathbb{T}, \mathcal{B}(\mathbb{T}), \mathcal{E})$  with the sup norm, which, for all  $L \in \mathbb{F}_b(\mathbb{T}, \mathcal{B}(\mathbb{T}), \mathcal{E})$ , we denote by

$$\sup(L) = \sup_{\lambda \in \mathbb{T}} \|L(\lambda)\|_{\mathcal{E}} .$$

Then, for all  $g$  valued in  $\mathcal{S}_1(\mathcal{H}_0)$  and  $L, R \in \mathbb{F}_b(\mathbb{T}, \mathcal{B}(\mathbb{T}), \mathcal{L}_b(\mathcal{H}_0, \mathcal{G}_0))$ , we have  $\mathcal{Q}_g(L, R) \in \mathcal{S}_1(\mathcal{G}_0)$  with

$$\|\mathcal{Q}_g(L, R)\|_1 \leq \sup(L) \sup(R) \int \|g\|_1 d\text{Leb}_{\mathbb{T}} , \quad (4.5)$$

and similarly, for a trace-class p.o.v.m.  $\nu$  valued in  $\mathcal{S}_1^+(\mathcal{H}_0)$ , we have  $\mathcal{Q}_{\nu}(L, R) \in \mathcal{S}_1(\mathcal{G}_0)$  with

$$\|\mathcal{Q}_{\nu}(L, R)\|_1 \leq \sup(L) \sup(R) \|\nu\|_1(\mathbb{T}) . \quad (4.6)$$

We denote by  $\mathbb{F}_{b,b}(\mathcal{H}_0, \mathcal{G}_0)$  the product vector space  $\mathbb{F}_b(\mathbb{T}, \mathcal{B}(\mathbb{T}), \mathcal{L}_b(\mathcal{H}_0, \mathcal{G}_0)) \times \mathbb{F}_b(\mathbb{T}, \mathcal{B}(\mathbb{T}), \mathcal{L}_b(\mathcal{H}_0, \mathcal{G}_0))$ , endowed with the max norm defined by

$$\|(L, R)\|_{b,b} := \max(\sup(L), \sup(R)) \quad \text{for all } (L, R) \in \mathbb{F}_{b,b}(\mathcal{H}_0, \mathcal{G}_0) .$$

For any two metric spaces  $(\mathcal{E}_1, d_1)$  and  $(\mathcal{E}_2, d_2)$ ,  $\mathcal{C}(\mathcal{E}_1, \mathcal{E}_2)$  denotes the space of continuous functions from  $\mathcal{E}_1$  to  $\mathcal{E}_2$ . If  $F$  and  $F_n$  are in  $\mathcal{C}(\mathcal{E}_1, \mathcal{E}_2)$  for all  $n \in \mathbb{N}$  and

$$\lim_{n \rightarrow \infty} \sup_{x \in \mathcal{E}_1} d_2(F_n(x), F(x)) = 0 ,$$

we say that  $(F_n)_{n \in \mathbb{N}}$  converges to  $F$  uniformly in  $\mathcal{C}(\mathcal{E}_1, \mathcal{E}_2)$ .

Using these definitions, immediate properties of the quadratic functionals  $\mathcal{Q}_{I_n^{X_n^c}}$  and  $\mathcal{Q}_{\nu_X}$  are summarized in the following proposition.

**Proposition 4.1.** *For any  $X_1, \dots, X_n$  in  $\mathcal{H}_0$ , the mapping  $\mathcal{Q}_{I_n^{X_n^c}}$  is well defined and belongs to  $\mathcal{C}(\mathbb{F}_{b,b}(\mathcal{H}_0, \mathcal{G}_0), \mathcal{S}_1(\mathcal{G}_0))$ . If **(A-2)** holds, we also have  $\mathcal{Q}_{\nu_X} \in \mathcal{C}(\mathbb{F}_{b,b}(\mathcal{H}_0, \mathcal{G}_0), \mathcal{S}_1(\mathcal{G}_0))$ .*

*Proof.* Observe that, under the given assumptions,  $X_{n,1}^c, \dots, X_{n,n}^c$  all are in  $\mathcal{H}_0$ , and so is  $d_n^{X_n^c}(\lambda)$ . Moreover,  $\left\| d_n^{X_n^c}(\lambda) \right\|_{\mathcal{H}_0}$  is bounded independently of  $\lambda$ . Consequently,  $I_n^{X_n^c}$  is valued in  $\mathcal{S}^1(\mathcal{H}_0)$  and its trace-norm is integrable over  $\mathbb{T}$ . The result on  $\mathcal{Q}_{I_n^{X_n^c}}$  thus follows from (4.5).

Under **(A-2)**,  $\nu_X$  is a trace-class p.o.v.m.  $\nu$  valued in  $\mathcal{S}_1^+(\mathcal{H}_0)$ , with  $\|\nu\|_1(\mathbb{T}) = \mathbb{E} \left[ \|X_0\|_{\mathcal{H}_0}^2 \right] < \infty$ . The result on  $\mathcal{Q}_{\nu_X}$  thus follows from (4.6).  $\square$

Our next result only exploits **(A-1)** and **(A-2)**, and is thus of independent interest. It is a uniform convergence result for integral quadratic functionals based on the periodogram. It applies to a parameterized pair of bounded operators functions defined on  $(\mathbb{T}, \mathcal{B}(\mathbb{T}))$  and valued in  $\mathcal{L}_b(\mathcal{H}_0, \mathcal{G}_0)$ . More precisely, for  $L$  and  $R$  in  $\mathcal{C}(\Theta \times \mathbb{T}, \mathcal{L}_b(\mathcal{H}_0, \mathcal{G}_0))$ , for all  $n \geq 1$ , we define  $\tilde{\mathcal{Q}}_{I_n^{X_n^c}}^{(L,R)} : \theta \mapsto \mathcal{Q}_{I_n^{X_n^c}}(L(\theta, \cdot), R(\theta, \cdot))$ . We also define  $\tilde{\mathcal{Q}}_{\nu_X}^{(L,R)} : \theta \mapsto \mathcal{Q}_{\nu_X}(L(\theta, \cdot), R(\theta, \cdot))$ , with  $\mathcal{Q}_{I_n^{X_n^c}}$  and  $\mathcal{Q}_{\nu_X}$  as in Proposition 4.1.

We can now state a first result on the convergence of the periodogram quadratic functional, in the case where the left and right operator functions' arrival spaces are finite-dimensional.

**Theorem 4.2.** *Let  $\mathcal{H}_0$  be a separable Hilbert space and  $\mathcal{G}_0$  be a finite dimensional space. Let  $X = (X_t)_{t \in \mathbb{Z}}$  be a process defined on  $(\Omega, \mathcal{F}, \mathbb{P})$  and valued in  $\mathcal{H}_0$  satisfying **(A-1)** and **(A-2)** and let  $(\Theta, \Delta)$  be a compact metric space. Let  $L$  and  $R$  in  $\mathcal{C}(\Theta \times \mathbb{T}, \mathcal{L}_b(\mathcal{H}_0, \mathcal{G}_0))$ . Then, we have*

$$\lim_{n \rightarrow \infty} \tilde{\mathcal{Q}}_{I_n^{X_n^c}}^{(L,R)} = \tilde{\mathcal{Q}}_{\nu_X}^{(L,R)} \quad \text{uniformly in } \mathcal{C}(\Theta, \mathcal{S}_1(\mathcal{G}_0)) , \quad \mathbb{P}\text{-a.s.} \quad (4.7)$$

We now consider the case where  $\mathcal{G}_0 = \mathcal{H}_0$  and  $\mathcal{H}_0$  is an infinite-dimensional separable Hilbert space. In order to obtain the convergence in  $\mathcal{S}_1(\mathcal{H}_0)$ , we will rely on an additional assumption in this case. To this end, for any sequence  $s = (s_k)_{k \in \mathbb{N}} \in [1, \infty)^{\mathbb{N}}$  and any orthonormal sequence  $(\phi_k)_{k \in \mathbb{N}}$  of  $\mathcal{H}_0$ , we set

$$\mathcal{H}_0^s = \left\{ x \in \overline{\text{Span}}^{\mathcal{H}_0}(\phi_k, k \in \mathbb{N}) : \sum_{k \in \mathbb{N}} s_k^2 \left| \langle x, \phi_k \rangle_{\mathcal{H}_0} \right|^2 < \infty \right\} . \quad (4.8)$$

A typical example of such spaces are the Sobolev spaces with index  $\alpha > 0$  where  $(\phi_k)_{k \in \mathbb{N}}$  is a well chosen Hilbert basis (i.e. orthonormal and complete in  $\mathcal{H}_0$ ) and  $s_k = (1+k)^\alpha$ . The space  $\mathcal{H}_0^s$  is a subspace of  $\mathcal{H}_0$  and is itself a separable Hilbert space endowed with the inner product

$$\langle x, y \rangle_{\mathcal{H}_0^s} = \sum_{k \in \mathbb{N}} s_k^2 \langle x, \phi_k \rangle_{\mathcal{H}_0} \overline{\langle y, \phi_k \rangle_{\mathcal{H}_0}}. \quad (4.9)$$

Setting  $\xi_k = s_k^{-1} \phi_k$  for all  $k \in \mathbb{N}$ , we note that  $(\xi_k)_{k \in \mathbb{N}}$  is a Hilbert basis of  $\mathcal{H}_0^s$ .

Using the space  $\mathcal{H}_0^s$  that we have just introduced, we have the following result for the infinite-dimensional case.

**Theorem 4.3.** *Let  $\mathcal{H}_0$  be an infinite-dimensional separable Hilbert space. Let  $X = (X_t)_{t \in \mathbb{Z}}$  be a process defined on  $(\Omega, \mathcal{F}, \mathbb{P})$  and valued in  $\mathcal{H}_0$  satisfying **(A-1)** and **(A-2)** and let  $(\Theta, \Delta)$  be a compact metric space. Let  $L$  and  $R$  in  $\mathcal{C}(\Theta \times \mathbb{T}, \mathcal{L}_b(\mathcal{H}_0))$ . Let  $s = (s_k)_{k \in \mathbb{N}} \in [1, \infty)^\mathbb{N}$  and  $(\phi_k)_{k \in \mathbb{N}}$  be an orthonormal sequence of  $\mathcal{H}_0$ . Define the Hilbert space  $\mathcal{H}_0^s$  by (4.8) and (4.9). We suppose that the three following assertions hold.*

- (i) *The sequence  $s = (s_k)_{k \in \mathbb{N}} \in [1, \infty)^\mathbb{N}$  is non-decreasing and goes to  $\infty$ .*
- (ii) *We have  $X_0 \in \mathcal{H}_0^s$   $\mathbb{P}$ -a.s. with  $\mathbb{E}[\|X_0\|_{\mathcal{H}_0^s}^2] < \infty$ .*
- (iii) *Defining  $L_s$  and  $R_s$  by  $L_s(\theta, \lambda) = L(\theta, \lambda)|_{\mathcal{H}_0^s}$  and  $R_s(\theta, \lambda) = R(\theta, \lambda)|_{\mathcal{H}_0^s}$  for all  $(\theta, \lambda) \in \Theta \times \mathbb{T}$ , we have  $L_s$  and  $R_s$  in  $\mathcal{C}(\Theta \times \mathbb{T}, \mathcal{L}_b(\mathcal{H}_0^s))$ .*

Then, the following convergence holds.

$$\lim_{n \rightarrow \infty} \tilde{Q}_{I_n^{X_n^c}}^{(L,R)} = \tilde{Q}_{\nu_X}^{(L,R)} \quad \text{uniformly in } \mathcal{C}(\Theta, \mathcal{S}_1(\mathcal{H}_0)), \quad \mathbb{P}\text{-a.s.} \quad (4.10)$$

In fact, as shown by Lemma 5.7 in Section 5.3, Assumptions **(A-1)** and **(A-2)** imply Conditions (i) and (ii) of Theorem 4.3 for a well chosen  $s$ . This fact is useful to prove in Theorem 1.1, by applying Theorem 4.3 for a specific choice of  $L$  and  $R$ , for which (iii) holds for any sequence  $s$ , see the proof of Theorem 1.1 in Section 5.4.

## 4.2 FIARMA prediction and estimation

A common tool for  $M$ -estimation for finite dimensional time series is the Whittle contrast, which relies on a Gaussian approximation of  $d_n^X$  as  $n \rightarrow \infty$ , hence suggesting to use a Gaussian likelihood contrast for  $d_n^X$ , based on its asymptotic covariance operator. We still have such an approximation for time series valued in a Hilbert space, see [3, Theorem 1]. However, using the Whittle approach directly in infinite dimension does not seem to be directly applicable. Indeed, Gaussian distributions are generally singular to each other in infinite dimension. In particular, the log determinant term of the noise covariance matrix appearing in the Whittle contrast (for example, see the first term in  $\bar{L}_N$  on Page 344 of [10]), is not well defined when this matrix is replaced by an infinite dimensional covariance operator.

There are two possible ways of circumventing this issue. The first one is to project the data on some finite-dimensional subspace for statistical inference and then study the behavior of estimators as the dimension of this subspace diverges. The second one is to work on a least square criterion, which does not include the optimization of the noise covariance operator. Here we investigate this second approach. We first derive the best one-step ahead predictor of a FIARMA process (see Theorem 4.4). We then show that, under some condition, such a predictor can be estimated from the data (see Theorem 4.6).

Recall that, for any integer  $d \in \mathbb{N}$ ,  $\mathcal{P}_d(\mathcal{H}_0)$  denotes the set of polynomials  $\mathfrak{p}$  of degree  $d$  with coefficients in  $\mathcal{L}_b(\mathcal{H}_0)$ , and such that  $\mathfrak{p}(0) = \text{Id}_{\mathcal{H}_0}$ . In the following, we further denote by  $\mathcal{P}_d^\dagger(\mathcal{H}_0)$ , the set of all  $\mathfrak{p} \in \mathcal{P}_d(\mathcal{H}_0)$ , which are invertible on the closed unit disk  $\bar{\mathbb{D}}$ .

**Theorem 4.4.** *Let  $\mathcal{H}_0$  be a separable Hilbert space and  $p, q$  be two non-negative integers. Let  $Y$  be an  $\mathcal{H}_0$ -valued FIARMA process, as in Definition 3.2, with long-memory operator  $D \in \mathcal{N}(\mathcal{H}_0)$ , MA polynomial  $\Theta \in \mathcal{P}_d^\dagger(\mathcal{H}_0)$ , AR polynomial  $\Phi \in \mathcal{P}_p^\dagger(\mathcal{H}_0)$  and an  $\mathcal{H}_0$ -valued centered white noise  $Z$ . Define, for all  $z \in \mathbb{C} \setminus [1, \infty)$ ,*

$$\Phi_{\Theta, \Phi, D}^\dagger(z) = \text{Id}_{\mathcal{H}_0} - \{\Theta(z)\}^{-1} \Phi(z) (1-z)^D. \quad (4.11)$$

Then,  $\lambda \mapsto \Phi_{\Theta, \Phi, D}^\dagger(e^{-i\lambda})$  belongs to  $\hat{\mathcal{H}}^Y$ , and, for all  $t \in \mathbb{Z}$ , the best linear predictor of  $Y_t$  given its past  $\{Y_s : s \leq t-1\}$  is given by

$$\text{proj} \left( Y_t | \mathcal{H}_{t-1}^Y \right) = \int_{\mathbb{T}} e^{i\lambda t} \Phi_{\Theta, \Phi, D}^\dagger(e^{-i\lambda}) \hat{Y}(d\lambda), \quad (4.12)$$

where  $\text{proj} \left( Y_t | \mathcal{H}_{t-1}^Y \right)$  denotes the orthogonal projection of  $Y_t$  onto the closed space

$$\mathcal{H}_{t-1}^Y = \overline{\text{span}}^{\mathcal{M}(\Omega, \mathcal{F}, \mathcal{H}_0, \mathbb{P})} (\text{P} Y_s, s = t-1, t-2, \dots, \text{P} \in \mathcal{L}_b(\mathcal{H}_0)) .$$

In the following we derive the best predictor among a collection of FIARMA predictors from a finite sample  $X_1, \dots, X_n$ . We will consider ARMA predictors positive long memory FIARMA predictors. More precisely, Define

$$\mathcal{N}^\dagger(\mathcal{H}_0) := \{D \in \mathcal{N}(\mathcal{H}_0) : \Re(D) \text{ is invertible}\}, \quad (4.13)$$

where  $\Re(D) = (D + D^H)/2$  denote the real part of  $D$ . Our main assumption on the FIARMA parameters is the following.

**(A-3)** The FIARMA parameters  $\aleph = (D_\theta, \Phi_\theta, \Theta_\theta)_{\theta \in \Theta}$  are valued in  $\mathcal{N}(\mathcal{H}_0) \times \mathcal{P}_p^\dagger(\mathcal{H}_0) \times \mathcal{P}_q^\dagger(\mathcal{H}_0)$  for some  $p, q \in \mathbb{N}^2$  and indexed by a compact metric space  $(\Theta, \Delta)$ . Moreover, the mappings  $\theta \mapsto D_\theta$ , (resp.  $(\theta, \lambda) \mapsto \Phi_\theta(e^{-i\lambda})$  and  $(\theta, \lambda) \mapsto \Theta_\theta(e^{-i\lambda})$ ) are continuous from  $\Theta$  (resp.  $\Theta \times \mathbb{T}$ ) to  $(\mathcal{L}_b(\mathcal{H}_0), \|\cdot\|_\infty)$ . In addition, one of the following assertions hold.

- (i) For all  $\theta \in \Theta$ ,  $D_\theta = 0$ .
- (ii) For all  $\theta \in \Theta$ ,  $D_\theta \in \mathcal{N}^\dagger(\mathcal{H}_0)$ .

In **(A-3)**, Condition (i) corresponds to using an ARMA predictor. Therefore, we will call  $\aleph$  an ARMA predictor model in this case. Condition (ii) corresponds to using a FIARMA predictor with *positive* long-memory. Therefore, we will call  $\aleph$  a positive FIARMA predictor model in this case.

Our goal is now to derive, based on a finite sample  $X_1, \dots, X_n$ , an approximation of the best possible  $\aleph$ -prediction of a weakly stationary process  $X$  taken among the ARMA or FIARMA predictors defined by the collection  $\aleph$  of **(A-3)**. We precise what we mean by this best prediction in the following result, for a centered weakly stationary process  $Y$ . We treat both the case where the model is well specified and the case where it is not. Recall that we say that the model  $\aleph$  of **(A-3)** is well-specified for  $Y$  when  $Y$  is indeed a FIARMA process with a FIARMA parameter  $(\theta, \Phi, D)$  among  $\aleph$ .

**Proposition 4.5** (Definition of  $\aleph$ -best prediction). *Let  $\mathcal{H}_0$  be an infinite-dimensional separable Hilbert space and  $p, q$  be two non-negative integers. Let  $Y = (Y_t)_{t \in \mathbb{Z}}$  be a centered weakly stationary process defined on  $(\Omega, \mathcal{F}, \mathbb{P})$  and valued in  $\mathcal{H}_0$  and let  $\aleph$  be a model satisfying **(A-3)**. Then, we have the following facts and definitions.*

- (i) For all  $\theta \in \Theta$ , there exists an absolutely summable  $\mathcal{L}_b(\mathcal{H}_0)$ -valued sequence  $(\text{P}_k^\dagger(\theta))_{k \geq 1}$  such that, for all  $\lambda \in \mathbb{T} \setminus \{0\}$ ,

$$\text{F}^\dagger(\theta, \lambda) := \text{Id}_{\mathcal{H}_0} - \sum_{k=1}^{\infty} \text{P}_k^\dagger(\theta) e^{-i\lambda k} = \left[ \Theta_\theta(e^{-i\lambda}) \right]^{-1} \Phi_\theta(e^{-i\lambda}) \left( 1 - e^{-i\lambda} \right)^{D_\theta}. \quad (4.14)$$

- (ii) For all  $t \in \mathbb{Z}$  and  $\theta \in \Theta$ , we can define

$$\hat{Y}_t(\theta) := \sum_{k=1}^{\infty} \text{P}_k^\dagger(\theta) Y_{t-k} \in \mathcal{H}_{t-1}^Y. \quad (4.15)$$

- (iii) The best  $\aleph$ -prediction quadratic risk of  $Y$  is defined as

$$\mathbb{E}^2(Y, \aleph) = \inf_{\theta \in \Theta} \mathbb{E} \left[ \left\| Y_t - \hat{Y}_t(\theta) \right\|_{\mathcal{H}_0}^2 \right]. \quad (4.16)$$

which does not depend on  $t$  by weak stationarity of  $Y$ .

(iv) The inf in (4.16) is attained in  $\Theta$  (hence is a minimum) and we call the argmin set the set of best  $\aleph$ -predictors for  $Y$ , denoted by

$$\Theta_Y^* := \left\{ \theta \in \Theta : \mathbb{E} \left[ \left\| Y_t - \hat{Y}_t(\theta) \right\|_{\mathcal{H}_0}^2 \right] = \mathbb{E}^2(Y, \aleph) \right\} \quad (4.17)$$

Then,  $\Theta_Y^*$  is a compact subset of  $\Theta$ .

- (v) If there exists  $\hat{Y}_t^* \in \mathcal{H}_{t-1}^Y$  such that the subset  $\{\hat{Y}_t(\theta) : \theta \in \Theta_Y^*\}$  of  $\mathcal{H}_{t-1}^Y$  reduces to the singleton  $\hat{Y}_t^*$ , we call  $\hat{Y}_t^*$  the best  $\aleph$ -predictor of  $Y_t$ . Otherwise we say that the best  $\aleph$ -predictor of  $Y_t$  is not well defined.
- (vi) When the best  $\aleph$ -predictor of  $Y_t$  is well defined for one  $t$  it is well defined for all  $t$ . Moreover, in this case, there exists a set of probability one on which, for all  $t \in \mathbb{Z}$  and  $\theta \in \Theta_Y^*$ ,  $\hat{Y}_t^* = \hat{Y}_t(\theta)$ .
- (vii) In the well-specified case, the best  $\aleph$ -predictor  $\hat{Y}_t^*$  is always well defined and coincides with the best predictor in  $\mathcal{H}_{t-1}^Y$ , that is,

$$\mathbb{E}^2(Y, \aleph) = \inf_{V \in \mathcal{H}_{t-1}^Y} \mathbb{E} \left[ \left\| Y_t - V \right\|_{\mathcal{H}_0}^2 \right], \quad (4.18)$$

$$\hat{Y}_t^* = \text{proj} \left( Y_t \mid \mathcal{H}_{t-1}^Y \right). \quad (4.19)$$

The next results shows how to estimate a predictor which converges to the best predictor that we have just introduced. We now introduce our estimation procedure.

Using  $F_\theta^\dagger$  defined by (4.14) and the periodogram  $I_n^{X_n^c}$  defined in Section 4.1, we consider a sequence of estimators  $(\hat{\theta}_n)_{n \in \mathbb{N}}$  satisfying

$$\limsup_{n \rightarrow \infty} \left( \Lambda_n(\hat{\theta}_n) - \inf_{\theta \in \Theta} \Lambda_n(\theta) \right) = 0, \quad (4.20)$$

where, for all  $n \in \mathbb{N}$  and  $\theta \in \Theta$ ,

$$\Lambda_n(\theta) := \text{Tr} \left( \tilde{Q}_{I_n^{X_n^c}}^{(F_\theta^\dagger, F_\theta^\dagger)}(\theta) \right) = \text{Tr} \left( \int F_\theta^\dagger(\theta, \lambda) I_n^{X_n^c}(\lambda) \left( F_\theta^\dagger(\theta, \lambda) \right)^H d\text{Leb}_T \right). \quad (4.21)$$

Finally, define a predictor of  $X_{n+1}$  from the sample  $X_1, \dots, X_n$  associated to the estimator  $\hat{\theta}_n$  by

$$\hat{X}_{n+1, n} = \frac{1}{n} \sum_{k=1}^n X_k + \sum_{k=1}^n P_k^\dagger(\hat{\theta}_n) X_{n, n+1-k}^c,$$

where  $P_k^\dagger(\theta)$  is defined in Proposition 4.5. Note that the predictor  $\hat{X}_{n+1, n}$  can be written as

$$m + \sum_{k=1}^n P_k^\dagger(\theta) (X_{n+1-k} - m)$$

by taking  $m \in \mathcal{H}_0$  and  $\theta \in \Theta$  equal to  $\frac{1}{n} \sum_{k=1}^n X_k$  and  $\hat{\theta}_n$ , respectively. In the following theorem, defining, for all  $n \geq 1$ ,  $m \in \mathcal{H}_0$  and  $\theta \in \Theta$ , the quadratic prediction risk of a predictor of this form by

$$\mathbb{E}_{X, n}^2(m, \theta) = \mathbb{E} \left[ \left\| X_{n+1} - \left( m + \sum_{k=1}^n P_k^\dagger(\theta) (X_{n+1-k} - m) \right) \right\|_{\mathcal{H}_0}^2 \right], \quad (4.22)$$

we show that, as a predictor of  $X_{n+1}$ ,  $\hat{X}_{n+1, n}$  asymptotically achieves the same prediction risk as the optimal risk for predicting the centered process  $Y = X - \mathbb{E}[X_0]$  from its past.

**Theorem 4.6.** *Let  $\mathcal{H}_0$  be an infinite-dimensional separable Hilbert space and  $p, q$  be two non-negative integers. Let  $X = (X_t)_{t \in \mathbb{Z}}$  be a process defined on  $(\Omega, \mathcal{F}, \mathbb{P})$  and valued in  $\mathcal{H}_0$  satisfying **(A-1)** and **(A-2)**. Let  $\aleph$  be a model satisfying **(A-3)** with compact parameter metric space  $(\Theta, \Delta)$ . Let  $s = (s_k)_{k \in \mathbb{N}} \in [1, \infty)^\mathbb{N}$  and  $(\phi_k)_{k \in \mathbb{N}}$  be an orthonormal sequence of  $\mathcal{H}_0$ . Define the Hilbert space  $\mathcal{H}_0^s$  by (4.8) and (4.9). We suppose that (i) and (ii) of Theorem 4.3 hold and that*

(iii) Defining, for all  $z \in \mathbb{C}$ ,  $D_{\theta,s}$ ,  $\mathbb{f}_{\theta,s}(z)$  and  $\mathbb{t}_{\theta,s}(z)$  the restrictions of  $D_\theta$ ,  $\mathbb{f}_\theta(z)$  and  $\mathbb{t}_\theta(z)$  to  $\mathcal{H}_0^s$ , we have that  $\theta \mapsto D_{\theta,s}$ ,  $(\theta, \lambda) \mapsto \mathbb{f}_{\theta,s}(e^{-i\lambda})$  and  $(\theta, \lambda) \mapsto \mathbb{t}_{\theta,s}(e^{-i\lambda})$  are continuous from  $\Theta$  (resp.  $\Theta \times \mathbb{T}$ ) to  $(\mathcal{L}_b(\mathcal{H}_0^s), \|\cdot\|_\infty)$ . Under **(A-3)(ii)**, assume, in addition, that, for all  $\theta \in \Theta$ ,  $D_{\theta,s} \in \mathcal{N}^\dagger(\mathcal{H}_0^s)$ .

Finally, let  $(\hat{\theta}_n)_{n \in \mathbb{N}}$  be a sequence of estimators satisfying (4.20). Then, we have

$$\lim_{n \rightarrow \infty} \Delta \left( \hat{\theta}_n, \Theta_Y^* \right) = 0, \quad \mathbb{P}\text{-a.s.}, \quad (4.23)$$

$$\lim_{n \rightarrow \infty} \mathbb{E}_{X,n}^2 \left( \frac{1}{n} \sum_{k=1}^n X_k, \hat{\theta}_n \right) = \mathbb{E}^2(Y, \aleph), \quad \mathbb{P}\text{-a.s.} \quad (4.24)$$

where  $Y = (Y_t)_{t \in \mathbb{Z}}$  denotes the centered process defined by  $Y_t = X_t - \mathbb{E}[X_t]$ ,  $\Theta_Y^*$  is defined by (4.17),  $\mathbb{E}_{X,n}^2$  in (4.22), and  $\mathbb{E}^2(Y, \aleph)$  in (4.16).

Moreover, if  $\hat{Y}_t^*$  is well defined (as defined in Proposition 4.5 (v)), we further have

$$\limsup_{n \rightarrow \infty} \mathbb{E} \left[ \left\| X_{n+1} - \hat{X}_{n+1,n} \right\|_{\mathcal{H}_0}^2 \right] \leq \mathbb{E}^2(Y, \aleph). \quad (4.25)$$

## 5 Postponed proofs

### 5.1 Proofs of Section 3.2

#### 5.1.1 Proofs of Lemma 3.2 and Theorem 3.3

Lemma 3.2 is used to introduce the definition of joint kernels as in Definition 3.3.

**Proof of Lemma 3.2.** Let  $(\phi_i)_{0 \leq i < N}$  denote a Hilbert basis of  $L^2(\mathcal{V}, \mathcal{V}, \xi)$ , assumed to be of dimension  $N \in \{1, 2, \dots, \infty\}$ . Define  $\mathcal{X}_n : (v, v'; \lambda) \mapsto \sum_{0 \leq i, j \leq n} \phi_i^H K(\lambda) \phi_j \phi_i(v) \bar{\phi}_j(v')$  on  $\mathcal{V}^2 \times \mathbb{T}$  and, for all  $\epsilon > 0$ ,  $N_\epsilon : \lambda \mapsto \inf \left\{ n < N : \sum_{i \text{ or } j > n} |\phi_i^H K(\lambda) \phi_j|^2 \leq \epsilon \right\}$  on  $\mathbb{T}$ . Note that, for all  $\lambda \in \mathbb{T}$ ,  $N_\epsilon(\lambda)$  is well defined and finite since  $\sum_{0 \leq i, j < N} |\phi_i^H K(\lambda) \phi_j|^2 = \|K(\lambda)\|_2 < \infty$ . Now let us define, for all  $v, v' \in \mathcal{V}$  and  $\lambda \in \mathbb{T}$ ,  $\mathcal{X}(v, v'; \lambda) := \lim_{n \rightarrow \infty} \mathcal{X}_{N_{2^{-n}}(\lambda)}(v, v'; \lambda)$  whenever this limit exists in  $\mathbb{C}$  and set  $\mathcal{X}(v, v'; \lambda) = 0$  otherwise. Since  $(\phi_k \otimes \bar{\phi}_{k'})_{0 \leq k, k' < N}$  is a Hilbert basis of  $L^2(\mathcal{V}^2, \mathcal{V}^{\otimes 2}, \xi^{\otimes 2})$ , we immediately have that, for any  $\lambda \in \Lambda$ ,  $\mathcal{X}_{N_{2^{-n}}(\lambda)}(\cdot, \cdot; \lambda)$  converges in the sense of this  $L^2$  space to  $\sum_{0 \leq i, j < N} \phi_i^H K(\lambda) \phi_j \phi_i \otimes \bar{\phi}_j$ , and thus, this limit must be equal to  $\mathcal{X}(\cdot, \cdot; \lambda)$ ,  $\xi^{\otimes 2}$ -a.e. It follows that, that for any  $\lambda \in \Lambda$ , for all  $i, j \in \mathbb{N}$ ,  $\int \mathcal{X}(v, v'; \lambda) \bar{\phi}_i(v) \phi_j(v') \xi(dv) \xi(dv') = \phi_i^H K(\lambda) \phi_j$ , which gives that  $K(\lambda)$  is an integral operator associated to the kernel  $\mathcal{X}(\cdot, \cdot; \lambda)$ . Since  $(v, v', \lambda) \mapsto \mathcal{X}(v, v'; \lambda)$  is measurable by definition, this concludes the proof of the existence of the  $\Lambda$ -joint kernel of  $K$ . If, moreover,  $K \in \mathcal{L}^2(\Lambda, \mathcal{A}, \mathcal{S}_2(\mathcal{H}_0), \mu)$ , then  $\mathcal{X}_n$  converges in  $L^2(\mathcal{V}^2 \times \Lambda, \mathcal{V}^{\otimes 2} \otimes \mathcal{A}, \xi^{\otimes 2} \otimes \mu)$  and the limit must be equal to  $\mathcal{X}$ ,  $\xi^{\otimes 2} \otimes \mu$ -a.e. since for each  $\lambda \in \Lambda$ ,  $\mathcal{X}_n(\cdot, \cdot; \lambda)$  converges to  $\mathcal{X}(\cdot, \cdot; \lambda)$  in  $L^2(\mathcal{V}^2, \mathcal{V}^{\otimes 2}, \xi^{\otimes 2})$ . Hence, we get that  $\mathcal{X} \in L^2(\mathcal{V}^2 \times \Lambda, \mathcal{V}^{\otimes 2} \otimes \mathcal{A}, \xi^{\otimes 2} \otimes \mu)$ .  $\square$

We now prove Theorem 3.3.

**Proof of Theorem 3.3.** We assume without loss of generality that  $\mu(\{0\}) = 0$  (since it affects none of the given assertions). By Proposition 2.1, Assertion (i) is equivalent to

$$\int_{\mathbb{T}} \left\| (1 - e^{-i\lambda})^{-D} g_X(\lambda) \left[ (1 - e^{-i\lambda})^{-D} \right]^H \right\|_1 \mu(d\lambda) < \infty. \quad (5.1)$$

Using the singular values decomposition (3.5), and since  $U$  is unitary from  $\mathcal{H}_0$  to  $L^2(\mathcal{V}, \mathcal{V}, \xi)$ , we get that, for all  $\lambda \in \mathbb{T} \setminus \{0\}$ ,  $\left\| (1 - e^{-i\lambda})^{-D} g_X(\lambda) \left[ (1 - e^{-i\lambda})^{-D} \right]^H \right\|_1 = \left\| U^H M_{(1-e^{-i\lambda})^{-d}} U g_X(\lambda) U^H M_{(1-e^{-i\lambda})^{-d}} U \right\|_1 = \left\| M_{(1-e^{-i\lambda})^{-d}} h(\lambda) \right\|_2^2$ . Hence (5.1) holds if and only if  $\int_{\mathbb{T}} \left\| M_{(1-e^{-i\lambda})^{-d}} h(\lambda) \right\|_2^2 \mu(d\lambda) < \infty$ , which, using the  $\mathbb{T}$ -joint kernel  $\mathcal{h}$  of  $h$ , reads  $\int \left| (1 - e^{-i\lambda})^{-d(v)} \mathcal{h}(v, v'; \lambda) \right|^2 \xi(dv) \xi(dv') \mu(d\lambda) < \infty$ . Applying Lemma A.1 to  $z = -d(v)$ , since  $d$  is a  $\mu$ -essentially bounded function, we get that Assertion (i) is equivalent to

$$\int_{\mathcal{V}^2 \times (-\pi, \pi]} |\lambda|^{-2\Re(d(v))} \left| \mathcal{h}(v, v'; \lambda) \right|^2 \xi(dv) \xi(dv') \mu(d\lambda) < \infty. \quad (5.2)$$



Using that  $|\lambda|^{2\Re-(d(v))}$  is bounded over  $\lambda \in (-\pi, \pi]$  and that

$$\int |\mathcal{K}(v, v'; \lambda)|^2 \xi(dv)\xi(dv')\mu(d\lambda) = \int \|h(\lambda)\|_2^2 \mu(d\lambda) < \infty,$$

Condition (5.2) is equivalent to Assertion (iii). Using that, for any  $\eta \in (0, \pi)$ ,  $|\lambda|^{-2\Re(d(v))}$  is bounded independently of  $v$  on  $\lambda \in (-\pi, \pi] \setminus (-\eta, \eta)$ , we also get that Condition (5.2) is equivalent to Assertion (ii).  $\square$

### 5.1.2 Proof of Theorem 3.4

We start with a useful result on ARMA processes.

**Lemma 5.1.** *Let  $\mathcal{H}_0$  be a separable Hilbert space and  $X$  be an ARMA( $p, q$ ) process defined by  $\hat{X}(d\lambda) = [\hat{\Phi}(e^{-i\lambda})]^{-1}\hat{\Theta}(e^{-i\lambda})\hat{Z}(d\lambda)$  with  $\hat{\Theta} \in \mathcal{P}_q(\mathcal{H}_0)$ ,  $\hat{\Phi} \in \mathcal{P}_p^*(\mathcal{H}_0)$  and  $Z$  an  $\mathcal{H}_0$ -valued white noise with covariance operator  $\Sigma$ . Then, there exists  $\eta \in (0, \pi)$  such that*

$$\sum_{n \in \mathbb{N}} \frac{\eta^n}{n!} \|\mathbb{P}_n(\hat{\Phi}, \hat{\Theta})\|_\infty < \infty, \quad (5.3)$$

where  $\mathbb{P}_n(\hat{\Phi}, \hat{\Theta})$  is defined in (3.8). Moreover, for Leb-a.e.  $\lambda \in (-\eta, \eta)$ , we have

$$g_X(\lambda) = h(\lambda) [h(\lambda)]^H \quad \text{with} \quad h(\lambda) = \sum_{n \in \mathbb{N}} \frac{(-i\lambda)^n}{n!} \mathbb{P}_n(\hat{\Phi}, \hat{\Theta}) \Sigma^{1/2}, \quad (5.4)$$

where  $h$  can be seen as a power series valued in  $\mathcal{S}_2(\mathcal{H}_0)$  with a convergence radius at least equal to  $\eta$ .

*Proof.* Since  $z \mapsto [\hat{\Phi}(z)]^{-1}\hat{\Theta}(z)$  is holomorphic in an open ring containing  $\mathbb{U}$  and the exponential function is holomorphic on  $\mathbb{C}$ , by [13, Theorem 1.8.5], there exists  $\eta > 0$  such that (5.3) holds and  $[\hat{\Phi}(e^z)]^{-1}\hat{\Theta}(e^z)$  coincides with the  $\mathcal{L}_b(\mathcal{H}_0)$ -valued power series  $\sum_{n=0}^{\infty} \mathbb{P}_n(\hat{\Phi}, \hat{\Theta}) z^n/n!$  on the set  $\{z \in \mathbb{C} : |z| \leq \eta\}$ .

Finally, we observe that  $g_X = hh^H$  with  $h(\lambda) = [\hat{\Phi}(e^{-i\lambda})]^{-1}\hat{\Theta}(e^{-i\lambda})\Sigma^{1/2}$  and the given expression of  $h$  in (5.4) follows from (5.3) and the usual bound

$$\left\| \mathbb{P}_n(\hat{\Phi}, \hat{\Theta}) \Sigma^{1/2} \right\|_2 \leq \|\mathbb{P}_n(\hat{\Phi}, \hat{\Theta})\|_\infty \left\| \Sigma^{1/2} \right\|_2 = \|\mathbb{P}_n(\hat{\Phi}, \hat{\Theta})\|_\infty \|\Sigma\|_1^{1/2}.$$

This concludes the proof.  $\square$

**Proof of Theorem 3.4.** Before proving the claimed implications, we start with some preliminary facts that follow from Lemma 5.1, Lemma B.2 and Theorem 3.3. First observe that the process  $UX = (UX_t)_{t \in \mathbb{Z}}$  is the  $\mathcal{G}_0$ -valued ARMA( $p, q$ ) process defined by  $\widehat{UX}(d\lambda) = [\hat{\Phi}(e^{-i\lambda})]^{-1}\hat{\Theta}(e^{-i\lambda})\widehat{UZ}(d\lambda)$ , where  $\hat{\Theta} := U\hat{\Theta}U^H \in \mathcal{P}_q(\mathcal{G}_0)$  and  $\hat{\Phi} := U\hat{\Phi}U^H \in \mathcal{P}_p^*(\mathcal{G}_0)$ , and  $UZ = (UZ_t)_{t \in \mathbb{Z}}$  is a  $\mathcal{G}_0$ -valued white noise. Then, applying Lemma 5.1 with  $\mu$  as the Lebesgue measure, we get that, for some  $\eta > 0$ ,  $\nu_{UX}$  has density  $h(\lambda)[h(\lambda)]^H$  on  $(-\eta, \eta)$  with  $h$  a power series valued in  $\mathcal{S}_2(\mathcal{G}_0)$  with radius of convergence at least  $\eta > 0$ ,

$$h(\lambda) = \sum_{n \in \mathbb{N}} \frac{(-i\lambda)^n}{n!} U \mathbb{P}_n(\hat{\Phi}, \hat{\Theta}) \Sigma^{1/2} U^H. \quad (5.5)$$

Now, define, for any  $\eta' \in (0, \eta)$ ,

$$I(\eta') := \int_{\mathbb{V}^2 \times (-\eta', \eta')} |\lambda|^{-2\Re(d(v))} |\mathcal{K}(v, v'; \lambda)|^2 \xi(dv)\xi(dv') \frac{d\lambda}{2\pi},$$

where  $\mathcal{K}$  is the  $\mathbb{T}$ -joint kernel of  $h$  in (5.5). By Theorem 3.3, Assertion (i) holds if, and only if, there exists  $\eta' \in (0, \eta)$  such that  $I(\eta') < \infty$  which is itself equivalent to having  $I(\eta') < \infty$  for all  $\eta' \in (0, \eta)$ . Using (5.5), we have

$$I(\eta') = \int_{\mathbb{V}^2 \times (-\eta', \eta')} |\lambda|^{-2\Re(d(v))} \left| \sum_{n \in \mathbb{N}} \frac{(-i\lambda)^n}{n!} \mathcal{K}_n(v, v') \right|^2 \xi(dv)\xi(dv') \frac{d\lambda}{2\pi}, \quad (5.6)$$

where  $\ell_n$  denotes the kernel of  $UP_n(\Phi, \Theta) \Sigma^{1/2} U^H \in \mathcal{S}_2(\mathcal{G}_0)$ . In particular we have by Lemma B.2 that

$$\sigma_n(v) = (\mathbb{E} [|W_n(v, \cdot)|^2])^{1/2} = \|\ell_n(v, \cdot)\|_{\mathcal{G}_0}. \quad (5.7)$$

Denote

$$I_n(\eta') := \int_{\mathbb{V}^2 \times (-\eta', \eta')} |\lambda|^{2n-2\Re(d(v))} |\ell_n(v, v')|^2 \xi(dv) \xi(dv') \frac{d\lambda}{2\pi}, \quad (5.8)$$

$$\bar{d} := \sup(\Re(d)) \quad \text{and} \quad \bar{m} := \inf \{m \in \mathbb{N} : m > \bar{d} - 1/2\}.$$

Note that  $\bar{d}$  and  $\bar{m}$  are finite since  $d$  is bounded. Defining, moreover,

$$\bar{I}(\eta') := \int_{\mathbb{V}^2 \times (-\eta', \eta')} |\lambda|^{-2\Re(d(v))} \left| \sum_{0 \leq n < \bar{m}} \frac{(-i\lambda)^n}{n!} \ell_n(v, v') \right|^2 \xi(dv) \xi(dv') \frac{d\lambda}{2\pi}$$

and  $R(\eta') := \sum_{n=\bar{m}}^{\infty} \frac{I_n^{1/2}(\eta')}{n!},$

and applying the Minkowski inequality in (5.6), for any  $\eta' \in (0, \eta)$ , we have that, if  $R(\eta') < \infty$ ,

$$I(\eta') < \infty \Leftrightarrow \bar{I}(\eta') < \infty.$$

Let us pick  $\eta' \in (0, 1 \wedge \eta)$ . Then, for all  $\lambda \in (-\eta', \eta')$  and  $n \in \mathbb{N}$ , we have  $|\lambda|^{2n-2\Re(d(v))} \leq |\lambda|^{2n-2\bar{d}}$  and thus, for all  $n \geq \bar{m}$ ,

$$I_n(\eta') \leq \frac{\eta'^{(1+2n-2\bar{d})}}{\pi(1+2n-2\bar{d})} \int_{\mathbb{V}^2} |\ell_n(v, v')|^2 \xi(dv) \xi(dv')$$

$$= \frac{\eta'^{(1+2n-2\bar{d})}}{\pi(1+2n-2\bar{d})} \left\| P_n(\Phi, \Theta) \Sigma^{1/2} \right\|_2^2.$$

Using that  $\left\| P_n(\Phi, \Theta) \Sigma^{1/2} \right\|_2 \leq \|P_n(\Phi, \Theta)\|_{\infty} \|\Sigma\|_1^{1/2}$  and the bound (5.3) of Lemma 5.1, we get that  $R(\eta') < \infty$ . We thus conclude that Assertion (i) is equivalent to have, for some  $\eta' \in (0, 1 \wedge \eta)$ ,

$$\bar{I}(\eta') < \infty. \quad (5.9)$$

Next, we show that, for any  $\eta' \in (0, 1)$ , Condition (ii) is, in fact, equivalent to have

$$I_n(\eta') < \infty \quad \text{for all } n \in \mathbb{N}. \quad (5.10)$$

Indeed, integrating w.r.t.  $v'$  and  $\lambda$  in the definition of  $I_n(\eta')$  in (5.8), we have that, for all  $n \in \mathbb{N}$ ,

$$I_n(\eta') = \int_{\{\Re(d) < n+1/2\}} \frac{\eta'^{(1+2n-2\Re(d(v)))} \sigma_n^2(v)}{1+2n-2\Re(d(v))} \frac{\xi(dv)}{\pi},$$

if  $\Re(d) < n+1/2$   $\xi$ -a.e. on  $\{\sigma_n > 0\}$ , or is equal to  $\infty$  otherwise. Since  $d$  is bounded on  $\mathbb{V}$ , so is  $\eta'^{(1+2n-2\Re(d(v)))}$  on  $v \in \mathbb{V}$  and we conclude that Condition (ii) is equivalent to (5.10).

We are now ready to prove each implication of the claimed equivalence successively.

**Proof of (i)  $\Leftarrow$  (ii).** This is now trivial, since applying the Minkowski inequality in the integral defining  $\bar{I}$  in (5.9) and the definition of  $I_n$  in (5.8), we immediately see that Condition (5.9) is implied by (5.10).

**Proof of (i)  $\Rightarrow$  (ii).** The proof of this implication is a bit more complex. The first step is to prove that Assertion (i) implies

$$\text{for all } n \in \mathbb{N}, \Re(d) < n+1/2 \quad \xi\text{-a.e. on } \{\sigma_n > 0\}. \quad (5.11)$$

Then, we show that it must also imply (3.9) for all  $n \in \mathbb{N}$  in a second and last step.

**Step 1.** Suppose that (5.11) does not hold; let us show that Assertion (i) cannot hold. Since it is equivalent to (5.9), it is sufficient to show that  $\bar{I}(\eta') = \infty$  for any arbitrary  $\eta' > 0$ . Let  $\underline{m}$  be the smallest  $n \in \mathbb{N}$  for which  $\xi(\{\Re(d) \geq n+1/2\} \cap \{\sigma_n > 0\}) > 0$ . Note that by their mere definitions, we have  $\underline{m} < \bar{m}$ . In addition, for all  $0 \leq n < \underline{m}$ , we have

$$\xi(\{\Re(d) \geq \underline{m}+1/2\} \cap \{\sigma_n > 0\}) \leq \xi(\{\Re(d) \geq n+1/2\} \cap \{\sigma_n > 0\}) = 0.$$

Hence,  $\ell_n(v, v') = 0$  for  $\xi^{\otimes 2}$ -a.e.  $(v, v') \in \{\Re(d) \geq \underline{m} + 1/2\} \times \mathbf{V}$ . Now we have, using the definition of  $\bar{I}(\eta')$  in (5.9) and what we just deduced,

$$\begin{aligned} \bar{I}(\eta') &\geq \int_{\{\Re(d) \geq \underline{m} + 1/2\} \times \mathbf{V} \times (-\eta', \eta')} |\lambda|^{-2\Re(d(v))} \left| \sum_{0 \leq n < \bar{m}} \frac{(-i\lambda)^n}{n!} \ell_n(v, v') \right|^2 \xi(dv) \xi(dv') \frac{d\lambda}{2\pi} \\ &= \int_{\{\Re(d) \geq \underline{m} + 1/2\} \times \mathbf{V} \times (-\eta', \eta')} |\lambda|^{-2\Re(d(v))} \left| \sum_{\underline{m} \leq n < \bar{m}} \frac{(-i\lambda)^n}{n!} \ell_n(v, v') \right|^2 \xi(dv) \xi(dv') \frac{d\lambda}{2\pi}. \end{aligned}$$

Note that, for all  $(v, v') \in \mathbf{V}^2$  and  $\lambda \in \mathbb{R}$ , we have

$$|\lambda|^{-2\Re(d(v))} \left| \sum_{\underline{m} \leq n < \bar{m}} \frac{(-i\lambda)^n}{n!} \ell_n(v, v') \right|^2 = |\lambda|^{2\underline{m} - 2\Re(d(v))} \left( \frac{|\ell_{\underline{m}}(v, v')|^2}{(\underline{m}!)^2} + o(1) \right),$$

where  $o$ -term tends to 0 as  $\lambda \rightarrow 0$ . We get that, for all  $(v, v') \in (\{\Re(d) \geq \underline{m} + 1/2\} \times \mathbf{V}) \cap \{|\ell_{\underline{m}}|^2 > 0\}$  and  $\eta' > 0$ ,

$$\int_{(-\eta', \eta')} |\lambda|^{-2\Re(d(v))} \left| \sum_{\underline{m} \leq n < \bar{m}} \frac{(-i\lambda)^n}{n!} \ell_n(v, v') \right|^2 \frac{d\lambda}{2\pi} = \infty.$$

With the previous lower bound on  $\bar{I}(\eta')$ , we deduce that  $\bar{I}(\eta') = \infty$  if

$$\xi^{\otimes 2}(\{\Re(d) \geq \underline{m} + 1/2\} \times \mathbf{V}) \cap \{|\ell_{\underline{m}}|^2 > 0\} > 0.$$

By definition of  $\sigma_n$  in (5.7), we have, for all  $v \in \mathbf{V}$ ,

$$g(v) := \int_{\mathbf{V}} \mathbb{1}_{\{|\ell_{\underline{m}}(v, v')|^2 > 0\}} \xi(dv') > 0 \Leftrightarrow \sigma_{\underline{m}}(v) > 0.$$

Hence,

$$\xi^{\otimes 2}(\{\Re(d) \geq \underline{m} + 1/2\} \times \mathbf{V}) \cap \{|\ell_{\underline{m}}|^2 > 0\} = \int_{\{\Re(d) \geq \underline{m} + 1/2\}} g \, d\xi$$

is positive if and only if  $\xi(\{\Re(d) \geq \underline{m} + 1/2\} \cap \{\sigma_{\underline{m}} > 0\}) > 0$ , which is true by definition of  $\underline{m}$ . This concludes the first step.

**Step 2.** Suppose now that (5.11) does hold but (3.9) does not hold for all  $n \in \mathbb{N}$  and let us show again that Assertion (i) cannot hold. Let us define  $\tilde{m}$  as the smallest  $n \in \mathbb{N}$  such that (3.9) does not hold. Again by definition of  $\bar{m}$ , we must have  $\tilde{m} < \bar{m}$ , since (3.9) holds for  $n = \bar{m}$  by definition of  $\bar{m}$ . Take now an arbitrary  $\eta' \in (0, 1 \wedge \eta)$ . We have shown in the preamble of the proof that if (5.11) is satisfied, then (3.9) is equivalent to  $I_n(\eta') < \infty$ . Hence, we can also see  $\tilde{m}$  as the smallest  $n \in \mathbb{N}$  such that  $I_n(\eta') = \infty$ . Thus, we have  $I_k(\eta') < \infty$  for all  $0 \leq k < \tilde{m}$  and  $I_{\tilde{m}}(\eta') = \infty$ . It follows that Assertion (i) is not only equivalent to having  $\bar{I}(\eta') < \infty$  as in (5.9) but also to the condition

$$\tilde{I}(\eta') := \int_{\mathbf{V}^2 \times (-\eta', \eta')} |\lambda|^{-2\Re(d(v))} \left| \sum_{\tilde{m} \leq n < \bar{m}} \frac{(-i\lambda)^n}{n!} \ell_n(v, v') \right|^2 \xi(dv) \xi(dv') \frac{d\lambda}{2\pi} < \infty. \quad (5.12)$$

Now, we observe that

$$\tilde{I}(\eta') \geq \int_{\{\Re(d) < \tilde{m} + 1/2\} \times \mathbf{V} \times (-\eta', \eta')} |\lambda|^{-2\Re(d(v))} \left| \sum_{\tilde{m} \leq n < \bar{m}} \frac{(-i\lambda)^n}{n!} \ell_n(v, v') \right|^2 \xi(dv) \xi(dv') \frac{d\lambda}{2\pi}.$$

Therefore, to conclude that  $\tilde{I}(\eta') = \infty$  (implying that Assertion (i) does not hold), by the Minkowski inequality, it is sufficient to show that

$$\tilde{I}_{\tilde{m}}(\eta') = \infty \quad \text{and, for all } n > \tilde{m}, \quad \tilde{I}_n(\eta') < \infty, \quad (5.13)$$

where, for all  $n \in \mathbb{N}$ , we denoted

$$\tilde{I}_n(\eta') := \int_{\{\Re(d) < \tilde{m} + 1/2\} \times \mathcal{V} \times (-\eta', \eta')} |\lambda|^{2n - 2\Re(d(v))} |\#_n(v, v')|^2 \xi(dv) \xi(dv') \frac{d\lambda}{2\pi}.$$

For all  $n \geq \tilde{m}$  we have, as in the previous computation of  $I_n$  that

$$\tilde{I}_n(\eta') < \infty \Leftrightarrow \int_{\{\Re(d) < \tilde{m} + 1/2\}} \frac{\sigma_n^2(v)}{1 + 2n - 2\Re(d(v))} \xi(dv) < \infty.$$

For an integer  $n > \tilde{m}$ , we have  $1 + 2n - 2\Re(d(v)) \geq 2$  on  $\{\Re(d) < \tilde{m} + 1/2\}$ , hence the right-hand side of (5.13) follows as a consequence of  $\int \sigma_n^2 d\xi < \infty$ . For  $n = \tilde{m}$ , the left-hand side of (5.13) follows as a consequence of (3.9) not being satisfied for  $n = \tilde{m}$  by definition of  $\tilde{m}$ .  $\square$

## 5.2 Proofs of Section 3.3

**Proof of Lemma 3.5.** Since  $\epsilon$  is a white noise, as explained in Remark 3.2 (2), Assertion (ii) of Theorem 3.4 only needs to be checked for  $n = 0$ . The result follows since this case precisely corresponds to the conditions in (3.11) with  $D = \text{Id}_{\mathcal{H}_0} - N$ .  $\square$

The proof of Proposition 3.6 relies on the following lemma where we recall that the open and closed complex unit discs of  $\mathbb{C}$  are denoted by  $\mathbb{D}$  and  $\overline{\mathbb{D}}$ , respectively. This lemma will also be useful in the proofs of Section 4.2.

**Lemma 5.2.** *Let  $\mathcal{H}_0$  be a separable Hilbert space and  $N$  be in  $\mathcal{N}(\mathcal{H}_0)$ . Let  $\varrho$  and  $\varsigma$  such that*

$$\varrho \leq \inf \left\{ \langle \Re(N) x, x \rangle_{\mathcal{H}_0} : x \in \mathcal{H}_0, \|x\|_{\mathcal{H}_0} = 1 \right\} \leq \|N\|_{\infty} \leq \varsigma, \quad (5.14)$$

where  $\Re(N) = (N^H + N)/2$ . Then, there exist  $Q \in \mathcal{L}_b(\mathcal{H}_0)$  and  $(P_k^*)_{k \in \mathbb{N}} \in \mathcal{L}_b(\mathcal{H}_0)^{\mathbb{N}}$  such that, for all  $z \in \mathbb{D}$ ,

$$(1 - z)^{N - \text{Id}} = Q \left( \sum_{k=0}^{\infty} (k+1)^{-N} z^k \right) + \sum_{k=0}^{\infty} P_k^* z^k, \quad (5.15)$$

where the two infinite sums on the right-hand side are  $\mathcal{L}_b(\mathcal{H}_0)$ -valued power series with a convergence radius at least equal to 1.

There further exist  $C_{\varsigma}, C_{\varrho} > 0$  only depending on  $\varsigma$  and  $\varrho$  respectively such that

$$\|Q\|_{\infty} \leq C_{\varsigma} \quad \text{and, for all } k \geq 0, \quad \|P_k^*\|_{\infty} \leq C_{\varsigma} C_{\varrho} (k+1)^{-1-\varrho}. \quad (5.16)$$

Moreover, if  $\varrho > 0$ , then (5.15) continues to hold for all  $z \in \overline{\mathbb{D}} \setminus \{1\}$  with the two infinite sums converging in  $\mathcal{L}_b(\mathcal{H}_0)$ .

*Proof.* In the following, we denote by  $\mathfrak{n}$  the singular value function of  $N$ , defined on  $\mathcal{G}_0 := L^2(\mathcal{V}, \mathcal{V}, \xi)$  with decomposition operator  $U$ . Then, Condition (5.14) can be rewritten as

$$\varrho \leq \xi\text{-essinf}_{v \in \mathcal{V}} \Re(\mathfrak{n}(v)) \leq \xi\text{-esssup}_{v \in \mathcal{V}} |\mathfrak{n}(v)| \leq \varsigma. \quad (5.17)$$

We now proceed in three steps. We first show Relation (5.15) for all  $z \in \mathbb{D}$ , then that the bounds (5.16) hold, and, finally, we extend (5.15) to  $z \in \overline{\mathbb{D}} \setminus \{1\}$  when  $\varrho > 0$ .

**Step 1.** Let  $z \in \mathbb{D}$ , then

$$(1 - z)^{N - \text{Id}} = \text{Id} + \sum_{k \geq 1} N_k z^k \quad \text{with} \quad N_k = \prod_{j=1}^k \left( \text{Id} - \frac{N}{j} \right), \quad \text{for all } k \geq 1. \quad (5.18)$$

Define the integer  $k_0 \geq 1$  by the condition  $\varsigma < k_0 \leq \varsigma + 1$ . Then, for all  $j \geq k_0$ ,  $\text{Id} - \frac{N}{j} = \exp\left(\ln\left(\text{Id} - \frac{N}{j}\right)\right) = \exp\left(-\sum_{\ell \geq 1} \frac{N^{\ell}}{\ell j^{\ell}}\right)$  and therefore, for all  $k \geq k_0$ ,

$$\begin{aligned} N_k &= \prod_{j=1}^{k_0-1} \left( \text{Id} - \frac{N}{j} \right) \exp\left(-\sum_{\ell \geq 1} \frac{N^{\ell}}{\ell} \sum_{j=k_0}^k \frac{1}{j^{\ell}}\right) \\ &= \prod_{j=1}^{k_0-1} \left( \text{Id} - \frac{N}{j} \right) \exp\left(-N \sum_{j=k_0}^k \frac{1}{j}\right) \exp\left(-\sum_{\ell \geq 2} \frac{N^{\ell}}{\ell} \sum_{j=k_0}^k \frac{1}{j^{\ell}}\right). \end{aligned} \quad (5.19)$$

Moreover, we have the following asymptotic expansions. For all  $k \geq k_0$ ,

$$\begin{aligned} \sum_{j=k_0}^k \frac{1}{j} &= \sum_{j=1}^k \frac{1}{j} - \sum_{j=1}^{k_0-1} \frac{1}{j} = \ln(k+1) + \gamma_e - \sum_{j=1}^{k_0-1} \frac{1}{j} + \frac{\alpha_k}{k}, \\ \sum_{j=k_0}^k \frac{1}{j^\ell} &= \sum_{j=k_0}^{\infty} \frac{1}{j^\ell} - \sum_{j=k+1}^{\infty} \frac{1}{j^\ell} = \frac{\beta_\ell}{k_0^\ell} + \frac{\eta_{k,\ell}}{(\ell-1)k^{\ell-1}}, \text{ for all } \ell \geq 2, \end{aligned}$$

where  $\gamma_e$  is Euler's constant,  $\beta_\ell := \sum_{k=k_0}^{\infty} \left(\frac{k_0}{k}\right)^\ell$ , and  $(\alpha_k)_{k \geq 1}$  and  $(\eta_{k,\ell})_{k \geq 1, \ell \geq 2}$  are some universal constants satisfying

$$\sup_{k \geq 1} |\alpha_k| < \infty \quad \text{and} \quad \sup_{k \geq 1, \ell \geq 2} |\eta_{k,\ell}| < \infty. \quad (5.20)$$

Also note that

$$\sup_{\ell \geq 2} \beta_\ell = \beta_2 < \infty. \quad (5.21)$$

Using these definitions in (5.19), we obtain, for all  $k \geq k_0$ ,

$$N_k = Q(k+1)^{-N} \exp \left( -N \frac{\alpha_k}{k} - \sum_{\ell \geq 2} \frac{N^\ell \eta_{k,\ell}}{(\ell-1)k^{\ell-1}} \right)$$

where

$$Q = \prod_{j=1}^{k_0-1} \left( \text{Id} - \frac{N}{j} \right) \exp \left( -N \left( \gamma_e - \sum_{t=1}^{k_0-1} \frac{1}{t} \right) \right) \exp \left( - \sum_{\ell \geq 2} \left( \frac{N}{k_0} \right)^\ell \frac{\beta_\ell}{\ell} \right). \quad (5.22)$$

Using the previous equations in (5.18), for all  $z \in \mathbb{D}$ , we can write  $(1-z)^{N-\text{Id}}$  as

$$\text{Id} + \sum_{k=1}^{k_0-1} \prod_{j=1}^k \left( \text{Id} - \frac{N}{j} \right) z^k + Q \sum_{k \geq k_0} (k+1)^{-N} \exp \left( -N \frac{\alpha_k}{k} - \sum_{\ell \geq 2} \frac{N^\ell \eta_{k,\ell}}{(\ell-1)k^{\ell-1}} \right) z^k.$$

Thus Relation (5.15) follows by setting

$$P_0^* := \text{Id} - Q, \quad (5.23)$$

$$P_k^* := \prod_{j=1}^k \left( \text{Id} - \frac{N}{j} \right) - Q (k+1)^{-N}, \quad \text{for } 1 \leq k \leq k_0 - 1, \quad (5.24)$$

$$P_k^* := Q(k+1)^{-N} \left[ \exp \left( -N \frac{\alpha_k}{k} - \sum_{\ell \geq 2} \frac{N^\ell \eta_{k,\ell}}{(\ell-1)k^{\ell-1}} \right) - \text{Id} \right] \quad \text{for } k \geq k_0. \quad (5.25)$$

**Step 2.** By (5.22), (5.14) and (5.21), first note that  $\|Q\|_\infty$  can be bounded by a constant only depending on  $\varsigma$  (since  $k_0$  only depends on  $\varsigma$  as well). Hence, by (5.23) and (5.24), for  $k < k_0 \leq \varsigma + 1$  we can again bound  $\|P_k^*\|$  by a constant only depending on  $\varsigma$ . Now let  $k \geq k_0$ , defining

$$\Phi_k := -N \frac{\alpha_k}{k} - \sum_{\ell \geq 2} \frac{N^\ell \eta_{k,\ell}}{(\ell-1)k^{\ell-1}},$$

Relation (5.25) yields

$$\|P_k^*\|_\infty \leq \|Q\|_\infty \left\| (k+1)^{-N} \right\|_\infty \sum_{t \geq 1} \frac{\|\Phi_k\|_\infty^t}{t!}.$$

Using the singular value function  $n$  with (5.17), we have

$$\left\| (k+1)^{-N} \right\|_\infty = \left\| (k+1)^{-M_n} \right\|_\infty = \|M_{(k+1)-n}\|_\infty = \xi\text{-essup}_{v \in V} \left| (k+1)^{-n(v)} \right| \leq (k+1)^{-e}.$$

Using the upper bound of operator norm of  $N$  in (5.14), we have

$$\begin{aligned}\|\Phi_k\|_\infty &\leq \varsigma \frac{|\alpha_k|}{k} + \sum_{\ell \geq 2} \frac{\varsigma^\ell \eta_{k,\ell}}{(\ell-1)k^{\ell-1}} \\ &= \frac{\varsigma}{k} \left( |\alpha_k| + \sum_{\ell \geq 1} \frac{\varsigma^\ell}{\ell} \eta_{k,\ell+1} k^{1-\ell} \right) \\ &\leq \frac{\varsigma}{k} \left( |\alpha_k| + \left( \sup_{k \geq 1, \ell \geq 2} |\eta_{k,\ell}| \right) \varsigma \left( 1 - \frac{\varsigma}{k_0} \right)^{-1} \right).\end{aligned}$$

By (5.20), this bound only depend on  $\varsigma$  (since  $k_0$  does as well). Gathering the obtained bounds we get (5.16).

**Step 3.** We now assume  $\varrho > 0$  and extend (5.15) to  $\overline{\mathbb{D}} \setminus \{1\}$ , that is to the case  $z = e^{-i\lambda}$  for some  $\lambda \in \mathbb{T} \setminus \{0\}$ . For such a  $\lambda$ , we already have, for all  $0 < a < 1$ ,

$$(1 - ae^{-i\lambda})^{N-\text{Id}} = Q \sum_{k \geq 0} (k+1)^{-N} a^k e^{-i\lambda k} + \sum_{k \geq 0} P_k^* a^k e^{-i\lambda k}.$$

Moreover,  $(1 - e^{-i\lambda})^{N-\text{Id}} = \lim_{a \uparrow 1} (1 - ae^{-i\lambda})^{N-\text{Id}}$  by continuity of  $z \mapsto (1-z)^{N-\text{Id}}$  in  $\overline{\mathbb{D}} \setminus \{1\}$  and  $\sum_{k \geq 0} P_k^* e^{-i\lambda k} = \lim_{a \uparrow 1} \sum_{k \geq 0} P_k^* a^k e^{-i\lambda k}$  because  $\sum_{k \geq 0} \|P_k^*\|_\infty < \infty$ . It remains to show that  $\sum_{k \geq 0} (k+1)^{-N} z$  is well defined on  $\mathbb{U} \setminus \{1\}$  and that, for  $\lambda \in \mathbb{T} \setminus \{0\}$ ,  $\sum_{k \geq 0} (k+1)^{-N} a^k e^{-i\lambda k}$  converges to  $\sum_{k \geq 0} (k+1)^{-N} e^{-i\lambda k}$  as  $a \uparrow 1$ . We prove these facts at once by applying Lemma A.3 with  $a_k = (k+1)^{-N}$ . We already used in Step 2 that, for all  $k \in \mathbb{N}$ , we have  $\|(k+1)^{-N}\|_\infty \leq (k+1)^{-\varrho}$ . Since  $\varrho > 0$ , we get that  $\|(k+1)^{-N}\|_\infty \rightarrow 0$  as  $k \rightarrow \infty$ . Hence, to apply Lemma A.3 it only remains to show

$$\sum_{k \in \mathbb{N}} \left\| (k+1)^{-N} - (k+2)^{-N} \right\|_\infty < \infty. \quad (5.26)$$

Note that we have, for all  $k \in \mathbb{N}$ ,

$$\left\| (k+1)^{-N} - (k+2)^{-N} \right\|_\infty = \xi\text{-esssup}_{v \in \mathbb{V}} \left| (k+1)^{-n(v)} - (k+2)^{-n(v)} \right|. \quad (5.27)$$

Moreover, for all  $k \in \mathbb{N}$ , and  $\xi$ -a.e.  $v \in \mathbb{V}$ , since  $\Re(n(v)) \geq \varrho > 0$ , we have

$$\begin{aligned}\left| (k+1)^{-n(v)} - (k+2)^{-n(v)} \right| &= |k+1|^{-\Re(n(v))} \left| 1 - \exp\left(-\ln\left(1 + \frac{1}{k+1}\right) n(v)\right) \right| \\ &\leq \varsigma \alpha(\varsigma \ln(2)) (k+1)^{-\varrho} \ln\left(1 + \frac{1}{k+1}\right),\end{aligned}$$

where, here, we set, for any  $r > 0$ ,  $\alpha(r) := \sup \left\{ \left| \frac{1-e^{-z}}{z} \right| : z \in \mathbb{C} \ 0 < |z| \leq r \right\}$ . This leads to the asymptotic bound, as  $k \rightarrow \infty$ ,  $\xi\text{-esssup}_{v \in \mathbb{V}} \left| (k+1)^{-n(v)} - (k+2)^{-n(v)} \right| = O((k+1)^{-\varrho-1})$ . Hence, with (5.27) and the assumption  $\varrho > 0$ , we obtain (5.26).  $\square$

**Proof of Proposition 3.6.** The first condition in (3.11) gives that (5.14) holds with  $\varrho \geq 1/2$ . Applying Lemma 5.2, there exists  $Q \in \mathcal{L}_b(\mathcal{H}_0)$  and  $(P_k^*)_{k \in \mathbb{N}} \in \mathcal{L}_b(\mathcal{H}_0)^{\mathbb{N}}$  with  $\|P_k^*\|_\infty = O(k^{-3/2})$  such that  $(1 - e^{-i\lambda})^{N-\text{Id}} = Q \sum_{k=0}^\infty (k+1)^{-N} e^{-i\lambda k} + \sum_{k=0}^\infty P_k^* e^{-i\lambda k}$  in  $\mathcal{L}_b(\mathcal{H}_0)$  for all  $\lambda \in \mathbb{T} \setminus \{0\}$ , thus concluding the proof.  $\square$

## 5.3 Proofs of Section 4.1

### 5.3.1 Preliminary results

In the following, for two separable Hilbert spaces  $\mathcal{H}_0$  and  $\mathcal{I}_0$  and a finite non-negative measure  $\mu$  on  $(\mathbb{T}, \mathcal{B}(\mathbb{T}))$ , we denote by  $\|\cdot\|_{1,1}$  the natural norm of the Bochner space  $L^{1,1}(\mathcal{H}_0, \mathcal{I}_0, \mu) := L^1(\mathbb{T}, \mathcal{B}(\mathbb{T}), \mathcal{S}_1(\mathcal{H}_0, \mathcal{I}_0), \mu)$ , that is,

$$\|g\|_{1,1} := \int \|g(\lambda)\|_1 d\mu.$$

We use the notation

$$B_{1,1}(r, \mathcal{H}_0, \mathcal{I}_0, \mu) = \left\{ g \in L^1(\mathbb{T}, \mathcal{B}(\mathbb{T}), \mathcal{S}_1(\mathcal{H}_0, \mathcal{I}_0), \mu) : \|g\|_{1,1} \leq r \right\},$$

for the ball of radius  $r$  in the Banach space  $L^{1,1}(\mathcal{H}_0, \mathcal{I}_0, \mu)$ . As usual, if  $\mathcal{I}_0 = \mathcal{H}_0$ , we drop  $\mathcal{I}_0$  in the notation, thus writing  $L^{1,1}(\mathcal{H}_0, \mu)$  and  $B_{1,1}(r, \mathcal{H}_0, \mu)$  in this case. Also if  $\mu = \text{Leb}_{\mathbb{T}}$  we drop the measure in the notation, thus writing  $L^{1,1}(\mathcal{H}_0, \mathcal{I}_0)$  and  $B_{1,1}(r, \mathcal{H}_0, \mathcal{I}_0)$ , or  $L^{1,1}(\mathcal{H}_0)$  and  $B_{1,1}(r, \mathcal{H}_0)$  if  $\mathcal{I}_0 = \mathcal{G}_0$ , in this case.

These definitions and those introduced in Section 4.1 (such as  $\mathcal{Q}$  and  $\tilde{\mathcal{Q}}$ ) will be useful in the following. Recall, in particular, that  $\mathbb{F}_b(\mathbb{T}, \mathcal{B}(\mathbb{T}), \mathcal{E})$  denotes the set of bounded measurable functions from  $(\mathbb{T}, \mathcal{B}(\mathbb{T}))$  to  $(\mathcal{E}, \mathcal{B}(\mathcal{E}))$ . We will also need to define  $\mathbb{F}_{b,b}((\mathcal{H}_0, \mathcal{G}_0), (\mathcal{I}_0, \mathcal{J}_0))$  as the product vector space  $\mathbb{F}_b(\mathbb{T}, \mathcal{B}(\mathbb{T}), \mathcal{L}_b(\mathcal{I}_0, \mathcal{J}_0)) \times \mathbb{F}_b(\mathbb{T}, \mathcal{B}(\mathbb{T}), \mathcal{L}_b(\mathcal{H}_0, \mathcal{G}_0))$ , endowed with the max norm

$$\|(L, R)\|_{b,b} := \max(\sup(L), \sup(R)) \quad \text{for all } (L, R) \in \mathbb{F}_{b,b}((\mathcal{H}_0, \mathcal{G}_0), (\mathcal{I}_0, \mathcal{J}_0)).$$

This simply extends the definition of  $\mathbb{F}_{b,b}(\mathcal{H}_0, \mathcal{G}_0)$  already introduced in Section 4.1, which can be seen as a short-hand notation for  $\mathbb{F}_{b,b}((\mathcal{H}_0, \mathcal{G}_0), (\mathcal{H}_0, \mathcal{G}_0))$ .

We now derive a series of useful lemmas.

**Lemma 5.3.** *Let  $\mu$  be a finite non-negative measure on  $(\mathbb{T}, \mathcal{B}(\mathbb{T}))$  and  $\mathcal{H}_0, \mathcal{G}_0, \mathcal{I}_0$  and  $\mathcal{J}_0$  be four separable Hilbert spaces and  $g \in L^{1,1}(\mathcal{H}_0, \mathcal{I}_0, \mu)$ . Then the mapping  $\mathcal{Q}_{g,\mu}$  defined by*

$$\mathcal{Q}_{g,\mu}(L, R) = \int L g R^{\text{H}} d\mu \tag{5.28}$$

is a continuous sesquilinear mapping from  $\mathbb{F}_{b,b}((\mathcal{H}_0, \mathcal{G}_0), (\mathcal{I}_0, \mathcal{J}_0))$  to  $\mathcal{S}_1(\mathcal{G}_0, \mathcal{J}_0)$  satisfying

$$\sup \left\{ \|\mathcal{Q}_{g,\mu}(L, R)\|_1 : \|(L, R)\|_{b,b} \leq 1 \right\} \leq \|g\|_{1,1}.$$

Consequently, for any positive radius  $r$ , the set  $\{\mathcal{Q}_{g,\mu} : g \in B_{1,1}(r, \mathcal{H}_0, \mathcal{I}_0, \mu)\}$  is equicontinuous in  $\mathcal{C}(\mathbb{F}_{b,b}((\mathcal{H}_0, \mathcal{G}_0), (\mathcal{I}_0, \mathcal{J}_0)), \mathcal{S}_1(\mathcal{G}_0, \mathcal{I}_0))$ .

*Proof.* For all  $(L, R) \in \mathbb{F}_{b,b}((\mathcal{H}_0, \mathcal{G}_0), (\mathcal{I}_0, \mathcal{J}_0))$ ,  $g \in L^{1,1}(\mathcal{H}_0, \mathcal{I}_0)$  and  $\lambda \in \mathbb{T}$ , we have

$$\left\| L(\lambda) g(\lambda) R(\lambda)^{\text{H}} \right\|_1 \leq \|L(\lambda)\|_{\infty} \|g(\lambda)\|_1 \left\| R(\lambda)^{\text{H}} \right\|_{\infty},$$

which is thus integrable with respect to  $\mu$ . Moreover, we obtain that

$$\|\mathcal{Q}_{g,\mu}(L, R)\|_1 \leq \|g\|_{1,1} \sup(L) \sup(R).$$

The given claim immediately follows as well as its consequence.  $\square$

We also derive the following lemma, which will be useful in the following.

**Lemma 5.4.** *Let  $\mathcal{H}_0, \mathcal{G}_0$  and  $\mathcal{I}_0$  be three separable Hilbert spaces,  $\Theta$  be a compact metric space, and let  $L$  and  $R$  be two continuous mappings from  $\Theta \times \mathbb{T}$  into  $\mathcal{L}_b(\mathcal{H}_0, \mathcal{G}_0)$  and  $\mathcal{L}_b(\mathcal{H}_0, \mathcal{I}_0)$ , respectively. Then, for any positive radius  $r$ , the set  $\mathcal{R}(r) := \left\{ \tilde{\mathcal{Q}}_g^{(L,R)} : g \in B_{1,1}(r, \mathcal{H}_0) \right\}$  is equicontinuous in  $\mathcal{C}(\Theta, \mathcal{S}_1(\mathcal{I}_0, \mathcal{G}_0))$ .*

*Proof.* Let  $r > 0$ . Since  $\Theta \times \mathbb{T}$  is compact,  $R$  and  $L$  are uniformly continuous on  $\Theta \times \mathbb{T}$  and we get that  $\theta \mapsto (L(\theta, \cdot), R(\theta, \cdot)) \in \mathcal{C}(\Theta, \mathbb{F}_{b,b}((\mathcal{H}_0, \mathcal{G}_0), (\mathcal{H}_0, \mathcal{I}_0)))$ . By Lemma 5.3, we get that  $\mathcal{R}(r)$  is equicontinuous in  $\mathcal{C}(\Theta, \mathcal{S}_1(\mathcal{I}_0, \mathcal{G}_0))$ .  $\square$

We next provide three last preliminary lemmas, one about the non-centered periodogram, and two dealing with the centering term.

**Lemma 5.5.** *Let  $\mathcal{H}_0$  and  $\mathcal{I}_0$  be two separable Hilbert spaces such that  $\mathcal{I}_0$  is continuously embedded in  $\mathcal{H}_0$ . Assume (A-1) and suppose moreover that  $X_0 \in \mathcal{I}_0$   $\mathbb{P}$ -a.s. with  $\mathbb{E} \left[ \|X_0\|_{\mathcal{I}_0}^2 \right] < \infty$ . Then, for all  $n \geq 1$ ,  $I_n^X \in L^{1,1}(\mathcal{H}_0, \mathcal{I}_0)$   $\mathbb{P}$ -a.s. and we have*

$$\sup_{n \geq 1} \left\| I_n^X \right\|_{1,1} < \infty \quad \mathbb{P}\text{-a.s.}$$

*Proof.* Note that, for all  $n \geq 1$ , we have, using the definition of  $I_n^X$ , and then that of  $d_n^X$ ,

$$\begin{aligned} \|I_n^X\|_{1,1} &= \int \|I_n^X\|_1 \, d\text{Leb}_{\mathbb{T}} \\ &\leq \int \|d_n^X(\lambda)\|_{\mathcal{I}_0} \|d_n^X(\lambda)\|_{\mathcal{H}_0} \, d\text{Leb}_{\mathbb{T}} \\ &\leq \left( \int \|d_n^X(\lambda)\|_{\mathcal{I}_0}^2 \, d\text{Leb}_{\mathbb{T}} \right)^{1/2} \left( \int \|d_n^X(\lambda)\|_{\mathcal{H}_0}^2 \, d\text{Leb}_{\mathbb{T}} \right)^{1/2} \\ &= \left( \frac{1}{n} \sum_{k=1}^n \|X_k\|_{\mathcal{I}_0}^2 \right)^{1/2} \left( \frac{1}{n} \sum_{k=1}^n \|X_k\|_{\mathcal{H}_0}^2 \right)^{1/2}, \end{aligned}$$

where in the right-hand side of the first line  $\|\cdot\|_1$  denotes the  $\mathcal{S}_1(\mathcal{H}_0, \mathcal{I}_0)$ -norm. By **(A-1)**, with the Birkhoff ergodic theorem, we get that the right-hand of the previous bound converges  $\mathbb{P}$ -a.s. The uniform bound follows  $\mathbb{P}$ -a.s.  $\square$

**Lemma 5.6.** *Recall that  $I_n^X$  and  $I_n^{X_n^c}$  denote the periodograms respectively computed from  $X_1, \dots, X_n$  and from  $X_{n,1}^c, \dots, X_{n,n}^c$ , as defined in (1.3). Suppose that  $X_1, \dots, X_n \in \mathcal{H}_0$ . Then,  $I_n^X$  and  $I_n^{X_n^c}$  belong to  $\mathcal{S}_1(\mathcal{H}_0)$  and we have, for all  $\lambda \in \mathbb{T}$ ,*

$$\|I_n^X(\lambda) - I_n^{X_n^c}(\lambda)\|_1 \leq \left\| \frac{1}{n} \sum_{j=1}^n X_j \right\|_{\mathcal{H}_0}^2 F_n(\lambda) + 2 \left\| \frac{1}{n} \sum_{j=1}^n X_j \right\|_{\mathcal{H}_0} \|d_n^X(\lambda)\|_{\mathcal{H}_0} (F_n(\lambda))^{1/2},$$

where  $F_n$  denotes the Fejér kernel defined by

$$F_n(\lambda) = \frac{1}{n} \left| \sum_{k=1}^n e^{-i\lambda k} \right|^2. \quad (5.29)$$

*Proof.* By (1.3) and (4.1), we have, for all  $\lambda \in \mathbb{T}$ ,

$$d_n^{X_n^c}(\lambda) = d_n^X(\lambda) - \left( \frac{1}{n} \sum_{j=1}^n X_j \right) \left( \frac{1}{\sqrt{n}} \sum_{k=1}^n e^{-i\lambda k} \right).$$

Computing  $I_n^{X_n^c}(\lambda) = d_n^{X_n^c}(\lambda) \otimes d_n^{X_n^c}(\lambda)$  and using that  $\|x \otimes y\|_1 = \|x\|_{\mathcal{H}_0} \|y\|_{\mathcal{H}_0}$  for all  $x, y \in \mathcal{H}_0$ , we easily get the result.  $\square$

We have the following result on the empirical mean.

**Lemma 5.7.** *Let  $\mathcal{H}_0$  be a separable Hilbert space. Assume **(A-1)** and **(A-2)**. Then, the process  $X$  is valued in a finite-dimensional space  $\mathcal{G}_0 \subset \mathcal{H}_0$  or, if it is not the case, there always exists an orthonormal sequence  $(\phi_k)_{k \in \mathbb{N}}$  of  $\mathcal{H}_0$  and a sequence  $s = (s_k)_{k \in \mathbb{N}} \in [1, \infty)^{\mathbb{N}}$  such that Conditions (i) and (ii) in Theorem 4.4 hold.*

*Proof.* Let  $\Gamma(0)$  denote the covariance operator of  $X_0$ . Since  $\Gamma(0) \in \mathcal{S}_1^+(\mathcal{H}_0)$ , there exists a finite or countable non-increasing sequence  $(\sigma_k)_{0 \leq k < K} \in (0, \infty)^{\mathbb{N}}$  and an orthonormal sequence  $(\phi_k)_{0 \leq k < K} \in \mathcal{H}_0^{\mathbb{N}}$  such that

$$\Gamma(0) = \sum_{0 \leq k < K} \sigma_k^2 \phi_k \otimes \phi_k \quad \text{and} \quad \sum_{0 \leq k < K} \sigma_k^2 < \infty.$$

In particular, we have that  $\mathbb{P}$ -a.s.,  $X_0$  is valued in  $\overline{\text{Span}}^{\mathcal{H}_0}(\phi_k, 0 \leq k < K)$ . If  $K$  is finite, then  $X_0$  is valued in the finite-dimensional space  $\mathcal{G}_0 = \text{Span}(\phi_k, 0 \leq k < K)$ .

From now on, we take  $K = \infty$ . By Lemma A.4, we can find  $s = (s_k)_{k \in \mathbb{N}} \in [1, \infty)^{\mathbb{N}}$ , non-decreasing and going to  $\infty$  (hence, satisfying Condition (i) in Theorem 4.4), such that  $\sum_{k \in \mathbb{N}} s_k^2 \sigma_k^2 < \infty$ . Defining  $\mathcal{H}_0^s$  by (4.8) and its inner product by (4.9), we get that

$$\mathbb{E} \left[ \|X_0\|_{\mathcal{H}_0^s}^2 \right] = \sum_{k \in \mathbb{N}} s_k^2 \sigma_k^2 < \infty.$$

We thus have Condition (ii) in Theorem 4.4.  $\square$



The following lemma is used to treat the centering term in the next result, and also to prove Theorem 1.1.

**Lemma 5.8.** *Let  $\mathcal{H}_0$  be a separable Hilbert space. Assume (A-1) and (A-2). Then,*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n X_j = \mathbb{E}[X_0] \quad \text{in } \mathcal{H}_0, \quad \mathbb{P}\text{-a.s.} \quad (5.30)$$

*Proof.* If we can find a finite-dimensional space  $\mathcal{G}_0 \subset \mathcal{H}_0$  such that  $X$  is valued in  $\mathcal{G}_0$ , then (5.30) follows immediately. If not, by Lemma 5.7, we can find an orthonormal sequence  $(\phi_k)_{k \in \mathbb{N}}$  of  $\mathcal{H}_0$  such that Conditions (i) and (ii) in Theorem 4.4 hold. As a consequence, by the Birkhoff ergodic theorem, we get that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \|X_j\|_{\mathcal{H}_0^s} = \mathbb{E} \left[ \|X_0\|_{\mathcal{H}_0^s} \right] \quad \mathbb{P}\text{-a.s.}$$

Since  $\left\| \frac{1}{n} \sum_{j=1}^n X_j \right\|_{\mathcal{H}_0^s} \leq \frac{1}{n} \sum_{j=1}^n \|X_j\|_{\mathcal{H}_0^s}$  for all  $n \geq 1$ , we get that,  $\mathbb{P}$ -a.s., there exists  $r > 0$  such that  $\sup_{n \geq 1} \left\| \frac{1}{n} \sum_{j=1}^n X_j \right\|_{\mathcal{H}_0^s} \leq r$ . Using that  $(s_k)_{k \in \mathbb{N}}$  is going to infinity, we get that the operator  $\sum_{k \in \mathbb{N}} s_k^{-2} \phi_k \otimes \phi_k$  belongs to  $\mathcal{S}_\infty(\mathcal{H}_0)$ . Since the image by this operator of the unit  $\mathcal{H}_0$ -ball is the unit  $\mathcal{H}_0^s$ -ball, we get that all  $\mathcal{H}_0^s$ -balls are compact in  $\mathcal{H}_0$ . In particular, we have that  $\left( \frac{1}{n} \sum_{j=1}^n X_j \right)_{n \geq 1}$  is  $\mathbb{P}$ -a.s. valued in a compact subset of  $\mathcal{H}_0$ . Therefore, it only remains to show that,  $\mathbb{P}$ -a.s., the only possible accumulation point of this sequence is  $\mathbb{E}[X_0]$ . This fact follows by using the Birkhoff ergodic theorem again, which gives us, for all  $x \in \mathcal{H}_0$ ,

$$\lim_{n \rightarrow \infty} \left\langle \frac{1}{n} \sum_{j=1}^n X_j, x \right\rangle_{\mathcal{H}_0} = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \langle X_j, x \rangle_{\mathcal{H}_0} = \mathbb{E} \left[ \langle X_0, x \rangle_{\mathcal{H}_0} \right] = \langle \mathbb{E}[X_0], x \rangle_{\mathcal{H}_0} \quad \mathbb{P}\text{-a.s.}$$

Hence we get the convergence (5.30).  $\square$

We can now state a result which applies both in the context of Theorem 4.2 and Theorem 4.3.

**Theorem 5.9.** *Let  $\mathcal{H}_0$  and  $\mathcal{G}_0$  be two separable Hilbert spaces. Assume (A-1) and (A-2), and suppose that  $(\Theta, \Delta)$  is a compact metric space. Let  $L$  and  $R$  in  $\mathcal{C}(\Theta \times \mathbb{T}, \mathcal{L}_b(\mathcal{H}_0, \mathcal{G}_0))$ . Moreover, suppose that,  $\mathbb{P}$ -a.s., there exists, for all  $\theta \in \Theta$ , a compact subset  $B \subset \mathcal{S}_1(\mathcal{G}_0)$  such that, for all  $n \geq 1$   $\tilde{\mathcal{Q}}_{I_n^X}^{(L,R)}(\theta) \in B$ . Then, we have*

$$\lim_{n \rightarrow \infty} \tilde{\mathcal{Q}}_{I_n^{X_n^c}}^{(L,R)} = \tilde{\mathcal{Q}}_{\nu_X}^{(L,R)} \quad \text{uniformly in } \mathcal{C}(\Theta, \mathcal{S}_1(\mathcal{G}_0)), \quad \mathbb{P}\text{-a.s.} \quad (5.31)$$

*Proof.* Thanks to the centering (1.3), we can replace all  $X_k$ 's by  $X_k - \mathbb{E}[X_0]$  without modifying  $X_{n,k}^c$  for any  $n, k \geq 1$ . Hence, without loss of generality, from now on in this proof, we assume that  $\mathbb{E}[X_0]$  is zero, that is, the process  $X$  is centered.

By Lemma 5.6 and using that  $\int F_n \, d\text{Leb}_{\mathbb{T}} = 1$ , first note that for all  $\theta \in \Theta$  we have, for all  $n \geq 1$ ,

$$\left\| \tilde{\mathcal{Q}}_{I_n^{X_n^c}}(\theta) - \tilde{\mathcal{Q}}_{I_n^X}(\theta) \right\|_1 \leq A_n (A_n + B_n) \sup_{\substack{\theta \in \Theta \\ \lambda \in \mathbb{T}}} \|L(\theta, \lambda)\|_\infty \sup_{\substack{\theta \in \Theta \\ \lambda \in \mathbb{T}}} \|R(\theta, \lambda)\|_\infty,$$

where we set

$$A_n := \left\| \frac{1}{n} \sum_{j=1}^n X_j \right\|_{\mathcal{H}_0} \quad \text{and} \quad B_n = 2 \int \left\| d_n^X \right\|_{\mathcal{H}_0} (F_n)^{1/2} \, d\text{Leb}_{\mathbb{T}}.$$

By the Cauchy-Schwartz inequality and then the Parseval identity, we have

$$B_n \leq 2 \left( \int \left\| d_n^X \right\|_{\mathcal{H}_0}^2 \, d\text{Leb}_{\mathbb{T}} \right)^{1/2} = 2 \left( \frac{1}{n} \sum_{k=1}^n \|X_k\|_{\mathcal{H}_0}^2 \right)^{1/2},$$

which, by the Birkhoff ergodic theorem and **(A-2)**, converges  $\mathbb{P}$ -a.s.. On the other hand, using Lemma 5.8, we have that,  $\mathbb{P}$ -a.s.,  $\lim_{n \rightarrow \infty} A_n = 0$ . Hence, we finally get that

$$\lim_{n \rightarrow \infty} \sup_{\theta \in \Theta} \left\| \tilde{\mathcal{Q}}_{I_n^{X_n^c}}(\theta) - \tilde{\mathcal{Q}}_{I_n^X}(\theta) \right\|_1 = 0 \quad \mathbb{P}\text{-a.s.}$$

Therefore, to prove (5.31), we can replace  $I_n^{X_n^c}$  by  $I_n^X$ , that is, it only remains to prove that

$$\lim_{n \rightarrow \infty} \tilde{\mathcal{Q}}_{I_n^X}^{(L,R)} = \tilde{\mathcal{Q}}_{\nu_X}^{(L,R)} \quad \text{uniformly in } \mathcal{C}(\Theta, \mathcal{S}_1(\mathcal{G}_0)) \text{ , } \mathbb{P}\text{-a.s.} \quad (5.32)$$

We immediately have from Proposition 4.1 that  $\tilde{\mathcal{Q}}_{I_n^X}^{(L,R)}$  and  $\tilde{\mathcal{Q}}_{\nu_X}^{(L,R)}$  belong to  $\mathcal{C}(\Theta, \mathcal{S}_1(\mathcal{G}_0))$ . The rest of the proof is now in three steps. First, we show that,  $\mathbb{P}$ -a.s., every sequence valued in  $\{\tilde{\mathcal{Q}}_{I_n^X}^{(L,R)} : n \geq 1\}$  admits a subsequence which converges uniformly in  $\mathcal{C}(\Theta, \mathcal{S}_1(\mathcal{G}_0))$ . Second, we show that, for all  $x, y \in \mathcal{G}_0$  and  $\theta \in \Theta$ , we have

$$\lim_{n \rightarrow \infty} x^H \tilde{\mathcal{Q}}_{I_n^X}^{(L,R)}(\theta) y = x^H \tilde{\mathcal{Q}}_{\nu_X}^{(L,R)}(\theta) y \text{ , } \mathbb{P}\text{-a.s.} \quad (5.33)$$

We then conclude in Step 3 from these two results.

**Step 1.** By **(A-2)**, we have  $\mathbb{E} \left[ \|X_0\|_{\mathcal{H}_0}^2 \right] < \infty$ , and we can apply Lemma 5.5 with  $\mathcal{I}_0 = \mathcal{H}_0$  and get that,  $\mathbb{P}$ -a.s., there exists  $C$  such that, for all  $n \in \mathbb{N}$ ,  $\|I_n^X\|_{1,1} \leq C$ , where here  $\|\cdot\|_{1,1}$  denotes the norm in  $L^{1,1}(\mathcal{H}_0)$ . We conclude with Lemma 5.4 that,  $\mathbb{P}$ -a.s.,  $\{\tilde{\mathcal{Q}}_{I_n^X}^{(L,R)} : n \geq 1\}$  is equicontinuous in  $\mathcal{C}(\Theta, \mathcal{S}_1(\mathcal{G}_0))$ . Using the last assumption of the theorem, we also have that,  $\mathbb{P}$ -a.s., for all  $\theta \in \Theta$ , there is a compact subset  $B$  of  $\mathcal{S}_1(\mathcal{G}_0)$  such that  $\{\tilde{\mathcal{Q}}_{I_n^X}^{(L,R)}(\theta) : n \geq 1\}$  is included in  $B$ . Thus,  $\mathbb{P}$ -a.s., the Ascoli-Arzelà theorem (see [27, Section 7.10]) applies and any sequence in  $\{\tilde{\mathcal{Q}}_{I_n^X}^{(L,R)} : n \geq 1\}$  admits a uniformly convergent subsequence in  $\mathcal{C}(\Theta, \mathcal{S}_1(\mathcal{G}_0))$ , which concludes the proof of Step 1.

**Step 2.** This step is similar to the scalar case and we follow the ideas of [14]. Let  $x, y \in \mathcal{G}_0$ ,  $\theta \in \Theta$ , and denote  $\mathbf{x} = x^H L(\theta, \cdot)$  and  $\mathbf{y} = y^H R(\theta, \cdot)$ , so that  $\mathbf{x}$  and  $\mathbf{y}$  are continuous functions from  $\mathbb{T}$  to  $\mathcal{H}_0^H = \mathcal{L}_b(\mathcal{H}_0, \mathbb{C})$ , and (5.33) can be written as

$$\lim_{n \rightarrow \infty} \mathcal{Q}_{I_n^X}(\mathbf{x}, \mathbf{y}) = \mathcal{Q}_{\nu_X}(\mathbf{x}, \mathbf{y}) \text{ , } \mathbb{P}\text{-a.s.} \quad (5.34)$$

We can write, for any  $N \in \mathbb{N}^*$ ,

$$\mathbf{x} = F_N \star \mathbf{x} + (\mathbf{x} - F_N \star \mathbf{x}) \text{ ,}$$

where  $\star$  denotes the convolution of locally integrable  $2\pi$ -periodic functions,

$$f \star g(\lambda) = \int f(\lambda') g(\lambda - \lambda') \text{Leb}_{\mathbb{T}}(d\lambda') \text{ , } \lambda \in \mathbb{T} \text{ ,}$$

and  $F_N$  is the Fejér kernel defined by (5.29). Using standard properties of Fejér's kernel and the fact that  $\lambda \mapsto \mathbf{x}(\lambda)$  is continuous on  $\mathbb{R}$ , we have, denoting  $\mathbf{e}_\ell(\lambda) = e^{i\lambda\ell}$  so that  $(\mathbf{e}_\ell)_{\ell \in \mathbb{Z}}$  is a Hilbert basis of  $\mathcal{L}_0 := L^2(\mathbb{T}, \mathcal{B}(\mathbb{T}), \text{Leb}_{\mathbb{T}})$ ,

$$F_N \star \mathbf{x} = \sum_{\ell=-N}^N \alpha_\ell(N) c_\ell(\mathbf{x}) \mathbf{e}_\ell \text{ ,} \quad (5.35)$$

$$\text{with } \alpha_\ell(N) = 1 - \frac{|\ell|}{N} \text{ and } c_\ell(\mathbf{x}) = \int \mathbf{x}(\lambda) e^{-i\ell\lambda} \text{Leb}_{\mathbb{T}}(d\lambda) \text{ ,}$$

$$\lim_{N \rightarrow \infty} \sup_{\lambda \in \mathbb{R}} \|\mathbf{x}(\lambda) - F_N \star \mathbf{x}(\lambda)\|_\infty = 0 \text{ .} \quad (5.36)$$

Eq (5.36) can be interpreted as saying that  $F_N \star \mathbf{x}$  converges to  $\mathbf{x}$  in  $\mathbb{F}_b(\mathbb{T}, \mathcal{B}(\mathbb{T}), \mathcal{L}_b(\mathcal{H}_0, \mathbb{C}))$ . The same holds with  $\mathbf{y}$  replacing  $\mathbf{x}$  and, applying Lemma 5.3 with  $\mathcal{I}_0 = \mathcal{H}_0$  and  $\mathcal{G}_0 = \mathcal{J}_0 = \mathbb{C}$  and  $\mu = \text{Leb}_{\mathbb{T}}$ , since by Step 1,  $\{I_n^X : n \geq 1\}$  remains in a ball of  $L^{1,1}$   $\mathbb{P}$ -a.s., we have,  $\mathbb{P}$ -a.s.,

$$\lim_{N \rightarrow \infty} \sup_{g \in \{I_n^X : n \geq 1\}} |\mathcal{Q}_g(F_N \star \mathbf{x}, F_N \star \mathbf{y}) - \mathcal{Q}_g(\mathbf{x}, \mathbf{y})| \text{ .} \quad (5.37)$$

Similarly by the continuity of  $\mathcal{Q}_{\nu_X} = \mathcal{Q}_{f_X, \mu}$  (with  $f_X$  the density of  $\nu_X$  with respect to  $\mu = \|\nu_X\|_1$ ) established in Lemma 5.3, we have

$$\lim_{N \rightarrow \infty} \mathcal{Q}_{\nu_X}(F_N \star \mathbf{x}, F_N \star \mathbf{y}) = \mathcal{Q}_{\nu_X}(\mathbf{x}, \mathbf{y}). \quad (5.38)$$

Next, using (5.35), we have

$$\begin{aligned} \mathcal{Q}_{I_n^X}(F_N \star \mathbf{x}, F_N \star \mathbf{y}) &= \sum_{\ell, \ell' = -N}^N \alpha_\ell(N) \alpha_{\ell'}(N) c_\ell(\mathbf{x}) \mathcal{Q}_{I_n^X}(\mathbf{e}_\ell, \mathbf{e}_{\ell'}) c_{\ell'}(\mathbf{y})^H \\ &= \sum_{\ell, \ell' = -N}^N \alpha_\ell(N) \alpha_{\ell'}(N) c_\ell(\mathbf{x}) \tilde{\Gamma}_n(\ell - \ell') c_{\ell'}(\mathbf{y})^H, \end{aligned} \quad (5.39)$$

where  $\tilde{\Gamma}_n$  denotes the empirical covariance defined as in (4.2), but with  $X_n^c$  replaced by  $X$ , that is,

$$\tilde{\Gamma}_n(s - t) = \int I_n^X(\lambda) e^{i(s-t)\lambda} \text{Leb}_{\mathbb{T}}(d\lambda) = \frac{1}{n} \sum_{\substack{1 \leq k, k' \leq n \\ k - k' = (s-t)}} X_k \otimes X_{k'}.$$

In particular, we have for any  $\ell, \ell' \in \{-N, \dots, N\}$ ,

$$c_\ell(\mathbf{x}) \tilde{\Gamma}_n(\ell - \ell') c_{\ell'}(\mathbf{y})^H = \frac{1}{n} \sum_{1 \vee (1 + \ell - \ell') \leq k \leq n \wedge (n + \ell - \ell')} c_\ell(\mathbf{x}) X_k X_{k - (\ell - \ell')}^H c_{\ell'}(\mathbf{y})^H$$

By **(A-1)**, with the Birkhoff ergodic theorem, we get, for any  $\ell, \ell' \in \{-N, \dots, N\}$ ,  $\mathbb{P}$ -a.s.,

$$\begin{aligned} \lim_{n \rightarrow \infty} c_\ell(\mathbf{x}) \tilde{\Gamma}_n(\ell - \ell') c_{\ell'}(\mathbf{y})^H &= \mathbb{E} \left[ c_\ell(\mathbf{x}) X_{\ell - \ell'} X_0^H c_{\ell'}(\mathbf{y})^H \right] \\ &= c_\ell(\mathbf{x}) \text{Cov}(X_{\ell - \ell'}, X_0) c_{\ell'}(\mathbf{y})^H \\ &= c_\ell(\mathbf{x}) \mathcal{Q}_{\nu_X}(\mathbf{e}_\ell, \mathbf{e}_{\ell'}) c_{\ell'}(\mathbf{y})^H. \end{aligned}$$

where we used that  $X$  is centered and that  $\nu_X$  is the spectral operator measure of  $X$ . From (5.39), using (5.35) again, we get that,  $\mathbb{P}$ -a.s., for any  $N \geq 1$ ,

$$\lim_{n \rightarrow \infty} \mathcal{Q}_{I_n^X}(F_N \star \mathbf{x}, F_N \star \mathbf{y}) = \mathcal{Q}_{\nu_X}(F_N \star \mathbf{x}, F_N \star \mathbf{y}).$$

This, with (5.37) and (5.38), concludes Step 2.

**Step 3.** From Step 1, there exists  $\Omega' \in \mathcal{F}$  with probability 1 such that on  $\Omega'$ , any sequence valued in  $\left\{ \tilde{\mathcal{Q}}_{I_n^X}^{(L,R)} : n \geq 1 \right\}$  admits a subsequence uniformly converging in  $\mathcal{C}(\Theta, \mathcal{S}_1(\mathcal{H}_0))$ . To obtain (5.32), we will exhibit  $\Omega'' \subset \Omega'$  with probability one such that, on  $\Omega''$ ,  $\tilde{\mathcal{Q}}_{\nu_X}^{(L,R)}$  is the only possible accumulation point of the sequence  $\left( \tilde{\mathcal{Q}}_{I_n^X}^{(L,R)} \right)_{n \geq 1}$ . Let  $E_0$  be a countable linearly dense subset of  $\mathcal{G}_0$  and let  $(\theta_j)_{j \in \mathbb{N}}$  be a dense sequence in  $\Theta$ , which exists since  $\Theta$  is compact. Then, from Step 2, we have,  $\mathbb{P}$ -a.s.,

$$\forall j \in \mathbb{N}, \forall x, y \in E_0, \lim_{n \rightarrow \mathbb{N}} x^H \tilde{\mathcal{Q}}_{I_n^X}^{(L,R)} y = x^H \tilde{\mathcal{Q}}_{\nu_X}^{(L,R)} y.$$

We can thus take  $\Omega'' \subset \Omega'$  with probability one, on which the previous display holds. Let  $\omega \in \Omega''$  and take an accumulation point  $\tilde{\mathcal{Q}}_\infty$  of  $\left( \tilde{\mathcal{Q}}_{I_n^X(\omega)}^{(L,R)} \right)_{n \geq 1}$  in  $\mathcal{C}(\Theta, \mathcal{S}_1(\mathcal{G}_0))$ . Then, for all  $j \in \mathbb{N}$ , using the previous display and the fact that  $\tilde{\mathcal{Q}}_\infty$  must also be an accumulation point for the weak operator topology, we get that, for all  $x, y \in E_0$ ,  $x^H \tilde{\mathcal{Q}}_\infty(\theta_j) y = x^H \tilde{\mathcal{Q}}_{\nu_X}^{(L,R)}(\theta_j) y$ , which implies  $\tilde{\mathcal{Q}}_\infty(\theta_j) = \tilde{\mathcal{Q}}_{\nu_X}^{(L,R)}(\theta_j)$ . Since  $\tilde{\mathcal{Q}}_\infty$  and  $\tilde{\mathcal{Q}}_{\nu_X}^{(L,R)}$  are continuous on  $\Theta$ , and  $(\theta_j)_{j \in \mathbb{N}}$  is dense in  $\Theta$ , we get that  $\tilde{\mathcal{Q}}_\infty$  and  $\tilde{\mathcal{Q}}_{\nu_X}^{(L,R)}$  coincide, which concludes the proof.  $\square$

### 5.3.2 Proof of Theorem 4.2

We can now prove Theorem 4.2 as a direct application of Theorem 5.9.

**Proof of Theorem 4.2.** Theorem 4.2 directly follows from Theorem 5.9, if we can prove that,  $\mathbb{P}$ -a.s., there exists  $B$ , a compact subset of  $\mathcal{S}_1(\mathcal{G}_0)$ , such that  $\tilde{Q}_{I_n^X}^{(L,R)}(\theta) \in B$  for all  $n \geq 1$  and  $\theta \in \Theta$ . Because  $\mathcal{G}_0$  is finite-dimensional, so is  $\mathcal{S}_1(\mathcal{G}_0)$ , and we only need to show that,  $\mathbb{P}$ -a.s., there exists  $C > 0$  such that  $\left\| \tilde{Q}_{I_n^X}^{(L,R)}(\theta) \right\|_1 \leq C$  for all  $n \geq 1$  and  $\theta \in \Theta$ . By Lemma 5.3, this follows from the fact that,  $\mathbb{P}$ -a.s., there exists  $r > 0$  such that  $\|I_n^X\|_1 \leq r$  for all  $n \geq 1$ , which has already been used in the proof of Theorem 5.9 and is a consequence of Lemma 5.5 with **(A-2)** in the case  $\mathcal{H}_0 = \mathcal{I}_0$ . This concludes the proof.  $\square$

### 5.3.3 Proof of Theorem 4.3

The proof of Theorem 4.3 essentially follows the same path as that of Theorem 4.2. However, in the infinite-dimensional case, we will need an additional result (Proposition 5.11) to prove the assumption involving the set  $B$  in Theorem 5.9. The result relies on the space  $\mathcal{H}_0^s$  introduced in Section 4.1. In this section, we will make extensive use of partial isometries as (see [7, Definition 3.8]). We recall that a partial isometry  $U$  on the Hilbert space  $\mathcal{H}_0$  onto another Hilbert space  $\mathcal{G}_0$  is a bounded operator which is an isometry on  $(\ker(U))^\perp$ . The subspaces  $(\ker(U))^\perp$  and  $\text{Im}(U)$  are respectively called the *initial space* and *final space* of  $U$ . We recall that, if  $U$  is a partial isometry, then  $U^H U$  and  $U U^H$  are the orthogonal projections onto the initial and the final space of  $U$  respectively.

Let us start with the following lemma, whose proof is straightforward, but which contains some important definitions

**Lemma 5.10.** *Let  $\mathcal{H}_0$  be a separable Hilbert space and let  $(\phi_k)_{k \in \mathbb{N}}$  be an orthonormal sequence in  $\mathcal{H}_0$ . Let  $s = (s_k)_{k \in \mathbb{N}} \in [1, \infty)^\mathbb{N}$  and define  $(\mathcal{H}_0^s, \langle \cdot, \cdot \rangle_{\mathcal{H}_0^s})$  by (4.8) and (4.9). Denote by  $J_s$  the continuous  $\mathcal{H}_0^s \hookrightarrow \mathcal{H}_0$ -inclusion map defined on  $\mathcal{H}_0^s$  onto  $\mathcal{H}_0$  by  $x \mapsto x$ . Further denote by  $U_s$  the partial isometry with initial space  $\overline{\text{Span}}^{\mathcal{H}_0}(\phi_k, k \in \mathbb{N})$  and final space  $\mathcal{H}_0^s$  such that, for all  $k \in \mathbb{N}$ ,  $U_s \phi_k = s_k^{-1} \phi_k$ . Finally, let us set  $\mathbb{J}_s = J_s U_s \in \mathcal{L}_b(\mathcal{H}_0)$ . Then, for all  $x \in \mathcal{H}_0$ , we have*

$$\mathbb{J}_s = \sum_{k \in \mathbb{N}} s_k^{-1} \phi_k \otimes \phi_k. \quad (5.40)$$

Moreover, suppose that  $s$  is non-decreasing and going to  $\infty$ . Then, we have  $\mathbb{J}_s \in \mathcal{S}_\infty(\mathcal{H}_0)$ .

We now have the following result.

**Proposition 5.11.** *Let  $\mathcal{H}_0$  be a separable Hilbert space and let  $(\phi_k)_{k \in \mathbb{N}}$  be a an orthonormal sequence in  $\mathcal{H}_0$ . Define  $\mathbb{J}_s$  as in Lemma 5.10 for  $s = (s_k)_{k \in \mathbb{N}} \in [1, \infty)^\mathbb{N}$  non-decreasing and going to  $\infty$ . Define  $\mathbb{B}_1 = \left\{ P \in \mathcal{S}_1(\mathcal{H}_0) : \|P\|_1 \leq 1 \text{ and } (\mathbb{J}_s P)^H = \mathbb{J}_s P \right\}$ . Then the set  $B_1 = \{\mathbb{J}_s P : P \in \mathbb{B}_1\}$  is compact in  $\mathcal{S}_1(\mathcal{H}_0)$ .*

*Proof.* Let  $(P_n)_{n \in \mathbb{N}}$  be a sequence valued in  $B_1$  and let us prove that it admits a subsequence which converges in  $B_1$ . By definition, we can write, for all  $n \in \mathbb{N}$ ,  $P_n = \mathbb{J}_s \tilde{P}_n$  with  $\tilde{P}_n \in \mathbb{B}_1$ . We use that  $\mathcal{S}_1(\mathcal{H}_0)$  is isometric to the dual of the space  $\mathcal{S}_\infty(\mathcal{H}_0)$  (see [7, Theorem 19.1]). Then, by the Banach-Alaoglu Theorem (see Theorem 3.1 in [6, Chapter V]), we get that the unit ball of  $\mathcal{S}_1(\mathcal{H}_0)$  is compact for the weak-star topology, that is the topology generated by the family of semi-norms  $\{P \mapsto |\text{Tr}(CP)| : C \in \mathcal{S}_\infty(\mathcal{H}_0)\}$ . This implies that  $(\tilde{P}_n)_{n \in \mathbb{N}}$  admits a subsequence  $(\tilde{P}_{a_n})_{n \in \mathbb{N}}$  converging to an element  $\tilde{Q}$  in the unit ball of  $\mathcal{S}_1(\mathcal{H}_0)$  in the sense of the weak-star topology, that is, for all  $C \in \mathcal{S}_\infty(\mathcal{H}_0)$ , we have

$$\lim_{n \rightarrow \infty} \text{Tr}(C \tilde{P}_{a_n}) = \text{Tr}(C \tilde{Q}).$$

Observe that for all  $x, y \in \mathcal{H}_0$ , the operator  $C = x \otimes y \mathbb{J}_s$  is a rank-one (hence compact) linear operator on  $\mathcal{H}_0$  onto  $\mathcal{H}_0$ . The last display thus gives that  $\tilde{P}_{a_n}$  converges to  $\tilde{Q}$  in weak operator topology (that is, for all  $x, y \in \mathcal{H}_0$ ,  $\langle \tilde{P}_{a_n} x, y \rangle_{\mathcal{H}_0}$  converges to  $\langle \tilde{Q} x, y \rangle_{\mathcal{H}_0}$ ). Since  $\mathbb{J}_s \tilde{P}_{a_n}$  is hermitian for all  $n$ , we get that  $\mathbb{J}_s \tilde{Q}$  is hermitian as well and we finally get that  $\tilde{Q}$  must be in  $\mathbb{B}_1$ . (In fact we have shown that  $\mathbb{B}_1$  is compact for the weak-star topology).

Let us set  $Q = \mathbb{J}_s \tilde{Q} \in B_1$  and for all  $n \in \mathbb{N}$ ,  $\Delta_n = P_{a_n} - Q = \mathbb{J}_s \tilde{\Delta}_n$  with  $\tilde{\Delta}_n = \tilde{P}_{a_n} - \tilde{Q}$ , and let us summarize our findings so far. We already know that  $P_{a_n}$  and  $Q$  are in  $B_1$  (hence are hermitian and so is  $\Delta_n$ ), that  $\|\tilde{\Delta}_n\|_1 \leq 2$  and that  $(\tilde{\Delta}_n)_{n \rightarrow \infty}$  converges to zero in  $\mathcal{S}_1(\mathcal{H}_0)$

for the weak-star topology, which also implies the convergence in weak operator topology. To conclude, we now proceed in two steps. First, we show that  $(\Delta_n)_{n \in \mathbb{N}}$  converges to 0 for the strong operator topology. Second, we use the first step to show that  $\lim_{n \rightarrow +\infty} \|\Delta_n\|_1 = 0$ .

**Step 1.** Let  $x \in \mathcal{H}_0$ , then, for all  $n \in \mathbb{N}$ , using (5.40), we have

$$\|\Delta_n x\|_{\mathcal{H}_0}^2 = \sum_{k \in \mathbb{N}} \left| \langle \Delta_n x, \phi_k \rangle_{\mathcal{H}_0} \right|^2 = \sum_{k \in \mathbb{N}} s_k^{-2} \left| \langle \tilde{\Delta}_n x, \phi_k \rangle_{\mathcal{H}_0} \right|^2 .$$

Since  $(\tilde{\Delta}_n)_{n \in \mathbb{N}}$  converges to 0 for the weak operator topology, we have, for all  $m \geq 1$ ,

$$\lim_{n \rightarrow \infty} \sum_{k=0}^{m-1} s_k^{-2} \left| \langle \tilde{\Delta}_n x, \phi_k \rangle_{\mathcal{H}_0} \right|^2 = 0 .$$

On the other hand, for all  $m, n \geq 1$ , using the fact that  $\|\tilde{\Delta}_n\|_{\infty} \leq \|\tilde{\Delta}_n\|_1 \leq 2$  and that  $s$  is non-decreasing, we get that

$$\begin{aligned} \sum_{k=m}^{\infty} s_k^{-2} \left| \langle \tilde{\Delta}_n x, \phi_k \rangle_{\mathcal{H}_0} \right|^2 &\leq s_m^{-2} \sum_{k=0}^{\infty} \left| \langle \tilde{\Delta}_n x, \phi_k \rangle_{\mathcal{H}_0} \right|^2 \\ &= s_m^{-2} \|\tilde{\Delta}_n x\|_{\mathcal{H}_0^s}^2 \leq 4 s_m^{-2} \|x\|_{\mathcal{H}_0}^2 , \end{aligned}$$

hence converges to 0 independently of  $n$  as  $m \rightarrow \infty$ , by assumption on  $s$ . With the two previous displays, we conclude that  $(\Delta_n x)_{n \in \mathbb{N}}$  converges to 0 in  $\mathcal{H}_0$ . Hence,  $(\Delta_n)_{n \in \mathbb{N}}$  converges to 0 for the strong operator topology.

**Step 2.** Let  $n \in \mathbb{N}$ . Since  $\tilde{\Delta}_n \in \mathcal{S}_1(\mathcal{H}_0)$  and  $\mathbb{J}_s \in \mathcal{L}_b(\mathcal{H}_0)$ , we have  $\Delta_n \in \mathcal{S}_1(\mathcal{H}_0)$ . Consider the polar decomposition of  $\Delta_n$ , that is  $\Delta_n = V_n |\Delta_n|$  where  $V_n$  is a partial isometry with initial space  $\ker(\Delta_n)^\perp = \overline{\text{Im}}(|\Delta_n|)$  and final space  $\overline{\text{Im}}(\Delta_n)$  (see §3.9 in [7]). Since  $\Delta_n$  is autoadjoint we have  $\ker(\Delta_n)^\perp = \overline{\text{Im}}(\Delta_n)$  and we get that  $\overline{\text{Im}}(|\Delta_n|) = \overline{\text{Im}}(\Delta_n) \subset \overline{\text{Im}}(\mathbb{J}_s) = \overline{\text{Span}}(\phi_k, k \in \mathbb{N})$ . Hence

$$\|\Delta_n\| = \text{Tr}(|\Delta_n|) = \sum_{k \in \mathbb{N}} \langle |\Delta_n| \phi_k, \phi_k \rangle_{\mathcal{H}_0} = \sum_{k \in \mathbb{N}} \langle \Delta_n \phi_k, V_n \phi_k \rangle_{\mathcal{H}_0} , \quad (5.41)$$

where the last equality comes from the fact that  $|\Delta_n| = V_n^H \Delta_n$ .

Now, note that for all  $m \geq 1$ ,

$$\sum_{k=0}^{m-1} \left| \langle \Delta_n \phi_k, V_n \phi_k \rangle_{\mathcal{H}_0} \right| \leq \sum_{k=0}^{m-1} \|\Delta_n \phi_k\|_{\mathcal{H}_0} ,$$

which converges to zero by Step 1. Thus,

$$\text{for all } m \geq 1, \lim_{n \rightarrow \infty} \sum_{k=0}^{m-1} \langle \Delta_n \phi_k, V_n \phi_k \rangle_{\mathcal{H}_0} = 0 . \quad (5.42)$$

On the other hand, using the fact that  $\Delta_n$  is hermitian and (5.40), we have, for all  $n, k \in \mathbb{N}$ ,

$$\begin{aligned} \langle \Delta_n \phi_k, V_n \phi_k \rangle_{\mathcal{H}_0} &= \langle \phi_k, \Delta_n V_n \phi_k \rangle_{\mathcal{H}_0} = \left\langle \phi_k, \mathbb{J}_s \tilde{\Delta}_n V_n \phi_k \right\rangle_{\mathcal{H}_0} = s_k^{-1} \left\langle \phi_k, \tilde{\Delta}_n V_n \phi_k \right\rangle_{\mathcal{H}_0} \\ &= s_k^{-1} \left\langle \tilde{\Delta}_n \phi_k, V_n \phi_k \right\rangle_{\mathcal{H}_0} . \end{aligned}$$

It follows that, for all  $m \geq 1$ ,

$$\sum_{k \geq m} \left| \langle \Delta_n \phi_k, V_n \phi_k \rangle_{\mathcal{H}_0} \right| \leq s_m^{-1} \sum_{k \in \mathbb{N}} \left| \langle \tilde{\Delta}_n \phi_k, V_n \phi_k \rangle_{\mathcal{H}_0} \right| \leq s_m^{-1} \|\tilde{\Delta}_n\|_1 ,$$

where we used [7, Corollary 18.12]. Since  $\|\tilde{\Delta}_n\|_1 \leq 1$  and  $s_m^{-1}$  converges to 0, we obtain that

$$\lim_{m \rightarrow \infty} \sup_{n \in \mathbb{N}} \sum_{k \geq m} \left| \langle \Delta_n \phi_k, V_n \phi_k \rangle_{\mathcal{H}_0} \right| = 0 .$$

This with (5.41) and (5.42) concludes the second and final step.  $\square$

We can now prove Theorem 4.3.

**Proof of Theorem 4.3.** By the polarization formula we can write  $\tilde{Q}_{I_n^{X_n^c}}^{(L,R)}$  as a linear combination of  $\tilde{Q}_{I_n^{X_n^c}}^{(W,W)}$  with  $W$  in  $\{L+R, L-R, L+iR, L-iR\}$ . The same formula holds for expressing  $\tilde{Q}_{\nu^X}^{(L,R)}$  using  $\tilde{Q}_{\nu^X}^{(W,W)}$  with the same  $W$ 's. Hence, to obtain the claimed result, it suffices to show that, for all  $W \in \{L+R, L-R, L+iR, L-iR\}$ , we have

$$\lim_{n \rightarrow \infty} \tilde{Q}_{I_n^{X_n^c}}^{(W,W)} = \tilde{Q}_{\nu^X}^{(W,W)} \quad \text{uniformly in } \mathcal{C}(\Theta, \mathcal{S}_1(\mathcal{H}_0)) \text{ , } \mathbb{P}\text{-a.s.} \quad (5.43)$$

So, take  $W \in \{L+R, L-R, L+iR, L-iR\}$ , and let us show (5.43).

By assumption on  $L$  and  $R$ , we have  $W \in \mathcal{C}(\Theta \times \mathbb{T}, \mathcal{L}_b(\mathcal{H}_0))$  and, using Condition (iii),  $W_s \in \mathcal{C}(\Theta \times \mathbb{T}, \mathcal{L}_b(\mathcal{H}_0^s))$  where  $W_s(\theta, \lambda) = W(\theta, \lambda)|_{\mathcal{H}_0^s}$  for all  $(\theta, \lambda) \in \Theta \times \mathbb{T}$ . By Condition (ii), we can apply Lemma 5.5 with  $\mathcal{I}_0 = \mathcal{H}_0^s$ , and obtain that,  $\mathbb{P}$ -a.s.,

$$\text{there exists } r_1 > 0 \text{ such that } \left\{ I_n^{X_n^c} : n \geq 1 \right\} \subset B_{1,1}(r_1, \mathcal{H}_0, \mathcal{H}_0^s) . \quad (5.44)$$

Now let us define  $J_s, U_s$  and  $\mathbb{J}_s$  as in Lemma 5.10. Applying these definitions carefully and using the fact that  $UU^H$  is the orthogonal projection onto  $\text{Im}(U_s) = \mathcal{H}_0^s$ , it straightforwardly yields that, for all  $(\theta, \lambda) \in \Theta \times \mathbb{T}$ , and  $x \in \mathcal{H}_0^s$ ,

$$\mathbb{J}_s U_s^H W_s(\theta, \lambda) x = J_s W_s(\theta, \lambda) x = W(\theta, \lambda) x .$$

Thus, for all  $n \in \mathbb{N}$ ,  $\mathbb{P}$ -a.s.,  $I_n^X \in L^{1,1}(\mathcal{H}_0, \mathcal{H}_0^s)$ , and for all  $\theta \in \Theta$ ,

$$\tilde{Q}_{I_n^X}^{(W,W)}(\theta) = \mathbb{J}_s \tilde{Q}_{I_n^X}^{(U_s^H W_s, W)}(\theta) . \quad (5.45)$$

Observe that  $W_s \in \mathcal{C}(\Theta \times \mathbb{T}, \mathcal{L}_b(\mathcal{H}_0^s))$  immediately implies that  $U_s^H W_s(\theta, \cdot) \in \mathbb{F}_b(\mathcal{H}_0^s, \mathcal{H}_0)$ . Thus, for all  $\theta$ , we can apply Lemma 5.3 with  $\mathcal{I}_0 = \mathcal{H}_0^s$  and  $\mathcal{G}_0 = \mathcal{J}_0 = \mathcal{H}_0$ ,  $L = U_s^H W_s(\theta, \cdot)$ ,  $R = W$ ,  $g = I_n^X$  and  $\mu = \text{Leb}_{\mathbb{T}}$ , which, with (5.44), gives us that,  $\mathbb{P}$ -a.s.: for all  $\theta \in \Theta$ , there exists  $r > 0$  such that  $\left\| \tilde{Q}_{I_n^X}^{(U_s^H W_s, W)}(\theta) \right\|_1 \leq r$ , where, here,  $\|\cdot\|_1$  denotes the trace-class norm in  $\mathcal{S}_1(\mathcal{H}_0)$ . With (5.45) and the fact that  $\tilde{Q}_{I_n^X}^{(W,W)}(\theta)$  is an hermitian operator for all  $\theta$ , we get that,  $\mathbb{P}$ -a.s.: for all  $\theta \in \Theta$ , there exists  $r > 0$  such that  $\tilde{Q}_{I_n^X}^{(W,W)}(\theta)$  belongs to the set  $B = r B_1$ , with  $B_1$  defined as in Proposition 5.11. Since  $B$  is compact by Proposition 5.11, we can apply Theorem 5.9 with  $L = R := W$  and we obtain (5.43), which concludes the proof.  $\square$

## 5.4 Proof of Theorem 1.1

Hereafter we prove (1.4) by relying on Theorem 4.3. Of course the uniform convergence in  $\theta$  does not matter in this case and it could be proven by relying directly on Proposition 5.11, by mimicking the argument used in the proof of Theorem 4.3 to show that  $\left( \hat{\Gamma}_n(h) \right)_{n \in \mathbb{N}}$  remains in a compact subset of  $\mathcal{S}_1(\mathcal{H}_0)$ ,  $\mathbb{P}$ -a.s. (since the convergence of  $\hat{\Gamma}_n(h)$  to  $\Gamma(h) = \text{Cov}(X_h, X_0)$  in weak operator topology is obvious under the assumptions of Theorem 1.1).

**Proof of Theorem 1.1.** Let  $h$  be a given lag in  $\mathbb{Z}$ . The claimed result follows from Theorem 4.3 with  $L(\theta, \lambda) = e^{ih\lambda} \text{Id}_{\mathcal{H}_0}$  and  $R(\theta, \lambda) = \text{Id}_{\mathcal{H}_0}$ , with  $\theta$  being an arbitrary point and  $\Theta$  the corresponding singleton. Indeed, with these definitions, the convergence (1.4) can be rewritten as (5.31). Thus, we only to show that  $(X_t)_{t \in \mathbb{Z}}$  and the above defined  $L$  and  $R$  satisfy the assumptions of Theorem 4.3. Obviously **(A-1)** and **(A-2)** are satisfied. As for Conditions (i) and (ii), they follow from **(A-1)** and **(A-2)** by Lemma 5.7, for well chosen  $s$  and  $(\phi_k)_{k \in \mathbb{N}}$ . Now, defining  $L_s$  and  $R_s$  as in Theorem 4.3 with  $L$  and  $R$  as above, we get  $L_s(\theta, \lambda) = e^{ih\lambda} \text{Id}_{\mathcal{H}_0^s}$  and  $R_s(\theta, \lambda) = \text{Id}_{\mathcal{H}_0^s}$ , and  $L, R, L_s$  and  $R_s$  obviously satisfy the assumptions of the theorem (including Condition (iii)). Hence Theorem 4.3 applies and the proof is finished.  $\square$

## 5.5 Proofs of Section 4.2

### 5.5.1 Preliminary results

**Lemma 5.12.** *Let  $\mathcal{H}_0$  be a separable Hilbert space and  $p, q$  be two non-negative integers. Let  $D \in \mathcal{L}_b(\mathcal{H}_0)$ ,  $\theta \in \mathcal{P}_q^\dagger(\mathcal{H}_0)$  and  $\phi \in \mathcal{P}_p^\dagger(\mathcal{H}_0)$ . Define  $\Phi_{\theta, \phi, D}^\dagger : \mathbb{C} \setminus [1, \infty) \rightarrow \mathcal{L}_b(\mathcal{H}_0)$  by (4.11). Then,  $\theta^{-1} : z \mapsto [\theta(z)]^{-1}$ ,  $\phi^{-1} : z \mapsto [\phi(z)]^{-1}$  and  $\Phi_{\theta, \phi, D}^\dagger$  are all holomorphic functions on the open unit disk  $\mathbb{D}$  onto  $\mathcal{L}_b(\mathcal{H}_0)$ . Moreover  $\theta^{-1}$  and  $\phi^{-1}$  are continuous on the closed unit disk  $\overline{\mathbb{D}}$  and  $\Phi_{\theta, \phi, D}^\dagger$  is continuous over  $\overline{\mathbb{D}} \setminus \{1\}$ .*

*Proof.* By holomorphic in Lemma 5.12 we mean the same as in [13, Definition 1.1.1]. Since  $\phi$  and  $\theta$  are polynomials they are holomorphic in  $\mathbb{C}$ . Because we assumed that they belong to  $\mathcal{P}_p^\dagger(\mathcal{H}_0)$ , we further have that they are valued in the space of invertible operators on  $\overline{\mathbb{D}}$  and so the inverted polynomials  $z \mapsto [\theta(z)]^{-1}$  and  $z \mapsto [\phi(z)]^{-1}$  are holomorphic on  $\mathbb{D}$  and continuous on  $\overline{\mathbb{D}}$ . Since the principal logarithmic function is holomorphic on  $\mathbb{C} \setminus [1, \infty)$ , so is  $z \mapsto (1-z)^D$ . The result follows.  $\square$

**Lemma 5.13.** *Let  $\mathcal{H}_0$  be a separable Hilbert space and  $p, q$  be two non-negative integers. Let  $D \in \mathcal{N}(\mathcal{H}_0)$ ,  $\theta \in \mathcal{P}_q^\dagger(\mathcal{H}_0)$  and  $\phi \in \mathcal{P}_p(\mathcal{H}_0)$ . Define  $\Phi_{\theta, \phi, D}^\dagger : \mathbb{C} \setminus [1, \infty) \rightarrow \mathcal{L}_b(\mathcal{H}_0)$  by (4.11). Let moreover  $X \in \mathcal{S}_{\text{FI}_D}(\Omega, \mathcal{F}, \mathbb{P})$  and denote by  $g_X$  its spectral operator density with respect to a non-negative measure  $\mu$  on  $(\mathbb{T}, \mathcal{B}(\mathbb{T}))$ . We assume that  $\mu$  has no mass at the origin,  $\mu(\{0\}) = 0$ . Then, we have*

$$\int_{-\pi/3}^{\pi/3} \sup_{0 \leq \rho \leq 1} \left\| \Psi(\rho, \lambda) g_X(\lambda) \Psi^H(\rho, \lambda) \right\|_1 \mu(d\lambda) < \infty, \quad (5.46)$$

in the two following cases:

$$\text{if } \Psi(\rho, \lambda) := \left(1 - \rho e^{-i\lambda}\right)^{-D} - \left(1 - e^{-i\lambda}\right)^{-D}. \quad (5.47)$$

$$\text{or if } \Psi(\rho, \lambda) := \left(\Phi_{\theta, \phi, D}^\dagger(\rho e^{-i\lambda}) - \Phi_{\theta, \phi, D}^\dagger(e^{-i\lambda})\right) \left(1 - e^{-i\lambda}\right)^{-D}. \quad (5.48)$$

*Proof.* We first consider  $\Psi(\rho, \lambda)$  as in (5.47). In this case, for all  $0 \leq \rho \leq 1$  and  $\lambda \neq 0$ ,

$$\left\| \Psi(\rho, \lambda) g_X(\lambda) \Psi^H(\rho, \lambda) \right\|_1 = \left\| \Psi(\rho, \lambda) (g_X(\lambda))^{1/2} \right\|_2^2 \leq 4 \sup_{0 \leq \rho \leq 1} \left\| \left(1 - \rho e^{-i\lambda}\right)^{-D} (g_X(\lambda))^{1/2} \right\|_2^2.$$

Thus, the bound (5.46) is implied by

$$\int_{-\pi/3}^{\pi/3} \sup_{0 \leq \rho \leq 1} \left\| \left(1 - \rho e^{-i\lambda}\right)^{-D} (g_X(\lambda))^{1/2} \right\|_2^2 \mu(d\lambda) < \infty, \quad (5.49)$$

Since  $D$  is assumed to be normal, we can proceed as in the proof of Theorem 3.3 and use its singular value function  $d$  on  $\mathcal{G}_0 := L^2(\mathbb{V}, \mathcal{V}, \xi)$  and decomposition operator  $U$  so that (5.49) is implied by

$$\int_{\mathbb{V}^2 \times (-\pi/3, \pi/3)} \sup_{0 \leq \rho \leq 1} \left| \left(1 - \rho e^{-i\lambda}\right)^{-d(v)} \mathcal{K}(v, v'; \lambda) \right|^2 \xi(dv) \xi(dv') \mu(d\lambda) < \infty,$$

where  $\mathcal{K}$  denote the  $\mathbb{T}$ -joint kernel function of  $h$  such that  $h(\lambda)[h(\lambda)]^H = U g_X(\lambda) U^H$  for  $\mu$ -a.e.  $\lambda \in \mathbb{T}$ . Now, by Lemma A.2, using that  $d$  is bounded over  $\mathbb{V}$ , the previous condition holds if

$$\int_{\mathbb{V}^2 \times (-\pi/3, \pi/3)} |\lambda|^{-2\Re_+(d(v))} \left| \mathcal{K}(v, v'; \lambda) \right|^2 \xi(dv) \xi(dv') \mu(d\lambda) < \infty. \quad (5.50)$$

On the other hand, we assumed that  $X \in \mathcal{S}_{\text{FI}_D}(\Omega, \mathcal{F}, \mathbb{P})$ , which, by Theorem 3.3 is equivalent to have (3.7), which implies (5.50) (since  $\mu(\{0\}) = 0$ ). We thus proved (5.46) for  $\Psi(\rho, \lambda)$  as in (5.47).

We now consider  $\Psi(\rho, \lambda)$  as in (5.48). By (4.11), for all  $\lambda \in \mathbb{T} \setminus \{0\}$  and  $0 \leq \rho \leq 1$ , the difference  $\Phi_{\theta, \phi, D}^\dagger(\rho e^{-i\lambda}) - \Phi_{\theta, \phi, D}^\dagger(e^{-i\lambda})$  can be written as

$$\left[\theta(e^{-i\lambda})\right]^{-1} \phi(e^{-i\lambda}) (1 - e^{-i\lambda})^D - \left[\theta(\rho e^{-i\lambda})\right]^{-1} \phi(\rho e^{-i\lambda}) (1 - \rho e^{-i\lambda})^D.$$

By Lemma 5.12,  $\theta^{-1}$  and  $\phi$  are continuous on  $\overline{\mathbb{D}}$ , hence have bounded operator norms over  $\overline{\mathbb{D}}$ . Thus, to get (5.46), it is thus sufficient to show that

$$\int_{-\pi/3}^{\pi/3} \sup_{0 \leq \rho \leq 1} \left\| \tilde{\Psi}(\rho, \lambda) g_X(\lambda) \tilde{\Psi}^H(\rho, \lambda) \right\|_1 \mu(d\lambda) < \infty, \quad (5.51)$$

$$\text{where } \tilde{\Psi}(\rho, \lambda) := \left(1 - \rho e^{-i\lambda}\right)^D \left(1 - e^{-i\lambda}\right)^{-D}. \quad (5.52)$$

Since  $D$  is assumed to be normal, we can proceed as in the proof of Theorem 3.3 and use its singular value function  $d$  on  $\mathcal{G}_0 := L^2(\mathcal{V}, \mathcal{V}, \xi)$  and decomposition operator  $U$  so that (5.51) is implied by

$$\int_{\mathcal{V}^2 \times (-\pi/3, \pi/3)} \sup_{0 \leq \rho \leq 1} \left| \left( \frac{1 - \rho e^{-i\lambda}}{1 - e^{-i\lambda}} \right)^{d(v)} \mathcal{R}(v, v'; \lambda) \right|^2 \xi(dv) \xi(dv') \mu(d\lambda) < \infty. \quad (5.53)$$

Now, by Lemma A.1 and Lemma A.2, we have for all  $z \in \mathbb{C}$  and  $\lambda \in [-\pi/3, \pi/3] \setminus \{0\}$ ,

$$\sup_{0 \leq \rho \leq 1} \left| \left( \frac{1 - \rho e^{-i\lambda}}{1 - e^{-i\lambda}} \right)^z \right| \leq \left( 2\pi / (3\sqrt{3}) \right)^{\Re_-(z)} (\pi/2)^{\Re_-(z)} |\lambda|^{-\Re_+(z)} e^{\pi|\Im(z)|}.$$

Plugging this bound in (5.53) and using that  $d$  is bounded on  $\mathcal{V}$ , we get that (5.53) is again implied by (5.50). Hence we proved (5.46) in the case given by (5.48).  $\square$

In the following, for any positive integer  $p$ , we endow the set of polynomials of degree less than or equal to  $p$  (or any of its subsets  $\mathcal{P}_p(\mathcal{H}_0)$ ,  $\mathcal{P}_p^\dagger(\mathcal{H}_0)$  or  $\mathcal{P}_p^*(\mathcal{H}_0)$ ) with the max of the  $\|\cdot\|_\infty$ -norms of its  $\mathcal{L}_b(\mathcal{H}_0)$  coefficients. For instance if  $\psi(z) := \sum_{k=0}^p A_k z^k$  we denote

$$\|\psi\| = \max \{ \|A_k\|_\infty : k = 0, \dots, p \}.$$

It is straightforward to show that the convergence of a  $\mathcal{P}_p(\mathcal{H}_0)$ -valued sequence in the obtained Banach space is equivalent to the uniform convergence of this sequence in  $\mathcal{C}(\mathbb{U}, \mathcal{L}_b(\mathcal{H}_0))$ . In particular the continuity of  $(\theta, \lambda) \mapsto \phi_\theta(e^{-i\lambda})$  and  $(\theta, \lambda) \mapsto \theta_\theta(e^{-i\lambda})$  on  $\Theta \times \mathbb{T}$  onto  $\mathcal{L}_b(\mathcal{H}_0)$  assumed in **(A-3)** imply the convergence of  $\theta \mapsto \phi_\theta$  and  $\theta \mapsto \theta_\theta$  on  $\Theta$  onto  $\mathcal{P}_p(\mathcal{H}_0)$ .

We have the following lemma.

**Lemma 5.14.** *Let  $\mathcal{H}_0$  be a separable Hilbert space and  $p, q$  be two non-negative integers. Let  $\Phi_{\theta, \phi, D}^\dagger$  be defined by (4.11). Then, for all  $(\theta, \phi, D) \in \mathcal{P}_q^\dagger(\mathcal{H}_0) \times \mathcal{P}_p(\mathcal{H}_0) \times \mathcal{L}_b(\mathcal{H}_0)$  and all  $k \in \mathbb{N}$ ,*

$$P_k^\dagger(\theta, \phi, D) := \frac{1}{2i\pi} \int_{z \in \mathbb{C}, |z|=\rho} \Phi_{\theta, \phi, D}^\dagger(z) z^{-k-1} dz \quad (5.54)$$

is well defined as  $\mathcal{L}_b(\mathcal{H}_0)$ -valued Bochner integral for any  $\rho \in (0, 1)$  and does not depend on  $\rho$ . Moreover, the following assertions hold.

(i) For all  $(\theta, \phi, D) \in \mathcal{P}_q^\dagger(\mathcal{H}_0) \times \mathcal{P}_p(\mathcal{H}_0) \times \mathcal{L}_b(\mathcal{H}_0)$  and  $z \in \mathbb{D}$ ,

$$\Phi_{\theta, \phi, D}^\dagger(z) = \sum_{k=1}^{\infty} P_k^\dagger(\theta, \phi, D) z^k. \quad (5.55)$$

(ii) For any  $k \geq 1$ ,  $(\theta, \phi, D) \mapsto P_k^\dagger(\theta, \phi, D)$  is continuous on  $\mathcal{P}_q^\dagger(\mathcal{H}_0) \times \mathcal{P}_p(\mathcal{H}_0) \times \mathcal{L}_b(\mathcal{H}_0)$ .

(iii) For all  $(\theta, \phi, D) \in \mathcal{P}_q^\dagger(\mathcal{H}_0) \times \mathcal{P}_p(\mathcal{H}_0) \times \mathcal{N}^\dagger(\mathcal{H}_0)$  and  $z \in \overline{\mathbb{D}} \setminus \{1\}$ ,

$$\Phi_{\theta, \phi, D}^\dagger(z) = \sum_{k=1}^{\infty} P_k^\dagger(\theta, \phi, D) z^k. \quad (5.56)$$

(iv) For any compact subset  $K \subset \mathcal{P}_q^\dagger(\mathcal{H}_0) \times \mathcal{P}_p(\mathcal{H}_0) \times (\{0\} \cup \mathcal{N}^\dagger(\mathcal{H}_0))$ , we have

$$\sum_{k=1}^{\infty} \sup_{(\theta, \phi, D) \in K} \left\| P_k^\dagger(\theta, \phi, D) \right\|_\infty < \infty. \quad (5.57)$$



*Proof.* Recall that  $\Phi_{\theta, \phi, D}^\dagger$  is defined for all  $(\theta, \phi, D) \in \mathcal{P}_q^\dagger(\mathcal{H}_0) \times \mathcal{P}_p(\mathcal{H}_0) \times \mathcal{L}_b(\mathcal{H}_0)$  by (4.11) as a holomorphic function defined on  $\mathbb{C} \setminus [1, \infty)$  onto the Banach space  $\mathcal{L}_b(\mathcal{H}_0)$ . Using Lemma 5.12, and [13, Theorem 1.8.5], we can expand  $\Phi_{\theta, \phi, D}^\dagger$  as a power series on the open unit disk  $\mathbb{D}$ , that is (5.55) holds with  $P_k^\dagger(\theta, \phi, D)$  (well) defined by (5.54) for any  $\rho \in (0, 1)$ . Note that the sum in the right-hand side of (5.55) starts at  $k = 1$  because  $P_0^\dagger(\theta, \phi, D) = 0$  since, by the Cauchy Formula (see [13, Theorem 1.5.1]), we have  $P_0^\dagger(\theta, \phi, D) = \Phi_{\theta, \phi, D}^\dagger(0)$ , which is the null operator following (4.11) and  $\phi(0) = \phi(0) = 1^D = \text{Id}_{\mathcal{H}_0}$ .

Assertion (ii) follows from (5.54) by dominated convergence, since  $(z, \theta, \phi, D) \mapsto \Phi_{\theta, \phi, D}^\dagger(z)$  is continuous on  $\mathbb{D} \times \mathcal{P}_q^\dagger(\mathcal{H}_0) \times \mathcal{P}_p(\mathcal{H}_0) \times \mathcal{L}_b(\mathcal{H}_0)$  by (4.11) and Lemma 5.12.

Let us now prove Assertions (iii) and (iv). In fact, Assertions (ii) and (iv) imply that the right-hand side of (5.56) is continuous on  $\overline{\mathbb{D}}$ . Since, by Lemma 5.12, the left-hand side is continuous on  $\overline{\mathbb{D}} \setminus \{1\}$ , with Assertion (i), we conclude that we get both Assertions (iii) and (iv) by proving the bound (5.57). Let  $K$  be a compact subset of  $\mathcal{P}_q^\dagger(\mathcal{H}_0) \times \mathcal{P}_p(\mathcal{H}_0) \times \{0\}$  or of  $\mathcal{P}_q^\dagger(\mathcal{H}_0) \times \mathcal{P}_p(\mathcal{H}_0) \times \mathcal{N}^\dagger(\mathcal{H}_0)$ . Then, there exists  $r > 1$  such that for all  $(\theta, \phi, D) \in K$ ,  $\theta$  does not vanish over the open disk of radius  $r$ . It follows that for all  $(\theta, \phi, D) \in K$ ,  $z \mapsto [\theta]^{-1}(z)\phi(z)$  is a power series with a radius of convergence at least equal to  $r$  and that, for any  $\rho_1 \in (r^{-1}, 1)$ , there exists  $c_1 > 0$  such that, for all  $(\theta, \phi, D) \in K$  and  $z \in r\mathbb{D}$ ,

$$[\theta]^{-1}(z)\phi(z) = \text{Id}_{\mathcal{H}_0} + \sum_{k=1}^{\infty} C_k z^k \quad \text{with} \quad \|C_k\|_\infty \leq c_1 \rho_1^k \quad (5.58)$$

If  $K \subset \mathcal{P}_q^\dagger(\mathcal{H}_0) \times \mathcal{P}_p(\mathcal{H}_0) \times \{0\}$ , we have  $P_k^\dagger(\theta, \phi, D) = -C_k$  and (5.57) follows.

We now consider the case where  $K \subset \mathcal{P}_q^\dagger(\mathcal{H}_0) \times \mathcal{P}_p(\mathcal{H}_0) \times \mathcal{N}^\dagger(\mathcal{H}_0)$ . Let  $\sigma$  be an upper bound of  $\|D\|_\infty$  and  $\rho$  a lower bound of the smallest eigenvalue of  $(D + D^H)/2$  over  $(\theta, \phi, D) \in K$ . Then, we have  $\varrho > 0$  by definition of  $\mathcal{N}^\dagger(\mathcal{H}_0)$  in (4.13) and since  $K$  is compact. Then, setting  $N = D + \text{Id}_{\mathcal{H}_0}$ , Condition (5.14) holds with  $\varrho = \rho + 1 > 1$ . and  $\varsigma = \sigma + 1$ . Thus, by Lemma 5.2, we have, for all  $z \in \overline{\mathbb{D}} \setminus \{1\}$  and  $D \in \mathcal{N}^\dagger(\mathcal{H}_0)$ ,

$$(1 - z)^D = \sum_{k=0}^{\infty} \left( Q(k+1)^{-\text{Id}_{\mathcal{H}_0} - D} + P_k^* \right) z^k,$$

where  $\|Q\|_\infty \leq C$  and  $\|P_k^*\| \leq Ck^{-2-\rho}$  for some constant  $C > 0$  only depending on  $\sigma$ . Moreover, note that  $\|(k+1)^{-\text{Id}_{\mathcal{H}_0} - D}\|_\infty \leq (k+1)^{-1-\rho}$ . By (4.11) and setting  $C_0 = \text{Id}_{\mathcal{H}_0}$ , we obtain, for all  $z \in \overline{\mathbb{D}} \setminus \{1\}$ ,

$$\Phi_{\theta, \phi, D}^\dagger(z) = - \sum_{\ell=1}^{\infty} \left( \sum_{k=0}^{\ell} C_{\ell-k} \left( Q(k+1)^{-\text{Id}_{\mathcal{H}_0} - D} + P_k^* \right) \right) z^\ell.$$

By (4.11), we thus have, for all  $k \geq 1$ ,

$$P_k^\dagger(\theta, \phi, D) = - \left( \sum_{k=0}^{\ell} C_{\ell-k} \left( Q(k+1)^{-\text{Id}_{\mathcal{H}_0} - D} + P_k^* \right) \right),$$

and the previous bounds further yield

$$\sup_{(\theta, \phi, D) \in K} \left\| P_k^\dagger(\theta, \phi, D) \right\|_\infty \leq 2c_1 C \sum_{k=0}^{\ell} \rho_1^{\ell-k} (k+1)^{-1-\rho}.$$

Therefore, we obtain (5.57).  $\square$

Finally, the following lemma will be useful.

**Lemma 5.15.** *Let  $\mathcal{H}_0$  be a separable Hilbert space and  $X = (X_t)_{t \in \mathbb{Z}}$  be an ergodic stationary process defined on  $(\Omega, \mathcal{F}, \mathbb{P})$  valued in  $\mathcal{H}_0$  such that  $\mathbb{E} \left[ \|X\|_{\mathcal{H}_0}^2 \right] < \infty$ . Then we have*

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[ \left\| \frac{1}{n} \sum_{k=1}^n X_k - \mathbb{E}[X] \right\|_{\mathcal{H}_0}^2 \right] = 0.$$

*Proof.* The assumptions imply that  $X$  is weakly stationary. Moreover, the space of shift-invariant elements in  $\mathcal{H}^X$  (where the shift is defined by  $X_t \mapsto X_{t-1}$ ) is the null set, otherwise  $X$  would not be ergodic (take  $V$  shift-invariant in the sense of  $\mathcal{H}^X$ , then, for all  $x \in \mathcal{H}_0$ ,  $\langle V, x \rangle_{\mathcal{H}_0}$  is shift-invariant in the  $\mathbb{P}^X$ -a.s. sense). Therefore, the result simply follows from the von Neumann ergodic theorem.  $\square$

### 5.5.2 Proof of main results

**Proof of Theorem 4.4.** By Definition 3.2, denoting by  $\Sigma$  the covariance operator of  $Z$ ,  $Y$  admits the spectral density

$$f^Y(\lambda) = \left(1 - e^{-i\lambda}\right)^{-D} [\Phi(e^{-i\lambda})]^{-1} \Theta(e^{-i\lambda}) \Sigma \left( \left(1 - e^{-i\lambda}\right)^{-D} [\Phi(e^{-i\lambda})]^{-1} \Theta(e^{-i\lambda}) \right)^H$$

with respect to the normalized Lebesgue measure  $\text{Leb}_{\mathbb{T}}$ . Let  $t \in \mathbb{Z}$ . We first show that the right-hand side of (4.12) is well defined, that is, that  $\lambda \mapsto e^{i\lambda t} \Phi_{\Theta, \Phi, D}^\dagger(e^{-i\lambda})$  belongs to  $\mathcal{H}^Y$ . Since this mapping is continuous on  $\mathbb{T} \setminus \{0\}$  onto  $\mathcal{L}_b(\mathcal{H}_0)$ , by Proposition 2.1, this is equivalent to have

$$\int \left\| \Phi_{\Theta, \Phi, D}^\dagger(e^{-i\lambda}) f^Y(\lambda) \Phi_{\Theta, \Phi, D}^\dagger(e^{-i\lambda}) \right\|_1^H \text{Leb}_{\mathbb{T}}(d\lambda) < \infty.$$

By definition of  $f^Y$ , we thus have to show that

$$\int \left\| \Phi_{\Theta, \Phi, D}^\dagger(e^{-i\lambda}) \left(1 - e^{-i\lambda}\right)^{-D} [\Phi(e^{-i\lambda})]^{-1} \Theta(e^{-i\lambda}) \Sigma^{1/2} \right\|_2^2 \text{Leb}_{\mathbb{T}}(d\lambda) < \infty. \quad (5.59)$$

By definition of  $\Phi_{\Theta, \Phi, D}^\dagger$  in (4.11), we have, for all  $\lambda \in \mathbb{T} \setminus \{0\}$ ,

$$\left( \text{Id}_{\mathcal{H}_0} - \Phi_{\Theta, \Phi, D}^\dagger(e^{-i\lambda}) \right) \left(1 - e^{-i\lambda}\right)^{-D} [\Phi(e^{-i\lambda})]^{-1} \Theta(e^{-i\lambda}) = \text{Id}_{\mathcal{H}_0}.$$

We thus get that, for all  $\lambda \in \mathbb{T} \setminus \{0\}$ ,  $\Phi_{\Theta, \Phi, D}^\dagger(e^{-i\lambda}) \left(1 - e^{-i\lambda}\right)^{-D} [\Phi(e^{-i\lambda})]^{-1} \Theta(e^{-i\lambda}) \Sigma^{1/2}$  can be expressed as

$$\left(1 - e^{-i\lambda}\right)^{-D} [\Phi(e^{-i\lambda})]^{-1} \Theta(e^{-i\lambda}) \Sigma^{1/2} - \Sigma^{1/2}.$$

Thus, since  $\left\| \Sigma^{1/2} \right\|_2^2 = \|\Sigma\|_1 < \infty$ , Condition (5.59) is implied by

$$\int \left\| \left(1 - e^{-i\lambda}\right)^{-D} [\Phi(e^{-i\lambda})]^{-1} \Theta(e^{-i\lambda}) \Sigma^{1/2} \right\|_2^2 \text{Leb}_{\mathbb{T}}(d\lambda) < \infty.$$

On the other hand, the square  $\mathcal{S}_2$ -norm inside the previous integral is equal to  $\|f^Y\|_1$  which is  $\text{Leb}_{\mathbb{T}}$ -integrable as a spectral density. We thus get that the right-hand side of (4.12) is well defined and, in the following, we denote

$$\hat{Y}_t = \int e^{i\lambda t} \Phi_{\Theta, \Phi, D}^\dagger(e^{-i\lambda}) \hat{Y}(d\lambda).$$

It only remains to show the two following assertions.

(i) We have  $\hat{Y}_t \in \mathcal{H}_{t-1}^Y$ .

(ii) We have  $Y_t - \hat{Y}_t \perp \mathcal{H}_{t-1}^Y$ .

Let us first prove Assertion (i). By Assertion (i) in Lemma 5.14, we immediately have that, for all  $\rho \in (0, 1)$ ,

$$\hat{Y}_t^{(\rho)} = \int e^{i\lambda t} \Phi_{\Theta, \Phi, D}^\dagger(\rho e^{-i\lambda}) \hat{Y}(d\lambda) \in \mathcal{H}_{t-1}^Y.$$

To conclude Assertion (i), it is thus sufficient to show that

$$\lim_{\rho \uparrow 1} \hat{Y}_t^{(\rho)} = Y_t \quad \text{in } \mathcal{M}(\Omega, \mathcal{F}, \mathcal{H}_0, \mathbb{P}). \quad (5.60)$$

By the Kolmogorov Gramian isometric theorem see [11, Theorem 4.3], setting

$$\Psi(\rho, \lambda) := \Phi_{\Theta, \Phi, D}^\dagger(\rho e^{-i\lambda}) - \Phi_{\Theta, \Phi, D}^\dagger(e^{-i\lambda}),$$

we can write, for any  $\eta \in (0, \pi)$  and  $\rho \in (0, 1)$ ,

$$\begin{aligned} \mathbb{E} \left[ \left\| \hat{Y}_t^{(\rho)} - \hat{Y}_t \right\|_{\mathcal{H}_0}^2 \right] &= \int \left\| \Psi(\rho, \lambda) f^Y(\lambda) \Psi^H(\rho, \lambda) \right\|_1 \text{Leb}_{\mathbb{T}}(d\lambda) \\ &\leq \left( \int \left\| f^Y \right\|_1 d\text{Leb}_{\mathbb{T}} \right) \sup_{\lambda \in \mathbb{T} \setminus [-\eta, \eta]} \left\| \Psi(\rho, \lambda) \right\|_{\infty}^2 \\ &\quad + \int_{-\eta}^{\eta} \sup_{0 \leq \rho \leq 1} \left\| \Psi(\rho, \lambda) f^Y(\lambda) \Psi^H(\rho, \lambda) \right\|_1 \text{Leb}_{\mathbb{T}}(d\lambda). \end{aligned}$$

By Lemma 5.12 the first term of this bound tends to zero as  $\rho \uparrow 1$  for all  $\eta \in (0, \pi)$ . It thus only remain to check that the second term can be made arbitrarily small as  $\eta \downarrow 0$ , which follows if there exists  $\eta > 0$  such that

$$\int_{-\eta}^{\eta} \sup_{0 \leq \rho \leq 1} \left\| \Psi(\rho, \lambda) f^Y(\lambda) \Psi^H(\rho, \lambda) \right\|_1 \text{Leb}_{\mathbb{T}}(d\lambda) < \infty.$$

By definition of  $f^Y$ , setting  $X$  as the ARMA( $p, q$ ) process defined by

$$\hat{X}(d\lambda) = [\Phi(e^{-i\lambda})]^{-1} \Theta(e^{-i\lambda}) \hat{Z}(d\lambda), \quad (5.61)$$

denoting by  $f_X$  the density of  $X$  with respect to  $\text{Leb}_{\mathbb{T}}$ , the previous condition is equivalent to

$$\int_{-\eta}^{\eta} \sup_{0 \leq \rho \leq 1} \left\| \tilde{\Psi}(\rho, \lambda) f_X(\lambda) \tilde{\Psi}^H(\rho, \lambda) \right\|_1 \text{Leb}_{\mathbb{T}}(d\lambda) < \infty,$$

where  $\tilde{\Psi}(\rho, \lambda) := \Psi(\rho, \lambda) (1 - e^{-i\lambda})^{-D} = \left( \Phi_{\Theta, \Phi, D}^{\dagger}(\rho e^{-i\lambda}) - \Phi_{\Theta, \Phi, D}^{\dagger}(e^{-i\lambda}) \right) (1 - e^{-i\lambda})^{-D}$ .

Since  $X \in \mathcal{S}_{\text{FI}_D}(\Omega, \mathcal{F}, \mathbb{P})$  by definition of  $Y$ , Lemma 5.13 gives us that the latter condition holds, which concludes the proof of (5.60) and thus of Assertion (i).

We now prove Assertion (ii). By definition of  $\hat{Y}_t$ , (4.11) and (3.4), we have

$$\begin{aligned} Y_t - \hat{Y}_t &= \int e^{i\lambda t} \left( \text{Id}_{\mathcal{H}_0} - \Phi_{\Theta, \Phi, D}^{\dagger}(e^{-i\lambda}) \right) d\hat{Y} \\ &= \int e^{i\lambda t} \left[ \Theta(e^{-i\lambda}) \right]^{-1} \Phi(e^{-i\lambda}) (1 - e^{-i\lambda})^D d\hat{Y} \\ &= \int e^{i\lambda t} d\hat{Z} = Z_t. \end{aligned}$$

Since  $Z$  is a white noise we have  $Z_t \perp \mathcal{H}_{t-1}^Z$ . To prove Assertion (ii), it thus only remains to show that  $\mathcal{H}_s^Y$  is included in  $\mathcal{H}_s^Z$  for all  $s \in \mathbb{Z}$ . Since we assumed  $\Phi \in \mathcal{P}_p^{\dagger}(\mathcal{H}_0)$ , we have that  $\Phi^{-1}$  is holomorphic in an open domain that includes  $\overline{\mathbb{D}}$  and thus can be written as a power series on the unit circle. It follows that the ARMA process  $X$  defined by (5.61) satisfies  $\mathcal{H}_s^X \subset \mathcal{H}_s^Z$  for all  $s \in \mathbb{Z}$ . To conclude, we now prove that  $\mathcal{H}_s^Y \subset \mathcal{H}_s^X$  for all  $s \in \mathbb{Z}$ . Observe that, for all  $s \in \mathbb{Z}$ ,

$$Y_s = \int e^{i\lambda s} (1 - e^{-i\lambda})^{-D} \hat{X}(d\lambda).$$

Using that  $z \mapsto (1 - z)^{-D}$  is holomorphic on  $\mathbb{C} \setminus [1, \infty)$ , it can be expanded as a power series on  $\mathbb{D}$  and it follows that, for all  $s \in \mathbb{Z}$  and  $\rho \in (0, 1)$ ,

$$Y_s^{(\rho)} = \int e^{i\lambda s} (1 - \rho e^{-i\lambda})^{-D} \hat{X}(d\lambda) \in \mathcal{H}_s^X.$$

Using the same trick as for the proof of Assertion (i), we write, for any  $\eta \in (0, \pi)$  and  $\rho \in (0, 1)$ ,

$$\begin{aligned} \mathbb{E} \left[ \left\| Y_s - Y_s^{(\rho)} \right\|_{\mathcal{H}_0}^2 \right] &\leq \left( \int \|f_X\|_1 d\text{Leb}_{\mathbb{T}} \right) \sup_{\lambda \in \mathbb{T} \setminus [-\eta, \eta]} \left\| (1 - \rho e^{-i\lambda})^{-D} - (1 - e^{-i\lambda})^{-D} \right\|_{\infty}^2 \\ &\quad + \int_{-\eta}^{\eta} \sup_{0 \leq \rho \leq 1} \left\| \Psi(\rho, \lambda) f_X(\lambda) \Psi^H(\rho, \lambda) \right\|_1 \text{Leb}_{\mathbb{T}}(d\lambda), \end{aligned}$$

$$\text{where, here, } \Psi(\rho, \lambda) := (1 - \rho e^{-i\lambda})^{-D} - (1 - e^{-i\lambda})^{-D}.$$

By continuity of  $z \mapsto (1 - z)^{-D}$  on  $\mathbb{C} \setminus [1, \infty)$ , the first term in the upper bound tends to zero for all  $\eta$ , while the second bound can be made arbitrarily small as  $\eta \downarrow 0$  as a consequence of Lemma 5.13. We thus get the claim  $\mathcal{H}_s^Y \subset \mathcal{H}_s^X$  and the proof is concluded.  $\square$

**Proof of Proposition 4.5.** All the assertions follow straightforwardly from Lemma 5.14 and other previous results. Observe indeed that (4.14) immediately follows from the fact that  $F^\dagger(\theta, \lambda) = \text{Id}_{\mathcal{H}_0} - \Phi_{D_\theta, \hat{\Phi}_\theta, \Theta_\theta}^\dagger(e^{-i\lambda})$ . As for the other claimed facts, here are some details.

Moreover, the continuity of  $(D, \Phi, \Theta) \mapsto P_k^\dagger(\Theta, \Phi, D)$  (and thus with **(A-3)**, that of  $\theta \mapsto P_k^\dagger(\theta)$ ), and the bound (5.57) gives us that  $\theta \mapsto \hat{Y}_t(\theta)$  is continuous on  $\Theta$  onto  $\mathcal{H}_{t-1}^Y$ , hence the expectation in the right-hand side of (4.16) is continuous in  $\theta$ , which shows that the inf is attained on a compact subset of  $\Theta$ .

When the best predictor  $\hat{Y}_t^*$  is well defined for one  $t \in \mathbb{Z}$ , by weak stationarity of  $Y$  it must be well defined for all  $t$ . Then, for all  $t \in \mathbb{Z}$  and all  $\theta \in \Theta_Y^*$ , we have  $\hat{Y}_t^* = \hat{Y}_t(\theta)$ ,  $\mathbb{P}$ -a.s.. Of course, since  $\mathbb{Z}$  is countable, we can exchange the  $\mathbb{P}$ -a.s. with the “for all  $t \in \mathbb{Z}$ ”. We can also exchange the  $\mathbb{P}$ -a.s. with the “for all  $\theta \in \Theta_Y^*$ ” as claimed in assertion (vi) of the proposition because the arguments above also give that,  $\mathbb{P}$ -a.s.,  $\theta \mapsto \hat{Y}_t(\theta)$  is continuous on  $\Theta$  onto  $\mathcal{H}_0$ , and consequently, is uniquely defined by its value on a dense countable subset of (the compact set)  $\Theta_Y^*$ .

Finally, in the well-specified case, we apply Theorem 4.4 and notice that, by Assertion (5.56) of Lemma 5.14,  $\hat{Y}_t(\theta)$  is the right-hand side of (4.12) with  $(\theta, \Phi, D)$  replaced by  $(D_\theta, \hat{\Phi}_\theta, \Theta_\theta)$ , which gives (4.19) and, consequently, (4.18).  $\square$

**Proof of Theorem 4.6.** By Lemma 5.14, we have that  $F^\dagger \in \mathcal{C}(\Theta \times \mathbb{T}, \mathcal{L}_b(\mathcal{H}_0))$  with  $F^\dagger$  defined in (4.14). Similarly, using Assertion (iii) in Theorem 4.6, we also have  $F_s^\dagger \in \mathcal{C}(\Theta \times \mathbb{T}, \mathcal{L}_b(\mathcal{H}_0^s))$ , with  $F_s^\dagger(\theta, \lambda) := F^\dagger(\theta, \lambda)|_{\mathcal{H}_0^s}$ . Hence Assertions (i), (ii) and (iii) of Theorem 4.3 hold with  $L = R = F^\dagger$ . Applying this theorem, with the fact that the trace is continuous on  $\mathcal{S}_1(\mathcal{H}_0)$ , we obtain that,  $\mathbb{P}$ -a.s.,

$$\lim_{n \rightarrow \infty} \Lambda_n = \text{Tr} \left( \tilde{Q}_{\nu_X}^{(F^\dagger, F^\dagger)} \right) \quad \text{uniformly in } \mathcal{C}(\Theta, \mathbb{R}).$$

Now, observe that by (4.14) and (4.15), for all  $\theta \in \Theta$ ,

$$\text{Tr} \left( \tilde{Q}_{\nu_X}^{(F^\dagger, F^\dagger)}(\theta) \right) = \mathbb{E} \left[ \left\| Y_0 - \hat{Y}_0(\theta) \right\|_{\mathcal{H}_0}^2 \right].$$

Therefore, with (4.20) and (4.16), we can write that,  $\mathbb{P}$ -a.s.,

$$\limsup_{n \rightarrow \infty} \Lambda_n(\hat{\theta}_n) \leq \limsup_{n \rightarrow \infty} \inf_{\theta \in \Theta} \Lambda_n(\theta) = \mathbb{E}^2(Y, \aleph).$$

Thus,  $\mathbb{P}$ -a.s., all accumulation points  $\theta$  of the  $\Theta$ -valued sequence  $(\hat{\theta}_n)_{n \geq 1}$  satisfy

$$\mathbb{E} \left[ \left\| Y_0 - \hat{Y}_0(\theta) \right\|_{\mathcal{H}_0}^2 \right] \leq \mathbb{E}^2(Y, \aleph), \quad \text{which implies } \theta \in \Theta_Y^*.$$

Since  $\Theta$  is compact, we obtain (4.23).

We now prove (4.24). Let us define, for all  $\theta \in \Theta$ ,

$$\mathbb{E}_{Y, \infty}^2(\theta) = \mathbb{E} \left[ \left\| Y_0 - \left( \sum_{k=1}^{\infty} P_k^\dagger(\theta) Y_{-k} \right) \right\|_{\mathcal{H}_0}^2 \right]. \quad (5.62)$$

By (4.22) and Minkowski's inequality, we have, for all  $(m, \theta) \in \mathcal{H}_0 \times \Theta$  and  $n \geq 1$ ,

$$|\mathbb{E}_{Y, \infty}(\theta) - \mathbb{E}_{X, n}(m, \theta)| \leq |\mathbb{E}_{Y, \infty}(\theta) - \mathbb{E}_{Y, n}(0, \theta)| + \left\| \left( \text{Id}_{\mathcal{H}_0} - \sum_{k=1}^{\infty} P_k^\dagger(\theta) \right) (\mathbb{E}[X_0] - m) \right\|_{\mathcal{H}_0}.$$

We further have, for all  $m \in \mathcal{H}_0$  and  $n \geq 1$ ,

$$\begin{aligned} \sup_{\theta \in \Theta} |\mathbb{E}_{Y, \infty}(\theta) - \mathbb{E}_{Y, n}(0, \theta)| &\leq \sup_{\theta \in \Theta} \left( \mathbb{E} \left[ \left\| \sum_{k=n+1}^{\infty} P_k^\dagger(\theta) Y_{-k} \right\|_{\mathcal{H}_0}^2 \right] \right)^{1/2} \\ &\leq \left( \sum_{k=n+1}^{\infty} \sup_{\theta \in \Theta} \|P_k^\dagger(\theta)\|_{\infty} \right) \left( \mathbb{E} \left[ \|Y_0\|_{\mathcal{H}_0}^2 \right] \right)^{1/2}, \end{aligned}$$

which, by Lemma 5.14, converges to zero as  $n \rightarrow \infty$ .

On the other hand, we have, for all  $m \in \mathcal{H}_0$ ,

$$\sup_{\theta \in \Theta} \left\| \left( \text{Id}_{\mathcal{H}_0} - \sum_{k=1}^{\infty} P_k^\dagger(\theta) \right) (\mathbb{E}[X_0] - m) \right\|_{\mathcal{H}_0} \leq \left( 1 + \sum_{k=1}^{\infty} \sup_{\theta \in \Theta} \|P_k^\dagger(\theta)\|_{\infty} \right) (\mathbb{E}[X_0] - m),$$

which, by Lemma 5.14 again, converges to zero as  $m \rightarrow \mathbb{E}[X_0]$ . Consequently, (4.24) follows if we can show that

$$\lim_{n \rightarrow \infty} \mathbb{E}_{Y, \infty}^2(\hat{\theta}_n) = \mathbb{E}^2(Y, \aleph) \quad \mathbb{P}\text{-a.s.} \quad (5.63)$$

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n X_k = \mathbb{E}[X_0] \quad \mathbb{P}\text{-a.s.} \quad (5.64)$$

Lemma 5.8 gives us (5.64). We now prove (5.63). Observe that, by (5.62) and by continuity of  $\theta \mapsto \sum_k P_k^\dagger(\theta)$  (see Lemma 5.14), we have that  $\theta \mapsto \mathbb{E}_{Y, \infty}^2(\theta)$  is continuous on  $\Theta$  onto  $\mathbb{R}_+$ . Now, since  $\Theta$  is compact, we have that,  $\mathbb{P}$ -a.s.,  $(\ell_n)_{n \geq 1}$  is a bounded sequence in  $\mathbb{R}_+$ , where we set, for all  $n \geq 1$ ,  $\ell_n := \mathbb{E}_{Y, \infty}^2(\hat{\theta}_n)$ . We now show that,  $\mathbb{P}$ -a.s., all accumulation points of this sequence is in fact equal to the right-hand side of (5.63). This follows from the fact that, by compactness of  $\Theta$ , from any increasing sequence  $(n_j)_{j \in \mathbb{N}}$  of positive integers, we can extract a subsequence  $(n'_j)_{j \in \mathbb{N}}$  such that  $\hat{\theta}_{n'_j}$  converges. Furthermore, by (4.23),  $\mathbb{P}$ -a.s., the limit of this sequence must belong to  $\Theta_Y^*$ . With the continuity of  $\theta \mapsto \mathbb{E}_{Y, \infty}^2(\theta)$  previously established, we conclude that,  $\mathbb{P}$ -a.s., all accumulation points of  $(\ell_n)_{n \geq 1}$  is of the form  $\mathbb{E}_{Y, \infty}^2(\theta)$  for some  $\theta \in \Theta_Y^*$ , hence is equal to  $\mathbb{E}^2(Y, \aleph)$  by (4.16) and definition of  $\Theta_Y^*$  in (4.17).

Let us show the last assertion of the theorem. To this end, we suppose that  $\hat{Y}_t^*$  is well defined and show that (4.25) holds. We first observe that, since  $X_{n,k}^c = Y_{n,k}^c$ ,

$$X_{n+1} - \hat{X}_{n+1,n} = Y_{n+1} - \sum_{k=1}^n P_k^\dagger(\hat{\theta}_n) Y_{n+1-k} - \left( \text{Id}_{\mathcal{H}_0} - \sum_{k=1}^n P_k^\dagger(\hat{\theta}_n) \right) \left( \frac{1}{n} \sum_{k=1}^n Y_k \right).$$

Therefore, to obtain (4.25), we only need to show that

$$\limsup_{n \rightarrow \infty} \mathbb{E} \left[ \left\| Y_{n+1} - \sum_{k=1}^n P_k^\dagger(\hat{\theta}_n) Y_{n+1-k} \right\|_{\mathcal{H}_0}^2 \right] \leq \mathbb{E}^2(Y, \aleph), \quad (5.65)$$

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[ \left\| \left( \text{Id}_{\mathcal{H}_0} - \sum_{k=1}^n P_k^\dagger(\hat{\theta}_n) \right) \left( \frac{1}{n} \sum_{k=1}^n Y_k \right) \right\|_{\mathcal{H}_0}^2 \right] = 0. \quad (5.66)$$

Let us start with (5.66). By Lemma 5.14, we have

$$\sup_{n \in \mathbb{N}} \sup_{\theta \in \Theta} \left\| \text{Id}_{\mathcal{H}_0} - \sum_{k=1}^n P_k^\dagger(\theta) \right\|_{\infty} \leq 1 + \sum_{k=1}^{\infty} \sup_{\theta \in \Theta} \|P_k^\dagger(\theta)\|_{\infty} < \infty.$$

Then, we get (5.66) by applying Lemma 5.15.

We now prove (5.65). For all  $n \geq 1$ , the squared norm in the left-hand side's expectation of (5.65) can be written as

$$\inf_{\theta \in \Theta_Y^*} \left\| Y_{n+1} - \sum_{k=1}^n P_k^\dagger(\theta) Y_{n+1-k} + \sum_{k=1}^n \left( P_k^\dagger(\theta) - P_k^\dagger(\hat{\theta}_n) \right) Y_{n+1-k} \right\|_{\mathcal{H}_0}^2.$$

We thus have, for all  $n \geq 1$ ,

$$\mathbb{E} \left[ \left\| Y_{n+1} - \sum_{k=1}^n P_k^\dagger(\hat{\theta}_n) Y_{n+1-k} \right\|_{\mathcal{H}_0}^2 \right] \leq \mathbb{E} [(A_n + B_n)^2], \quad (5.67)$$

where we set

$$A_n := \inf_{\theta \in \Theta_Y^*} \left\| \sum_{k=1}^n \left( P_k^\dagger(\theta) - P_k^\dagger(\hat{\theta}_n) \right) Y_{n+1-k} \right\|_{\mathcal{H}_0},$$

$$B_n := \sup_{\theta \in \Theta_Y^*} \left\| Y_{n+1} - \sum_{k=1}^n P_k^\dagger(\theta) Y_{n+1-k} \right\|_{\mathcal{H}_0}.$$

We are going to show, successively that

$$\lim_{n \rightarrow \infty} \mathbb{E} [A_n^2] = 0, \quad (5.68)$$

$$\lim_{n \rightarrow \infty} \mathbb{E} [B_n^2] = \mathbb{E}^2(Y, \aleph). \quad (5.69)$$

These two facts with (5.67) indeed imply (5.65). First observe that Lemma 5.14 straightforwardly yields

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[ \sup_{\theta \in \Theta} \left\| \sum_{k=n+1}^{\infty} P_k^\dagger(\theta) Y_{n+1-k} \right\|_{\mathcal{H}_0}^2 \right] = 0. \quad (5.70)$$

Thus, to have (5.68) and (5.69), and by stationarity of  $Y$ , we can use

$$A'_n := \inf_{\theta \in \Theta_Y^*} \left\| \sum_{k=1}^{\infty} \left( P_k^\dagger(\theta) - P_k^\dagger(\hat{\theta}_n) \right) Y_{-k} \right\|_{\mathcal{H}_0},$$

$$B' := \sup_{\theta \in \Theta_Y^*} \left\| Y_0 - \sum_{k=1}^{\infty} P_k^\dagger(\theta) Y_{-k} \right\|_{\mathcal{H}_0},$$

and prove instead

$$\lim_{n \rightarrow \infty} \mathbb{E} [A_n'^2] = 0, \quad (5.71)$$

$$\mathbb{E} [B'^2] = \mathbb{E}^2(Y, \aleph). \quad (5.72)$$

To get Relation (5.72), we observe that, with the assumption that the best  $\aleph$ -predictor is well defined, Assertion (vi) in Proposition 4.5 and (4.15) give that,  $\mathbb{P}$ -a.s., for all  $\theta \in \Theta_Y^*$ ,  $\hat{Y}_0(\theta) = \sum_{k=1}^{\infty} P_k^\dagger(\theta) Y_{-k} = \hat{Y}_0^*$ . Hence,  $\mathbb{P}$ -a.s.,  $B' = \|Y_0 - \hat{Y}_0^*\|$ . We thus have (5.72).

We conclude with the proof of (5.71). Note that

$$A'_n \leq 2 \sup_{\theta \in \Theta_Y^*} \left\| \sum_{k=1}^{\infty} P_k^\dagger(\theta) Y_{-k} \right\|_{\mathcal{H}_0}.$$

Note that the  $L^2$ -norm of this upper bound satisfies

$$\left( \mathbb{E} \left[ \sup_{\theta \in \Theta_Y^*} \left\| \sum_{k=1}^{\infty} P_k^\dagger(\theta) Y_{-k} \right\|_{\mathcal{H}_0}^2 \right] \right)^{1/2} \leq \left( \sum_{k=1}^{\infty} \sup_{\theta \in \Theta_Y^*} \|P_k^\dagger(\theta)\|_{\infty} \right) \left( \mathbb{E} [\|Y_0\|_{\mathcal{H}_0}^2] \right)^{1/2},$$

which is finite by Lemma 5.14. Thus, we can apply the dominated convergence theorem, and (5.71) follows from

$$\lim_{n \rightarrow \infty} A'_n = 0 \quad \mathbb{P}\text{-a.s.},$$

which we now prove by contradiction. Suppose that, with positive probability, we can find  $\eta > 0$  and an increasing sequence  $(n_j)_{j \in \mathbb{N}}$  of integers such that  $A'_{n_j} \geq \eta$  for all  $j$ . Then, by (4.23) and since  $\Theta$  is compact, with positive probability, there also exists a subsequence  $(n'_j)_{j \in \mathbb{N}}$  integers such that  $A'_{n'_j} \geq \eta$  for all  $j$  and  $\hat{\theta}_{n'_j}$  converges to some  $\theta \in \Theta_Y^*$  as  $j \rightarrow \infty$ . By Lemma 5.14, this latter fact implies that, for this  $\theta$ ,

$$\lim_{j \rightarrow \infty} \left\| \sum_{k=1}^{\infty} \left( P_k^\dagger(\theta) - P_k^\dagger(\hat{\theta}_{n'_j}) \right) Y_{-k} \right\|_{\mathcal{H}_0} = 0.$$

But since  $\theta \in \Theta_Y^*$ , this contradicts the assumption that yielded  $A'_{n'_j} \geq \eta > 0$  for all  $j$ . This finishes the proof.  $\square$

## A Technical lemmas

We start with two lemmas on complex analysis.

**Lemma A.1.** *For all  $z \in \mathbb{C}$  and  $\lambda \in [-\pi, \pi] \setminus \{0\}$ , we have*

$$(2/\pi)^{2\Re_+(z)} |\lambda|^{2\Re(z)} e^{-\pi|\Im(z)|} \leq \left| (1 - e^{-i\lambda})^z \right|^2 \leq (\pi/2)^{2\Re_-(z)} |\lambda|^{2\Re(z)} e^{\pi|\Im(z)|}, \quad (\text{A.1})$$

where  $\Re(z) = (z + \bar{z})/2$ ,  $\Re_+(z) = \max(\Re(z), 0)$  and  $\Re_-(z) = \max(-\Re(z), 0)$ .

*Proof.* Let  $z \in \mathbb{C}$  and  $\lambda \in (-\pi, \pi] \setminus \{0\}$ . By definition of the principal logarithm, we have

$$\text{for all } y \in \mathbb{C} \setminus \mathbb{R}_-, \quad |y^z|^2 = |\exp(z \ln(y))|^2 = |y|^{2\Re(z)} e^{-2\Im(z)b(y)}, \quad (\text{A.2})$$

where  $b(y)$  denotes the argument in the polar form of  $y$  in  $(-\frac{\pi}{2}, \frac{\pi}{2})$ . It follows that  $e^{-\pi|\Im(z)|} \leq e^{-2\Im(z)b(y)} \leq e^{\pi|\Im(z)|}$ . Applying (A.2) with  $y = 1 - e^{-i\lambda}$ , using that  $\frac{2|\lambda|}{\pi} \leq |2 \sin(\lambda/2)| = |1 - e^{-i\lambda}| \leq |\lambda|$  for all  $\lambda \in (-\pi, \pi)$  and separating the cases where  $\Re(z) \geq 0$  and where  $\Re(z) < 0$ , we get (A.1).  $\square$

**Lemma A.2.** *For all  $z \in \mathbb{C}$  and  $\lambda \in [-\pi/3, \pi/3] \setminus \{0\}$ , we have*

$$\sup_{0 \leq \rho \leq 1} \left| (1 - \rho e^{-i\lambda})^z \right|^2 \leq \left( 2\pi / (3\sqrt{3}) \right)^{2\Re_-(z)} |\lambda|^{-2\Re_-(z)} e^{\pi|\Im(z)|}, \quad (\text{A.3})$$

where  $\Re(z) = (z + \bar{z})/2$ ,  $\Re_+(z) = \max(\Re(z), 0)$  and  $\Re_-(z) = \max(-\Re(z), 0)$ .

*Proof.* Applying (A.2) with  $y = 1 - \rho e^{-i\lambda}$  and using that  $b(y) \in (-\frac{\pi}{2}, \frac{\pi}{2})$ , we get that

$$\sup_{0 \leq \rho \leq 1} \left| (1 - \rho e^{-i\lambda})^z \right|^2 \leq \sup_{0 \leq \rho \leq 1} \left| 1 - \rho e^{-i\lambda} \right|^{2\Re(z)} e^{\pi|\Im(z)|}.$$

Let now  $z \in \mathbb{C}$  and  $\lambda \in [-\pi/3, \pi/3] \setminus \{0\}$ . It is straightforward to show that, in this case,

$$\sin^2(\lambda) = \inf_{0 \leq \rho \leq 1} \left| (1 - \rho e^{-i\lambda}) \right|^2 \leq \sup_{0 \leq \rho \leq 1} \left| (1 - \rho e^{-i\lambda}) \right|^2 = 1.$$

Separating the cases where  $\Re(z) \geq 0$  and where  $\Re(z) < 0$ , and, in the latter case, using that  $|\sin(\lambda)| \geq 3\sqrt{3}|\lambda|/(2\pi)$  for  $|\lambda| \leq \pi/3$ , we easily get (A.3).  $\square$

The next two lemmas involve Banach-space-valued and non-negative series.

**Lemma A.3.** *Let  $E$  be a Banach space and  $(a_n)_{n \in \mathbb{N}} \in E^{\mathbb{N}}$  such that  $\|a_n\|_E \xrightarrow{n \rightarrow \infty} 0$  and the series  $\sum \|a_n - a_{n+1}\|_E$  converges. Then for all  $z_0 \in \mathbb{D} \setminus \{1\}$ , the series  $\sum_{n=0}^{\infty} a_n z_0^n$  converges in  $E$  and the mapping  $z \mapsto \sum_{n=0}^{\infty} a_n z^n$  is uniformly continuous on  $[0, z_0]$ .*

*Proof.* By assumption on  $(a_n)$ ,  $\sum a_n z^n$  is a power series valued in  $E$  with a convergence radius at least equal to 1, and hence is uniformly continuous on any compact subset of  $\mathbb{D}$ . When  $|z_0| = 1$ , the result follows using Abel's transform.  $\square$

**Lemma A.4.** *Let  $(u_k)_{k \in \mathbb{N}}$  be a non-negative non-increasing sequence such that  $\sum_{k \in \mathbb{N}} u_k < \infty$ . Then there exists a non-decreasing sequence  $(s_k)$ , going to  $\infty$  as  $k \rightarrow \infty$  such that  $\sum_{k \in \mathbb{N}} u_k s_k < \infty$ .*

*Proof.* Let  $k_0 = 0$ , and for all  $n \geq 1$ , define by induction

$$k_n = \min \left\{ j > k_{n-1} : \sum_{k=j}^{\infty} u_k \leq 4^{-n} \right\}.$$

Then  $(k_n)$  is a non-decreasing sequence going to  $\infty$  as  $n \rightarrow \infty$ . Define, for all  $n \geq 1$ , and for all  $k_{n-1} \leq k < k_n$ ,  $s_k = 2^n$ . Then  $(s_k)_{k \in \mathbb{N}}$  is a non-decreasing sequence going to  $\infty$  and we have, by definition of  $(s_k)$ , using that  $(u_k)$  is non-negative and then, by definition of  $(k_n)$ ,

$$\sum_{k \in \mathbb{N}} u_k s_k = \sum_{n=0}^{\infty} 2^n \left( \sum_{k_{n-1} \leq k < k_n} u_k \right) \leq \sum_{n=0}^{\infty} 2^n \left( \sum_{k=k_{n-1}}^{\infty} u_k \right) \leq 4 \sum_{n=0}^{\infty} 2^{-n} < \infty.$$

The proof is concluded.  $\square$

## B $L^2(\mathbf{V}, \mathcal{V}, \xi)$ -valued weakly stationary time series

Within this appendix, we set  $\mathcal{H}_0 = L^2(\mathbf{V}, \mathcal{V}, \xi)$  for a  $\sigma$ -finite measured space  $(\mathbf{V}, \mathcal{V}, \xi)$  and we assume that the Hilbert space  $\mathcal{H}_0$  is separable with dimension  $N \in \{1, 2, \dots, \infty\}$ . This will allow us to use a Hilbert basis  $(\phi_i)_{0 \leq i < N}$  of  $\mathcal{H}_0$ .

We first show that we can always find a version of an  $\mathcal{H}_0$ -valued random variable which is jointly measurable on  $\mathbf{V} \times \Omega$ .

**Proposition B.1.** *Let  $(\mathbf{V}, \mathcal{V}, \xi)$  be a  $\sigma$ -finite measured space. Assume that  $\mathcal{H}_0 = L^2(\mathbf{V}, \mathcal{V}, \xi)$  is separable and let  $Y$  be an  $\mathcal{H}_0$ -valued random variable defined on  $(\Omega, \mathcal{F}, \mathbb{P})$ . Then  $Y$  admits a version  $(v, \omega) \mapsto \tilde{Y}(v, \omega)$  jointly measurable on  $(\mathbf{V} \times \Omega, \mathcal{V} \otimes \mathcal{F})$ .*

*Proof.* Let us define for all  $0 \leq n < N$ ,  $\omega \in \Omega$ ,  $v \in \mathbf{V}$  and  $\epsilon > 0$ ,  $S_n^Y(v, \omega) := \sum_{k=0}^n \langle Y(\omega), \phi_k \rangle \phi_k(v)$  and  $N_\epsilon^Y(\omega) := \inf \left\{ n < N : \|S_n^Y(\cdot, \omega) - Y(\omega)\|_{\mathcal{H}_0}^2 \leq \epsilon \right\}$ . Then it is straightforward to show that, for all  $\omega \in \Omega$  and  $\epsilon > 0$ ,  $N_\epsilon^Y(\omega)$  is well defined in  $\mathbb{N}$  and that  $(N_{2^{-n}}^Y(\omega))_n$  is a non-decreasing sequence. We now defined  $\tilde{Y}$  on  $\mathbf{V} \times \Omega$  by  $\tilde{Y}(v, \omega) = \lim_{n \rightarrow \infty} S_{N_{2^{-n}}^Y(\omega)}^Y(v, \omega)$  if the limit exists in  $\mathbb{C}$  and 0 otherwise. It follows that, for all  $\omega \in \Omega$ ,  $S_{N_{2^{-n}}^Y(\omega)}^Y(\cdot, \omega)$  converges to  $Y$  in  $\mathcal{H}_0$  and  $S_{N_{2^{-n}}^Y(\omega)}^Y(v, \omega)$  converges to  $\tilde{Y}(v, \omega)$  for  $\xi$ -a.e.  $v \in \mathbf{V}$  as  $n \rightarrow \infty$  and that  $\tilde{Y}(\cdot, \omega) = Y(\omega)$  (as elements of  $\mathcal{H}_0$ ). The result follows since  $S_n^Y$  is jointly measurable on  $\mathbf{V} \times \Omega$  for all  $n \in \mathbb{N}$  and  $N_\epsilon^Y$  is measurable on  $\Omega$  for all  $\epsilon > 0$ .  $\square$

Hence, an  $\mathcal{H}_0$ -valued random variable  $Y$  can always be assumed to be represented by a  $\mathbf{V} \times \Omega \rightarrow \mathbb{C}$ -measurable function  $\tilde{Y}$ . If, moreover,  $Y \in L^2(\Omega, \mathcal{F}, \mathcal{H}_0, \mathbb{P})$ , then, by Fubini's theorem, we can see  $\tilde{Y}$  as an element of  $L^2(\mathbf{V} \times \Omega, \mathcal{V} \otimes \mathcal{F}, \xi \otimes \mathbb{P})$ , and we can write  $\tilde{Y}(v, \omega) = \sum_{0 \leq k < N} \langle Y(\omega), \phi_k \rangle \phi_k(v)$ , where the convergence holds in  $L^2(\mathbf{V} \times \Omega, \mathcal{V} \otimes \mathcal{F}, \xi \otimes \mathbb{P})$ . As expected, in this case, the covariance operator  $\text{Cov}(Y)$  is an integral operator with kernel  $(v, v') \mapsto \text{Cov}(\tilde{Y}(v, \cdot), \tilde{Y}(v', \cdot))$ . It is tempting to write that  $\text{Var}(\tilde{Y}(v, \cdot))$  is equal to the kernel of  $\text{Cov}(Y)$  on the diagonal  $\{v = v' : (v, v') \in \mathbf{V}^2\}$ . However, because this diagonal set has null  $\xi^{\otimes 2}$ -measure, this "equality" is meaningless. In the following lemma we make this statement rigorous by relying on a decomposition of the form  $\text{Cov}(Y) = KK^H$  for some  $K \in \mathcal{S}_2(\mathcal{H}_0)$ . In particular, this can be used to give a rigorous definition of  $\sigma_W$  in Theorem 3.4 or (3.11).

**Lemma B.2.** *Let  $(\mathbf{V}, \mathcal{V}, \xi)$  be a  $\sigma$ -finite measured space. Assume that  $\mathcal{H}_0 = L^2(\mathbf{V}, \mathcal{V}, \xi)$  is separable and let  $Y$  be an  $\mathcal{H}_0$ -valued random variable defined on  $(\Omega, \mathcal{F}, \mathbb{P})$ . Let  $K \in \mathcal{S}_2(\mathcal{H}_0)$  and denote by  $\mathcal{K}$  its kernel in  $L^2(\mathbf{V}^2, \mathcal{V}^{\otimes 2}, \xi^{\otimes 2})$ . Suppose that  $\text{Cov}(Y) = KK^H$ . Then, we have, for  $\xi$ -a.e.  $v \in \mathbf{V}$ ,*

$$\mathbb{E} \left[ \left| \tilde{Y}(v, \cdot) \right|^2 \right] = \int |\mathcal{K}(v, v')|^2 \xi(\mathrm{d}v') = \|\mathcal{K}(v, \cdot)\|_{\mathcal{H}_0}^2, \quad (\text{B.1})$$

where  $\tilde{Y}$  is a version of  $Y$  in  $L^2(\mathbf{V} \times \Omega, \mathcal{V} \otimes \mathcal{F}, \xi \otimes \mathbb{P})$ .

*Proof.* As explained before the lemma, we have that  $\tilde{Y}_n : (v, \omega) \mapsto \sum_{0 \leq k < n} \langle Y(\omega), \phi_k \rangle_{\mathcal{H}_0} \phi_k(v)$  converges to  $\tilde{Y}$  as  $n \rightarrow N$  in  $L^2(\mathbf{V} \times \Omega, \mathcal{V} \otimes \mathcal{F}, \xi \otimes \mathbb{P})$ . Let us define, for all  $v, v' \in \mathbf{V}$  and  $0 \leq n \leq N$ ,  $\mathcal{K}_n(v, v') = \sum_{0 \leq k < n} \langle \mathcal{K}(\cdot, v'), \phi_k \rangle_{\mathcal{H}_0} \phi_k(v)$ . Then, using that  $\mathcal{K} \in L^2(\mathbf{V}^2, \mathcal{V}^{\otimes 2}, \xi^{\otimes 2})$ , it is easy to show that  $\mathcal{K}_n$  converges to  $\mathcal{K}$  in  $L^2(\mathbf{V}^2, \mathcal{V}^{\otimes 2}, \xi^{\otimes 2})$  as  $n \rightarrow N$ . By the Cauchy-Schwartz inequality, the mappings  $(g, h) \mapsto [v \mapsto \mathbb{E} [g(v, \cdot) \overline{h(v, \cdot)}]]$  and  $(g, h) \mapsto [v \mapsto \langle g(v, \cdot), h(v, \cdot) \rangle_{\mathcal{H}_0}]$  sesquilinear continuous from  $L^2(\mathbf{V} \times \Omega, \mathcal{V} \otimes \mathcal{F}, \xi \otimes \mathbb{P})$  to  $L^1(\mathbf{V}, \mathcal{V}, \xi)$  and from  $L^2(\mathbf{V}^2, \mathcal{V}^{\otimes 2}, \xi^{\otimes 2})$  to  $L^1(\mathbf{V}, \mathcal{V}, \xi)$ , respectively. This, with the two previous convergence results, shows that  $v \mapsto \mathbb{E} \left[ \left| \tilde{Y}_n(v, \cdot) \right|^2 \right]$  and  $v \mapsto \|\mathcal{K}_n(v, \cdot)\|_{\mathcal{H}_0}^2$  both converge in  $L^1(\mathbf{V}, \mathcal{V}, \xi)$ , to  $\mathbb{E} \left[ \left| \tilde{Y}(v, \cdot) \right|^2 \right]$  and  $\|\mathcal{K}(v, \cdot)\|_{\mathcal{H}_0}^2$ , respectively, that is to the left-hand side and right-hand side of (B.1). Hence, to conclude, we only have to show that, for all  $v \in \mathbf{V}$  and  $0 \leq n < N$ ,  $\mathbb{E} \left[ \left| \tilde{Y}_n(v, \cdot) \right|^2 \right] = \|\mathcal{K}_n(v, \cdot)\|_{\mathcal{H}_0}^2$ .

To this end, we write  $\mathbb{E} \left[ \left| \tilde{Y}_n(v, \cdot) \right|^2 \right] = \mathbb{E} \left[ \sum_{0 \leq j, k < n} \langle Y, \phi_j \rangle_{\mathcal{H}_0} \langle \phi_k, Y \rangle_{\mathcal{H}_0} \phi_j(v) \overline{\phi_k(v)} \right] = \sum_{0 \leq j, k < n} \phi_j^H \text{Cov}(Y) \phi_k \phi_j(v) \overline{\phi_k(v)}$ . Using the fact that  $\text{Cov}(Y) = KK^H$  and Fubini's



theorem, we get  $\phi_j^H \text{Cov}(Y)\phi_k = \int \langle \mathcal{K}(\cdot, v''), \phi_j \rangle_{\mathcal{H}_0} \overline{\langle \mathcal{K}(\cdot, v''), \phi_k \rangle_{\mathcal{H}_0}} \xi(dv'')$ . Inserting this in the previous equation and moving the double sum inside the integral with respect to  $\xi(dv'')$ , this double sum becomes a product of two conjugate sums. Namely, we get that

$$\mathbb{E} \left[ \left| \tilde{Y}_n(v, \cdot) \right|^2 \right] = \int \left| \sum_{0 \leq k < n} \langle \mathcal{K}(\cdot, v''), \phi_k \rangle_{\mathcal{H}_0} \phi_k(v) \right|^2 \xi(dv'') = \|\mathcal{K}_n(v, \cdot)\|_{\mathcal{H}_0}^2,$$

which concludes the proof.  $\square$

## References

- [1] I. Berkes, L. Horváth, and G. Rice. On the asymptotic normality of kernel estimators of the long run covariance of functional time series. *Journal of Multivariate Analysis*, 144:150–175, 2016.
- [2] D. Bosq. *Linear processes in function spaces*, volume 149 of *Lecture Notes in Statistics*. Springer-Verlag, New York, 2000. Theory and applications.
- [3] C. Cerovecki and S. Hörmann. On the clt for discrete fourier transforms of functional time series. *Journal of Multivariate Analysis*, 154:282–295, 2017.
- [4] V. Characiejus and A. Račkauskas. The central limit theorem for a sequence of random processes with space-varying long memory. *Lithuanian mathematical journal*, 53(2):149–160, 2013.
- [5] V. Characiejus and A. Račkauskas. Operator self-similar processes and functional central limit theorems. *Stochastic Processes and their Applications*, 124(8):2605–2627, 2014.
- [6] J. B. Conway. *A course in functional analysis*, volume 96 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1990.
- [7] J. B. Conway. *A course in operator theory*, volume 21 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2000.
- [8] J. Diestel and J. J. Uhl. *Vector measures*. American Mathematical Society, Providence, R.I., 1977. With a foreword by B. J. Pettis, Mathematical Surveys, No. 15.
- [9] N. Dinculeanu. *Vector integration and stochastic integration in Banach spaces*, volume 48. John Wiley & Sons, 2011.
- [10] W. Dunsmuir and E. J. Hannan. Vector linear time series models. *Advances in Applied Probability*, 8(2):339–364, 1976.
- [11] A. Durand and F. Roueff. Weakly stationary stochastic processes valued in a separable Hilbert space: Gramian-Cramér representations and applications. working paper or preprint, June 2022.
- [12] M.-C. Düker. Limit theorems for Hilbert space-valued linear processes under long range dependence. *Stochastic Processes and their Applications*, 128(5):1439–1465, May 2018.
- [13] I. Gohberg and J. Leiterer. *Holomorphic operator functions of one variable and applications*, volume 192 of *Operator Theory: Advances and Applications*. Birkhäuser Verlag, Basel, 2009. Methods from complex analysis in several variables.
- [14] E. J. Hannan. The asymptotic theory of linear time-series models. *J. Appl. Probability*, 10:130–145, corrections, *ibid.* 10 (1973), 913, 1973.
- [15] S. Hörmann, L. Kidziński, and M. Hallin. Dynamic functional principal components. *J. R. Stat. Soc. Ser. B. Stat. Methodol.*, 77(2):319–348, 2015.
- [16] S. Hörmann and P. Kokoszka. Weakly dependent functional data. *The annals of statistics*, 38(3):1845–1884, 2010. Language of publication: en.

- [17] L. Horváth and P. Kokoszka. *Inference for Functional Data with Applications*. Springer Series in Statistics. Springer New York, 2012.
- [18] L. Horváth, P. Kokoszka, and 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):103–122, 2013.
- [19] J. Klepsch, C. Klüppelberg, and T. Wei. Prediction of functional ARMA processes with an application to traffic data. *Econometrics and Statistics*, 1:128–149, Jan. 2017.
- [20] P. Kokoszka. Dependent functional data. *ISRN Probability and Statistics*, 2012.
- [21] P. Kokoszka and N. Mohammadi Jouzdani. Frequency domain theory for functional time series: Variance decomposition and an invariance principle. *Bernoulli*, 26(3):2383–2399, 08 2020.
- [22] D. Li, P. M. Robinson, and H. L. Shang. Long-range dependent curve time series. *Journal of the American Statistical Association*, 115(530):957–971, 2020.
- [23] V. M. Panaretos and S. Tavakoli. Cramer-karhunen-loeve representation and harmonic principal component analysis of functional time series. *Stochastic Processes And Their Applications*, 123(7):29. 2779–2807, 2013.
- [24] V. M. Panaretos and S. Tavakoli. Fourier analysis of stationary time series in function space. *Ann. Statist.*, 41(2):568–603, 2013.
- [25] V. Pipiras and M. S. Taqqu. *Long-Range Dependence and Self-Similarity*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2017.
- [26] A. Račkauskas and C. Suquet. Operator fractional brownian motion as limit of polygonal lines processes in hilbert space. *Stochastics and Dynamics*, 11(01):49–70, 2011.
- [27] H. L. Royden. *Real analysis*. Macmillan Publishing Company, New York, third edition, 1988.
- [28] F. Spangenberg. Strictly stationary solutions of ARMA equations in Banach spaces. *Journal of Multivariate Analysis*, 121:127–138, Oct. 2013.
- [29] S. Tavakoli. *Fourier Analysis of Functional Time Series, with Applications to DNA Dynamics*. PhD thesis, Mathaa, Epfl, 2014.
- [30] A. van Delft. A note on quadratic forms of stationary functional time series under mild conditions. *Stochastic Processes and their Applications*, 2019.
- [31] J. Weidmann. *Linear operators in Hilbert spaces*, volume 68 of *Graduate Texts in Mathematics*. Springer-Verlag, New York-Berlin, 1980. Translated from the German by Joseph Szücs.