

Statistica Sinica Preprint No: SS-2024-0316	
Title	Bootstrapping Portmanteau Tests for Functional White Noise under Unknown Dependence
Manuscript ID	SS-2024-0316
URL	http://www.stat.sinica.edu.tw/statistica/
DOI	10.5705/ss.202024.0316
Complete List of Authors	Yu Miao, Muyi Li, Wai Keung Li and Xingbai Xu
Corresponding Authors	Muyi Li
E-mails	limuyi@xmu.edu.cn

Bootstrapping portmanteau tests for functional white noise under unknown dependence

Yu Miao^a, Mui Li^a, Wai Keung Li^b, Xingbai Xu^a

^a *Xiamen University*

^b *The Education University of Hong Kong*

Abstract: We propose portmanteau tests for functional white noise utilizing the sum of squared empirical autocorrelation functions of functional time series. By applying a Hilbert space approach, we establish the limiting properties of the test under the null hypothesis of uncorrelated but not necessarily independent processes. The test is non-pivotal due to unknown dependence within the sequence. To address this issue, we employ the blockwise random weighting bootstrap to obtain critical values and justify its validity. Furthermore, we extend this method for diagnostics of functional autoregressive model and demonstrate its effectiveness through extensive Monte Carlo simulations and a real data application. An accompanying R package is provided to facilitate checks for general functional white noise.

Key words and phrases: Blockwise random weighting bootstrap; Functional time series; Hilbert space; White noise checks; Model diagnostics.

1. Introduction

In recent years, functional data analysis (FDA) has garnered significant attention due to the necessity of analyzing data represented as random curves, along with important theoretical advancements from Euclidean to non-Euclidean spaces. The monographs by

Bosq (2000), Ramsay and Silverman (2005), and Ramsay and Hooker (2009) provided excellent introductions to FDA, while Horváth and Kokoszka (2012) offered an in-depth investigation of dependent functional data structures.

Building on these advances, researchers have extended FDA to functional time series (FTS), where functional observations are collected over time. A key initial task, as in classical time series, is testing for functional white noise to assess serial uncorrelatedness, which is crucial for both modeling and residual diagnostics. Existing methods fall into two main categories: time-domain portmanteau tests based on autocorrelations and frequency-domain tests using functional periodograms.

In the time domain, early contributions by Gabrys and Kokoszka (2007) introduced a portmanteau test by projecting functional observations onto principal components, thereby reducing the infinite dimensional problem to a multivariate one. This approach was later refined by Horváth et al. (2013), who proposed a test based on the L^2 -norms of empirical autocovariance kernels, allowing the lag order to grow with the sample size. Further generalizations included conditional heteroscedasticity adjusted test statistics by (Kokoszka et al., 2017) and novel autocorrelation measures, such as spherical autocorrelation introduced by Yeh et al. (2023). Meanwhile, Mestre et al. (2021) developed diagnostic plots of functional autocorrelation functions and partial autocorrelation functions with confidence bounds, facilitating visual assessment of dependence structures. However, a common limitation of these methods lies in their reliance on strong assumptions, typically requiring the functional sequence to be an independent and identically distributed (IID) process or a martingale difference sequence (MDS) to derive the asymptotic distribution of the test statistics.

In contrast, the spectral domain offers alternative approaches to functional white noise testing. For example, Zhang (2016) constructed a Cramér–von Mises type statistic based on a cumulative distance between the periodogram function and its integral with respect to the frequency. Due to the non-pivotal limiting null distribution, a block bootstrap procedure was used to obtain critical values. Bagchi et al. (2018) proposed a test based on the L^2 -distance between the estimated spectral density operator and its white noise counterpart, showing that the test statistic is asymptotically normal, thus enabling direct computation of critical values without bootstrapping. Characiejus and Rice (2020) introduced a test using kernel lag-window estimators of the spectral density operator for FTS in separable Hilbert spaces, however, the asymptotic results rely on stronger IID assumptions, compared to Zhang (2016) and Bagchi et al. (2018), where the IID assumption was relaxed. A recent survey by Kim et al. (2023) provided a comprehensive overview of functional white noise tests in both time and spectral domains.

Although spectral domain tests offer significant contributions, our paper highlights the advantages of the BLP portmanteau tests in the time domain for functional white noise, due to their practical implementation and accessibility for applied researchers. Previous studies such as those by Lobato et al. (2002), Francq and Raïssi (2007), Mainassara (2011), and Li and Zhang (2022) in finite-dimensional settings, have demonstrated that when data are uncorrelated but not necessarily independent, the asymptotic distribution of BLP portmanteau tests can deviate substantially due to changes in Bartlett’s formula for the autocorrelation covariance matrix. The discrepancy in asymptotic null distributions arises from the potential nonlinear dependence structure of the time series, which can result in severe size distortion if critical values derived under IID or MDS assump-

tions are used. In this paper, we address this issue in the context of functional time series analysis. We propose BLP portmanteau tests and visualization tools based on the empirical autocorrelation functions of stationary functional time series under the weak white noise hypothesis. Our approach accommodates nonlinear dependence within the white noise while avoiding projection onto a fixed-dimensional subspace. Under weak dependence conditions, we demonstrate that our test has a non-pivotal limiting distribution. To approximate this distribution, we introduce a blockwise random weighting bootstrap approach (Shao, 2011; Li and Zhang, 2022), extending its application from finite-dimensional time series to functional time series (FTS). We establish the consistency of the bootstrap approximation, which ensures our test is asymptotically valid for a broad class of processes, including functional GARCH-type processes and other uncorrelated linear or nonlinear processes. Leveraging the above results, we construct upper confidence bounds for empirical functional autocorrelation estimates.

The paper is organized as follows. Section 2 introduces the test statistics and outlines the theoretical framework. Section 3 investigates the asymptotic properties of the test statistic under both the null hypothesis and the global alternative. Section 4.1 presents a blockwise random weighting bootstrap method to approximate the critical values of the portmanteau test and establishes its asymptotic validity. Section 4.2 provides a by-product of confidence bounds for the single-lag functional autocovariance function (fACF). Monte Carlo simulations evaluating the performance of our test method on original observations, as well as diagnostic checks for fitted weak functional AR models, are presented in Sections 5.1 and 5.2, respectively. In Section 6, we illustrate the application of our method through a real data analysis. Section 7 concludes the paper with a summary of key find-

ings and a discussion of limitations with promising directions for future work. Technical proofs and additional numerical results are provided in the Supplementary Materials.

Before proceeding with the formal analysis, we first introduce the notations and definitions used throughout the paper. Let $\mathbb{H} = L^2[0, 1]$ denote the Hilbert space of real-valued square-integrable functions on the unit interval $[0, 1]$. The inner product in $\mathbb{H}$ is defined as $\langle f, g \rangle = \int_0^1 f(\tau)g(\tau)d\tau$, for any $f, g \in \mathbb{H}$, where $\int = \int_{[0,1]}$. The corresponding norm is $\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle}$. All random functions X are defined in a common probability space $(\Omega, \mathcal{A}, P)$. The L^p -norm of X is denoted by $\|\cdot\|_{L^p} = (\mathbb{E} \|X\|^p)^{1/p}$. For any compact operator, denote by $\|\cdot\|_S$ and $\|\cdot\|_{TR}$ the Hilbert-Schmidt norm and the trace norm respectively. For a real matrix B , $[B]_{ij}$ denotes the (i, j) -th element of the matrix B . We use the notation “ $\xrightarrow{d}$ ” for convergence in distribution, “ $\xrightarrow{p}$ ” for convergence in probability, and “ $\stackrel{D}{=}$ ” for identical distribution.

2. Test statistics

Without loss of generality, we consider a sequence of mean-zero, strict stationary FTS $\{X_t(\tau)\}_{t=1}^\infty$ defined on a compact interval $\mathcal{T}$. We assume that a finite stretch of observations, $X_1, \dots, X_T$, is available. Since any compact interval can be normalized to the unit interval $[0, 1]$, we rescale the argument τ such that $\mathcal{T} = [0, 1]$. Consequently, each X_t can be regarded as a random element in the Hilbert space $\mathbb{H} = L^2[0, 1]$. The main objective is to assess whether the observed series $\{X_t\}_{t=1}^T$ behaves as a functional white noise sequence.

To address this problem, we begin by defining the functional autocorrelation function

(fACF) at lag h , denoted by ρ_h , as introduced in Horváth et al. (2016):

$$\rho_h = \frac{\|\gamma_h\|}{\int \gamma_0(\tau, \tau) d\tau}, \quad \|\gamma_h\| = \left\{ \iint \gamma_h^2(\tau_1, \tau_2) d\tau_1 d\tau_2 \right\}^{1/2}, \quad (2.1)$$

$$\gamma_h(\tau_1, \tau_2) = E\{X_t(\tau_1)X_{t-h}(\tau_2)\}, \quad \tau_1, \tau_2 \in [0, 1], \quad (2.2)$$

where $\gamma_h(\tau_1, \tau_2)$ represents the lag- h autocovariance kernel function of the sequence $\{X_t\}$, and induces the corresponding autocovariance operator $\gamma_h(\cdot) : \mathbb{H} \mapsto \mathbb{H}$. Specifically, for each $g \in \mathbb{H}$ and $\tau_1 \in [0, 1]$,

$$\gamma_h g(\tau_1) = E\{\langle X_{t-h}, g \rangle X_t(\tau_1)\} = \int \gamma_h(\tau_1, \tau_2) g(\tau_2) d\tau_2.$$

The testing hypotheses are formally stated as follows:

$$\mathcal{H}_{0,K} : \rho_1 = \rho_2 = \cdots \rho_K = 0 \leftrightarrow \mathcal{H}_{a,K} : \rho_h \neq 0, \quad \exists \quad h \in \{1, \dots, K\}, \quad (2.3)$$

where K represents the maximal lag number. This test aims to identify the autocorrelation in the sequence up to lag K . By setting the number of lags in $\mathcal{H}_{0,K}$ to 1, we naturally obtain the following single-lag test:

$$\mathcal{H}_{0,h} : \rho_h = 0 \leftrightarrow \mathcal{H}_{a,h} : \rho_h \neq 0, \quad \forall \quad h \in \{1, \dots, K\}.$$

Before introducing the test statistic, we first calculate the sample autocovariance

kernel function $\hat{\gamma}_h$ and the sample fACF $\hat{\rho}_h$, as defined in Kokoszka et al. (2017):

$$\hat{\gamma}_h(\tau_1, \tau_2) = \frac{1}{T} \sum_{t=1+h}^T \{X_t(\tau_1)X_{t-h}(\tau_2)\}, \quad \hat{\rho}_h = \frac{\|\hat{\gamma}_h\|}{\int_0^1 \hat{\gamma}_0(\tau, \tau) d\tau}, \quad h \in \{1, \dots, K\}. \quad (2.4)$$

Then, the test statistics for $\mathcal{H}_{0,h}$ (single-lag test) and $\mathcal{H}_{0,K}$ (portmanteau test) based on these sample estimates are defined as follows:

$$Q_{T,h} = T\hat{\rho}_h^2, \quad V_{T,K} = T \sum_{h=1}^K \hat{\rho}_h^2.$$

We emphasize that our null hypothesis in (2.3) solely tests for the absence of autocorrelation up to lag K and does not preclude the possibility of other forms of dependence, such as higher-order nonlinear dependencies. To precisely characterize these potential dependencies, we introduce the following definition.

Definition 1. A sequence $\{X_t\}$ admits the representation

$$X_t = f(\varepsilon_t, \varepsilon_{t-1}, \dots), \quad t = 1, 2, \dots, \quad (2.5)$$

where the ε_i 's are IID elements taking values in a measurable space S , and f is a measurable function $f : S^\infty \rightarrow \mathbb{H}$. We define $\{\varepsilon_i^{(t)}\}$ as an IID sequence in S for each $t \in \mathbb{Z}$, independent of $\{\varepsilon_i\}$, but with the same distribution as $\{\varepsilon_i\}$, and

$$X_t^{(m)} = f(\varepsilon_t, \varepsilon_{t-1}, \dots, \varepsilon_{t-m+1}, \varepsilon_{t-m}^{(t)}, \varepsilon_{t-m-1}^{(t)}, \dots). \quad (2.6)$$

Then $\{X_t\}$ is called **L^p - m -approximable** (Hörmann and Kokoszka, 2010) if

$$\sum_{m=1}^{\infty} \|X_m - X_m^{(m)}\|_{L^p} < \infty. \quad (2.7)$$

Remark 1. Definition 1 provides a quantity measure of dependence in functional observations, connecting naturally with mixing conditions through consideration of σ -algebras separated by a temporal lag m tending to infinity. The key intuition underlying Definition 1 is that innovations ε_i from the distant past exert a negligible influence, permitting their replacement by independent copies. The magnitude of this replacement effect is explicitly captured by (2.7). Furthermore, as discussed in Hörmann and Kokoszka (2010), Definition 1 implies strict stationarity and ergodicity of the sequence X_t . These conditions have been verified for numerous well known stationary models, both linear and nonlinear, under mild parameter constraints. For example, Propositions 2.1–2.3 in Hörmann and Kokoszka (2010) provide sufficient conditions ensuring L^p - m -approximability for functional linear processes, functional bilinear models, and functional ARCH models. Consequently, Definition 1 has become fundamental in theoretical analyses of functional time series, as illustrated in the works of Horváth et al. (2014), Hörmann et al. (2015), Zhang (2016), and Kokoszka and Mohammadi Jouzdani (2020), among others. Zhang (2016), and Kokoszka and Mohammadi Jouzdani (2020), among others.

Definition 2. The pointwise definition of the n th-order **cumulant kernel** (Panaretos and Tavakoli, 2013) of the series $\{X_t\}$ is given by:

$$\text{Cum}\{X_{t_1}(\tau_1), \dots, X_{t_n}(\tau_n)\} = \sum_{p=1}^n \sum_{v=(v_1, \dots, v_p)} (-1)^{p-1} (p-1)! \prod_{l=1}^p \mathbb{E} \left\{ \prod_{j \in v_l} X_{t_j}(\tau_j) \right\}, \quad (2.8)$$

where the sum extends over all unordered partitions of $\{1, \dots, n\}$. A cumulant kernel of order $2k$ gives rise to a corresponding $2k$ th-order **cumulant operator** $\mathcal{R}_{t_1, t_2, \dots, t_{2k-1}}(\cdot) : L^2([0, 1]^k) \rightarrow L^2([0, 1]^k)$, defined as

$$(\mathcal{R}_{t_1, t_2, \dots, t_{2k-1}}g)(\tau_1, \dots, \tau_k) = \int_{[0, 1]^k} \text{Cum}\{X_{t_1}(\tau_1), \dots, X_{t_{2k-1}}(\tau_{2k-1}), X_0(\tau_{2k})\} \\ \times g(\tau_{k+1}, \dots, \tau_{2k}) d\tau_{k+1} \cdots d\tau_{2k}. \quad (2.9)$$

Remark 2. Consider Definition 2 with $k = 2$ as an illustrative example. We define $a_l(\tau_1, \tau_2) = E\{X_0(\tau_1)X_l(\tau_2)\}$. Consequently, the 4th-order cumulant function can be expressed as

$$\text{Cum}\{X_l(\tau_1), X_r(\tau_2), X_p(\tau'_1), X_0(\tau'_2)\} = E[X_l(\tau_1)X_r(\tau_2)X_p(\tau'_1)X_0(\tau'_2)] - a_l(\tau'_2, \tau_1)a_{p-r}(\tau_2, \tau'_1) \\ - a_r(\tau'_2, \tau_2)a_{p-l}(\tau_1, \tau'_1) - a_p(\tau'_2, \tau'_1)a_{r-l}(\tau_1, \tau_2).$$

It is worth noting that when functional observations degenerate to random variables, i.e., $X_t(\tau) = X_t$, the term $\text{Cum}\{X_l(\tau_1), X_r(\tau_2), X_p(\tau'_1), X_0(\tau'_2)\}$ reduces to the scalar 4th-order cumulant (Priestly, 1981).

3. Asymptotic properties

In this section, we explore the asymptotic properties of $V_{T,K}$ under both the null and alternative hypotheses. The asymptotic properties of $Q_{T,h}$ can be readily derived from $V_{T,K}$ with $K = 1$. Before presenting the main results, we first introduce the assumptions.

Assumption 1. The FTS $\{X_t\}$ is L^4 - m -approximable.

Assumption 2. The FTS $\{X_t\}$ satisfies the following conditions

- (i) $\sum_{s=-\infty}^{\infty} \|\gamma_s\|_{TR}^2 < \infty$;
- (ii) $\sum_{s_1, s_2, s_3=-\infty}^{\infty} \|\mathcal{R}_{s_1, s_2, s_3}\|_{TR} < \infty$.

Both assumptions impose specific restrictions on the dependence structure of the series $\{X_t\}$. Assumption 1 ensures the L^2 - m -approximability of the process $\{X_t(\tau_1)X_{t-h}(\tau_2)\}$, which is crucial for our analysis of asymptotic properties of the sample autocovariance kernel functions. Assumption 2 imposes summability conditions on the trace norms of the autocovariance operators and the fourth-order cumulant operators. These conditions are required to guarantee the tightness of the sequence of autocovariance kernels, enabling us to determine the limiting distribution through finite-dimensional projections, see, e.g., Panaretos and Tavakoli (2013) and Zhang (2016). Under these assumptions, we establish the asymptotic theory for our test statistics $V_{T,K}$ and $Q_{T,h}$.

Theorem 1. Under Assumptions 1, 2 and the null hypothesis $\mathcal{H}_{0,K}$,

$$V_{T,K} \xrightarrow{d} \frac{1}{\left[\int_0^1 \gamma_0(\tau, \tau) d\tau\right]^2} \sum_{l=1}^{\infty} \xi_{K,l} \mathcal{N}_l^2, \quad (3.1)$$

where $\{\mathcal{N}_l\}_{l=1}^{\infty}$ are IID $N(0, 1)$ random variables, and $\{\xi_{K,l}\}_{l=1}^{\infty}$ are a decreasing sequence of eigenvalues ($\xi_{K,1} \geq \xi_{K,2} \geq \dots$) of the covariance operator Ψ_K , which is defined as

$$\Psi_K(f)(\tau_1, \tau_2) = \iint \psi_K(\tau_1, \tau_2, \tau'_1, \tau'_2) f(\tau'_1, \tau'_2) d\tau'_1 d\tau'_2, \quad (3.2)$$

for any $f \in [0, 1]^2 \rightarrow \mathbb{R}^K$. The (i, j) -th element of the kernel function $\psi_K : [0, 1]^4 \rightarrow \mathbb{R}^{K \times K}$

is given by

$$[\psi_K(\tau_1, \tau_2, \tau'_1, \tau'_2)]_{ij} = \sum_{s=-\infty}^{\infty} E\{X_i(\tau_1)X_0(\tau_2)X_{s+j}(\tau'_1)X_s(\tau'_2)\}. \quad (3.3)$$

The coefficients $\{\xi_{K,l}\}_{l=1}^{\infty}$ are defined by

$$\Psi_K(\varphi_{K,l})(\tau_1, \tau_2) = \xi_{K,l}\varphi_{K,l}(\tau_1, \tau_2), \quad \text{with } \sum_{l=1}^{\infty} \xi_{K,l} < \infty,$$

where $\{\varphi_{K,l}(\tau_1, \tau_2), l \geq 1\}$ are the corresponding orthonormal basis of eigenfunctions. In short, these coefficients are the eigenvalues of a complex covariance operator defined in (3.2). As direct estimation of them is infeasible even for small values of K . Instead, we propose a bootstrap method (detailed in Section 4.1) to approximate the asymptotic distribution of $V_{T,K}$.

Remark 3. Theorem 1 shows the asymptotic distribution of the portmanteau test statistic $V_{T,K}$, which is a weighted chi-square distribution. This result can be viewed as a generalization of finite dimensional time series, with the weight coefficients $\{\xi_{K,l}\}$ now corresponding to the eigenvalues of an infinite dimensional covariance operator Ψ_K rather than a finite dimensional matrix. The proof of Theorem 3.1 relies fundamentally on the asymptotic distribution of the K dimensional process $(\sqrt{T}\hat{\gamma}_1(\tau_1, \tau_2), \dots, \sqrt{T}\hat{\gamma}_K(\tau_1, \tau_2))^T$. For a complete derivation of this key result, including the necessary technical conditions, we refer readers to Lemma ?? in the Supplementary Materials.

To test the hypothesis $\mathcal{H}_{0,h}$ at a specific lag h , we may directly apply the following proposition derived from Theorem 1.

Proposition 1. *Under the conditions of Theorem 1, we have*

$$\|\sqrt{T}\hat{\gamma}_h(\tau_1, \tau_2) - \Gamma_h(\tau_1, \tau_2)\|^2 = o_p(1), \quad \forall \quad h \in \{1, \dots, K\}, \quad (3.4)$$

where $\Gamma_h(\tau_1, \tau_2)$ is a mean-zero Gaussian process in $L^2([0, 1]^2)$ with the covariance structure specified by

$$\psi_{h,h} = \text{Cov}(\Gamma_h(\tau_1, \tau_2), \Gamma_h(\tau'_1, \tau'_2)) = \sum_{s=-\infty}^{\infty} \text{E}\{X_h(\tau_1)X_0(\tau_2)X_{s+h}(\tau'_1)X_s(\tau'_2)\}. \quad (3.5)$$

Here $\psi_{h,h}$ is the (h, h) -th element of the kernel ψ_K .

Then it follows directly that

$$Q_{T,h} \xrightarrow{d} \frac{\|\Gamma_h\|^2}{\left\{\int_0^1 \gamma_0(\tau, \tau) d\tau\right\}^2} \stackrel{D}{=} \frac{\sum_{l=1}^{\infty} \lambda_{h,l} \mathcal{N}_l^2}{\left\{\int_0^1 \gamma_0(\tau, \tau) d\tau\right\}^2}, \quad (3.6)$$

where $\{\lambda_{h,l}, l \geq 1\}$ and $\{\varphi_{h,l}(\tau_1, \tau_2), l \geq 1\}$ are the eigenvalues and eigenfunctions of the covariance operator induced by the kernel (3.5).

Proposition 1 provides the theoretical basis for constructing the upper confidence bound of fACF at any particular lag h , as detailed in Section 4.2. Since this result is a direct consequence of Theorem 1, we omit its proof here.

Remark 4. When sequence is an MDS, meaning that $\text{E}[X_t(\tau_1)|X_{t'}(\tau_2), \tau_2 \in [0, 1]] = 0$ for $t' < t$ and $\tau_1 \in [0, 1]$, the kernel $\psi_K(\tau_1, \tau_2, \tau'_1, \tau'_2)$ can be simplified as follows:

$$[\psi_K(\tau_1, \tau_2, \tau'_1, \tau'_2)]_{ij} = \text{E}\{X_i(\tau_1)X_0(\tau_2)X_j(\tau'_1)X_0(\tau'_2)\}, \quad \text{for } 1 \leq i, j \leq K.$$

This result is consistent with Theorem 2 in Kokoszka et al. (2017). Furthermore, if the sequence is assumed to be IID, the kernel $\psi_K(\tau_1, \tau_2, \tau'_1, \tau'_2)$ can be further simplified to

$$[\psi_K(\tau_1, \tau_2, \tau'_1, \tau'_2)]_{ij} = \begin{cases} E\{X_0(\tau_1)X_0(\tau'_1)\} E\{X_0(\tau_2)X_0(\tau'_2)\} & \text{for } i = j, \\ 0 & \text{for } i \neq j. \end{cases}$$

In this IID case, $\psi_K(\tau_1, \tau_2, \tau'_1, \tau'_2)$ forms a diagonal matrix.

These cases represent two specific instances of the general white noise hypothesis. Crucially, our proposed test maintains its validity even for uncorrelated processes that satisfy neither the IID nor MDS assumptions.

Next, we present the asymptotic behavior of the test $V_{T,K}$ under $\mathcal{H}_{a,K}$.

Theorem 2. *Assume that the FTS satisfies Assumption 1 and there exists $h \in \{1, \dots, K\}$ such that $\rho_h \neq 0$, then $V_{T,K} \xrightarrow{p} \infty$ as $T \rightarrow \infty$.*

Theorem 2 establishes the asymptotic power properties of our proposed testing method, guaranteeing its ability to detect departures from the null under suitable conditions. The proof of this theorem is directly adapted from Theorem 3 of Kokoszka et al. (2017), with additional details provided in the Supplementary Materials for completeness and clarity.

4. Blockwise random weighting bootstrap

4.1 Bootstrap procedure and its asymptotic validity

Due to the non-pivotal and computational complex of the limiting distribution of $V_{T,K}$, we employ a (blockwise) random weighting bootstrap method to approximate its critical

values. Since $Q_{T,h}$ is a special case of $V_{T,K}$, we restrict our attention to the bootstrap implementation for $V_{T,K}$. The bootstrap procedure consists of the following steps:

Step 1. Choose a block size b_T , such that $1 \leq b_T < T$. Divide the data into L_T blocks, denoted by $B_s = \{(s-1)b_T + 1, \dots, sb_T\}$ for $s = 1, \dots, L_T$, where $L_T = T/b_T$. For simplicity, we assume L_T is an integer.

Step 2. Generate IID random draws δ_s , $s = 1, 2, \dots, L_T$, independent of the data, from a common distribution W , where $E(W) = 0$, $E(W^2) = 1$ and $E(W^4) < \infty$. Define the auxiliary variables $w_t = \delta_s$ if $t \in B_s$, for $t = 1, \dots, T$.

Step 3. For each $h = 1, \dots, K$, compute

$$\hat{\gamma}_h^*(\tau_1, \tau_2) = \frac{1}{T} \sum_{t=1+h}^T w_t \{X_t(\tau_1)X_{t-h}(\tau_2) - \hat{\gamma}_h(\tau_1, \tau_2)\},$$

$$\text{then obtain } \hat{\rho}_h^* = \frac{\|\hat{\gamma}_h^*\|}{\int_0^1 \hat{\gamma}_0(\tau, \tau) d\tau}.$$

Step 4. Compute the bootstrapped test statistic $V_{T,K}^* = T \sum_{h=1}^K (\hat{\rho}_h^*)^2$.

Step 5. Repeat Steps 2-4 for B times and let $V_{T,\alpha}^*$ be the empirical $100(1-\alpha)\%$ quantile of $V_{T,K}^*$. We reject the null hypothesis at the significance level α if $V_{T,K} > V_{T,\alpha}^*$.

Remark 5. The blockwise random weighting bootstrap is actually a variation of the conventional wild bootstrap (Wu, 1986; Liu, 1988; Mammen, 1993). The key difference lies in the treatment of the random weights $\{w_t\}_{t=1}^T$, which are held constant within each block of size b_T while are IID across different blocks. This modification is crucial to capture the dependence between $\{X_t(\tau_1)X_{t-i}(\tau_2)\}$ and $\{X_{t'}(\tau_1)X_{t'-j}(\tau_2)\}$ for $1 \leq i, j \leq$

K when t and t' belong to the same block, as discussed Shao (2011), and Li and Zhang (2022). The essential motivation of this bootstrap method is based on maintaining the inherent dependence structure through blocking while generating bootstrap replicates via weighted autocovariance kernels. Notably, the special cases of IID or MDS processes are accommodated by setting $b_T = 1$, reducing the method to the standard wild bootstrap.

Remark 6. We highlight that our bootstrap method differs from the block bootstrap method employed by Zhang (2016). The key distinction manifests in Step 3, where we implement weighted de-meaned autocovariance structures, which may be compared with resampling blocks of the transformed data $\{\mathcal{Y}_{th}(\tau_1, \tau_2) = X_t(\tau_1)X_{t-h}(\tau_2) + X_t(\tau_2)X_{t-h}(\tau_1) - \hat{\gamma}_h(\tau_1, \tau_2) - \hat{\gamma}_h(\tau_2, \tau_1)\}$ in Zhang (2016). The advantages of our random weighting bootstrap become particularly pronounced when extending the method to model diagnostics (e.g., residual correlation analysis). As demonstrated in Section 5.2, our approach naturally incorporates parameter estimation effects through direct weighting of the estimation equations, yielding both computational simplicity and streamlined theoretical analysis. In contrast, adapting the block bootstrap would require explicit incorporation of estimation effects into $\mathcal{Y}_{th}(\tau_1, \tau_2)$, therefore, the theoretical justification and computation will be more complicated.

Establishing the theoretical validity of bootstrap methods in infinite-dimensional functional spaces introduces unique challenges that do not arise in finite-dimensional contexts. Existing works addressing these challenges include Politis and Romano (1994a,b), who established asymptotic validity for stationary bootstrap procedures applied to estimators in weakly dependent stationary Hilbert spaces. Dehling et al. (2015) proved the consistency of a non-overlapping block bootstrap under specific near-epoch dependence conditions

within Hilbert spaces. More recently, Pilavakis et al. (2020) developed a moving block bootstrap (MBB) approach for testing equality of lag-zero autocovariance operators across multiple independent functional time series. Our bootstrap method differs fundamentally from these existing infinite-dimensional techniques by incorporating random weights, necessitating novel proof techniques. Before presenting the relevant theoretical results, we first introduce assumptions.

Assumption 3. The FTS $\{X_t\}$ satisfies the following conditions

- (i) $\sum_{s=-\infty}^{\infty} |s| \cdot \|\gamma_s\|_{TR} < \infty$;
- (ii) $\sum_{s_1, s_2, s_3=-\infty}^{\infty} |s_i| \cdot \|\mathcal{R}_{s_1, s_2, s_3}\|_{TR} < \infty, \quad i = 1, 2, 3.$

Assumption 4. For each $J = 1, 2, \dots, 7$, and for each $j = 1, 2, \dots, J$, we have

$$\sum_{s_1, \dots, s_J=-\infty}^{\infty} |s_j| \cdot \|\text{Cum}(X_{s_1}, X_{s_2}, \dots, X_{s_J}, X_0)\|_{[0,1]^{J+1}} < \infty. \quad (4.1)$$

Remark 7. Assumption 3 is required to guarantee the tightness of the sequence of bootstrapped autocovariance kernels $(\sqrt{T}\hat{\gamma}_1^*(\tau_1, \tau_2), \dots, \sqrt{T}\hat{\gamma}_K^*(\tau_1, \tau_2))^T$. Assumption 4 specifies summability conditions on joint cumulants, crucial for deriving the second-order moment bound of the conditional variance of the bootstrapped test statistic. Both assumptions were also introduced in Zhang (2016). While direct verification of these assumptions in functional spaces may be challenging, their interpretation becomes more transparent in the univariate setting. In fact, Assumptions 1–4 are natural generalizations of conditions employed by Shao (2011) for univariate time series. Those conditions are derived from the GMC(α) conditions (Wu and Shao, 2004), which characterizes processes with exponential decay of dependence (short memory), including many widely used nonlinear models such

as bilinear, threshold, and various GARCH specifications. Consequently, the assumptions we adopt here can be viewed as functional extensions of the $\text{GMC}(\alpha)$ conditions.

To study the conditional convergence of $V_{T,K}^*$ given the sample, we adopt the concept “in distribution in probability”, following Definition 2 of Li et al. (2003). This concept provides a rigorous framework for analyzing the asymptotic behavior of bootstrap statistics conditioned on the observed data.

Based on this notion, the consistency of the proposed bootstrap procedure is established as follows.

Theorem 3. *Assume that $b_T \rightarrow +\infty$ and $b_T/T \rightarrow 0$ as $T \rightarrow +\infty$. Suppose the FTS $\{X_t\}$ satisfies Assumptions 1, 3 and 4. Then under both the null hypothesis $\mathcal{H}_{0,K}$ and alternative $\mathcal{H}_{a,K}$, we have*

$$d\left\{\mathcal{L}(V_{T,K}^*|\mathbb{X}_T), \mathcal{L}(V_{T,K})\right\} \rightarrow 0,$$

in probability as $T \rightarrow \infty$, where d is any metric that induces weak convergence in $L^2([0, 1]^2)^K$, and $\mathcal{L}(Z|\mathbb{X}_T)$ denotes the conditional distribution of the random element $Z \in L^2([0, 1]^2)^K$ given the sample $\mathbb{X}_T := \{X_t\}_{t=1}^T$.

Remark 8. To establish Theorem 3, it is sufficient to demonstrate that

$$d\left[\mathcal{L}\left\{\left(\sqrt{T}\hat{\gamma}_1^*(\tau_1, \tau_2), \dots, \sqrt{T}\hat{\gamma}_K^*(\tau_1, \tau_2)\right)^T \middle| \mathbb{X}_T\right\}, \mathcal{L}\left\{\left(\Gamma_1(\tau_1, \tau_2), \dots, \Gamma_K(\tau_1, \tau_2)\right)^T\right\}\right] \rightarrow 0,$$

in probability as $T \rightarrow \infty$. The proof of the above property follows the similar steps as Lemma ??, focusing on demonstrating the sequence $\left(\sqrt{T}\hat{\gamma}_1^*(\tau_1, \tau_2), \dots, \sqrt{T}\hat{\gamma}_K^*(\tau_1, \tau_2)\right)^T$.

Readers can refer to the Supplemental Materials for more details.

4.2 Upper confidence bound for single-lag fACF

The bootstrap methodology developed in Section 4.1 can be directly adapted to construct lag-wise upper confidence bounds for the functional autocorrelation function (fACF). This yields an informative graphical diagnostic for assessing serial dependence patterns in functional time series at any specified lag. Following the fACF formulation defined in Section 2, we calculate the lag h sample fACF $\hat{\rho}_h$ as :

$$\hat{\rho}_h = \frac{\|\hat{\gamma}_h\|}{\int \hat{\gamma}_0(\tau, \tau) d\tau} = \frac{1}{\sqrt{T}} \sqrt{Q_{T,h}}.$$

Let $Q_{T,h,\alpha}^*$ denote the empirical $100(1 - \alpha)\%$ quantile of $Q_{T,h}^*$ obtained from the bootstrap procedure in Section 4.1. The upper confidence bound for ρ_h is then given by:

$$\hat{C}_h(1 - \alpha) = \frac{1}{\sqrt{T}} \sqrt{Q_{T,h,\alpha}^*}.$$

For a given α , significant deviations of $\hat{\rho}_h$, $h = 1, \dots, K$, beyond this boundary can be used to detect serial correlation at lag h in the functional time series. In particular, under the null hypothesis that the sequence is weak white noise, the upper confidence bound for ρ_h varies with the lag h , contrasting with the constant upper bound derived under the IID assumption as presented in Mestre et al. (2021).

5. Simulation Study

In this section, we perform Monte Carlo simulations to assess the effectiveness of the proposed test and bootstrap method described in Sections 3 and 4. Although the asymptotic properties of the test have been rigorously analyzed for the observed data, simulations are conducted to demonstrate the finite-sample performance of the test when applied to the residuals, thereby evaluating the adequacy of the model.

5.1 White noise checking on observations

We consider the following data generating processes (DGPs):

- (a) IID-BM: $\{X_t(\tau), \tau \in [0, 1]\}_{t \in \mathbb{Z}}$ is a sequence of IID Brownian motion (BM).
- (b) FGARCH(1,1): $\{X_t(\tau)\}$ follows

$$X_t(\tau) = \sigma_t(\tau)\varepsilon_t(\tau), \quad \sigma_t^2(\tau) = \delta(\tau) + \alpha(X_{t-1}^2)(\tau) + \beta(\sigma_{t-1}^2)(\tau), \quad (5.1)$$

where $\delta = 0.01$ (a constant function), α and β are integral operators defined, for $x \in L^2([0, 1])$ and $\tau \in [0, 1]$, by

$$\begin{aligned} (\alpha x)(\tau_1) &= (\beta x)(\tau_1) := \int \alpha(\tau_1, \tau_2)x(\tau_2)d\tau_2 = \int \beta(\tau_1, \tau_2)x(\tau_2)d\tau_2 \\ &= \int 12\tau_1(1 - \tau_1)\tau_2(1 - \tau_2)x(\tau_2)d\tau_2, \end{aligned}$$

where $\varepsilon_t(\tau) = \frac{\sqrt{\ln(2)}}{2^{200\tau}}B_t(\frac{2^{400\tau}}{\ln 2})$, $\tau \in [0, 1]$, and $\{B_t(\tau), \tau \in [0, 1]\}_{t \in \mathbb{Z}}$ is an IID-BM process. This particular process setting is also used in Cerovecki et al. (2019) to mimic high-frequency intraday returns.

- (c) Fbilinear(1, U): $X_{t+1}(\tau) = \iint \phi_l(\tau, \tau_1, \tau_2) X_t(\tau_1) \varepsilon_t(\tau_2) d\tau_1 d\tau_2 + \varepsilon_{t+1}(\tau)$, $\tau \in [0, 1]$, where $\{\varepsilon_t(\tau)\}$ follows the IID-BM process. $\phi_l(\tau, \tau_1, \tau_2) = l \times \tau(1 - \tau)\tau_1(1 - \tau_1)\tau_2(1 - \tau_2)$, the constant l is determined such that $\|\phi_l\|^4 = U \in (0, 1)$, see Hörmann and Kokoszka (2010). Here we set $U = 0.3$ to ensure the approximability of the process.
- (d) FAR(1, S)-BM: $X_t(\tau_1) = \int \phi_c(\tau_1, \tau_2) X_{t-1}(\tau_2) d\tau_2 + \varepsilon_t(\tau_1)$, $\tau_1 \in [0, 1]$, where $\{\varepsilon_t(\tau_1)\}$ follows the IID-BM process. $\phi_c(\tau_1, \tau_2) = c \exp\{-(\tau_1^2 + \tau_2^2)/2\}$ is the Gaussian kernel, the constant c is chosen to satisfy $\|\phi_c\| = S \in (0, 1)$ and S represents the departure from the null hypothesis.

The first three DGPs: (a) the functional IID sequence (strong white noise), (b) the functional MDS (semi-strong white noise), and (c) the functional uncorrelated process (weak white noise, WWN) are used to assess the size of the test. DGP (d), which represents a functional autoregressive (FAR) process, is employed to evaluate the power of the test. For each DGP, simulations are performed with a sample size of T , using an equidistant grid of $J = 200$ points that span the unit interval. Integral calculations are performed using an equidistant grid of 50 points in the interval $[0, 1]$. Additionally, for DGPs (b), (c), and (d), a burn-in sample of length 50 is generated to reduce the effects of initial values.

The sample paths from DGP (a) to (d), with a sample size of $T = 100$, are illustrated in the upper panels of Figure 1. The corresponding autocorrelograms for each process are displayed in the lower panels. The blue dashed lines in the autocorrelograms denote the 95% upper functional WWN confidence bounds for the process, as obtained in Section 4.2. It is evident that the first three functional processes are serially uncorrelated, while the last one exhibits significant autocorrelations at the first two lags.

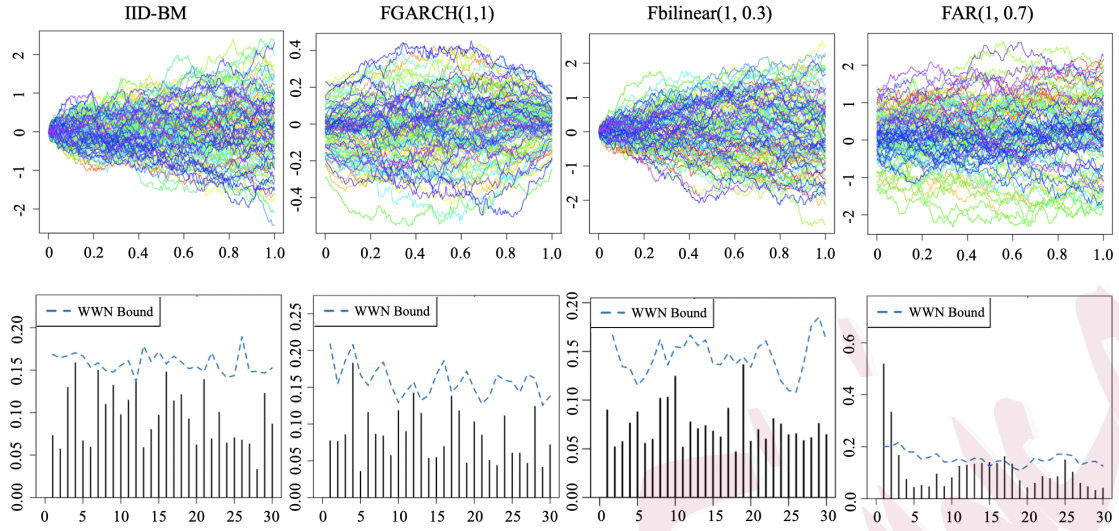


Figure 1. Upper panels: 100 sample paths of DGPs (a) to (d). Lower panels: sample autocorrelograms with 95% upper WWN confidence bounds.

We denote our (blockwise) random weighting bootstrap test as (B)RWB. The random weights w_t are generated from two types of distributions: (i) $N(0, 1)$; (ii) a Bernoulli distribution defined as (see Patilea et al., 2016):

$$w_t = \begin{cases} \frac{1 - \sqrt{5}}{2} & \text{w.p. } \frac{\sqrt{5} + 1}{2\sqrt{5}}, \\ \frac{1 + \sqrt{5}}{2} & \text{w.p. } \frac{\sqrt{5} - 1}{2\sqrt{5}}. \end{cases} \quad (5.2)$$

We compare it with the tests proposed by Gabrys and Kokoszka (2007) (thereafter GK) and Kokoszka et al. (2017) (thereafter KRS). The GK test is constructed based on functional principal components (FPCs), and its asymptotic distribution is obtained as a chi-squared distribution under the IID null hypothesis. The test statistic is defined as

$$G_{T,K,p} = T \sum_{h=1}^K \sum_{i,j=1}^p r_{f,h}(i, j) r_{b,h}(i, j),$$

where $r_{f,h}(i, j)$ and $r_{b,h}(i, j)$ represent the (i, j) -th entries of $\mathbf{C}_0^{-1}\mathbf{C}_h$ and $\mathbf{C}_h\mathbf{C}_0^{-1}$, respectively. Here, $\mathbf{C}_0$ and $\mathbf{C}_h$ are the sample covariance and lag- h autocovariance matrices constructed from empirical FPCs, and p denotes the number of FPCs. Usually, the value of p is chosen such that the cumulative variation explained by the first p principal components exceeds 90%.

The KRS test is constructed based on the cumulative sum of squared L^2 -norm of the autocovariance kernels of the first K lags, which is defined as

$$R_{T,K} = T \sum_{h=1}^K \|\hat{\gamma}_h\|^2.$$

The asymptotic distribution of $R_{T,K}$ is derived under moment conditions that are typically satisfied by functional GARCH-type models, which makes this approach robust for testing serial correlation in such sequences. Additionally, they employ the Welch-Satterthwaite method to approximate the limiting distribution of the test statistic by utilizing its first- and second-order moments.

We first assess the empirical sizes of various tests under the null hypothesis. The parameter K is set to values of 1, 3, 5, 8 and 10. Empirical rejection rates are calculated based on 1,000 independent replications, with each replication involving $B = 500$ bootstrapped iterations. For DGPs (a) and (b), we set $b_T = 1$, while for DGP (c), the blockwise random weighting bootstrap (BRWB) method is used. Although we have provided theoretical conditions for the selection of block size in Section 3, finding a universally applicable method in practice is challenging, and thus a data-driven approach is often recommended; see Zhang (2016) for a minimum volatility method. It is worth noting that our bootstrap

weights are applied to the sequence of centralized covariances, which exhibits a rapidly diminishing dependence. As a result, the block size can generally be chosen relatively small. Our simulation results demonstrate that the selection of block size has minimal impact on the outcomes, indicating that the proposed test is highly robust to variations in block size. To illustrate this robustness, we compare the results using different block sizes. Specifically, we use $b_T = 3$ and 6 for $T = 200$, and $b_T = 4$ and 8 for $T = 800$. The empirical results of the three different tests across various DGPs are presented in Table 1, corresponding to the nominal significance levels of 10%, 5%, and 1%, respectively.

The findings from Table 1 are summarized as follows:

(i) All three tests perform well under DGP (a) that the process is an IID sequence, although the GK test tends to be conservative for large K and small sample size, which is an intrinsic characteristic of finite-dimensional portmanteau tests. The KRS test exhibits some upward size distortion, but this issue diminishes as the sample size increases.

(ii) The GK test has a severe over-rejection problem under DGPs (b) and (c), suggesting that the critical values taken from the functional IID assumption are not valid for functional MDS or functional WWN.

(iii) The KRS test performs reasonably well for DGP (b), although it seems slightly conservative, especially for large K , which is consistent with the findings in Kokoszka et al. (2017). At the same time, it has an over-rejection problem for DGP (c), and the size distortion becomes more pronounced as the sample size increases. This finding indicates the invalidity of the KRS test to uncorrelated yet non-martingale difference sequences and further illustrates the significance of our proposed testing method.

(iv) Although the RWB/BRWB tests generally perform well across all DGPs, we

		$K = 1$			$K = 3$			$K = 5$			$K = 8$			$K = 10$			
Test	W	10%	5%	1%	10%	5%	1%	10%	5%	1%	10%	5%	1%	10%	5%	1%	
(a) IID-BM																	
$T = 200$	GK		9.6	5.0	1.4	8.4	5.1	1.4	10.1	5.3	1.5	8.3	4.2	0.7	8.3	4.2	1.0
	KRS		10.7	6.4	1.7	9.9	6.0	1.9	11.6	6.6	2.1	12.1	5.7	1.5	11.0	6.0	1.6
	RWB	$N(0, 1)$	9.8	4.5	0.7	9.6	4.4	0.7	10.0	4.5	0.8	7.7	3.7	0.5	9.9	4.9	0.9
		Bernoulli	9.6	4.4	0.7	9.3	4.5	0.6	10.3	4.4	1.2	8.5	4.5	0.6	11.7	5.5	0.8
$T = 800$	GK		9.8	5.4	1.4	10.8	5.2	2.1	10.0	5.9	1.2	9.4	5.7	1.6	9.5	5.3	1.8
	KRS		11.6	5.6	1.3	8.0	4.8	1.8	8.6	4.8	1.6	11.0	5.8	1.3	9.7	5.5	1.6
	RWB	$N(0, 1)$	10.6	4.7	1.2	10.4	5.0	0.9	11.0	5.4	1.4	10.2	5.8	1.0	10.1	5.1	0.4
		Bernoulli	11.2	5.1	0.9	11.3	4.9	1.2	11.2	5.1	1.2	10.6	5.6	1.1	10.5	5.6	0.6
(b) FGARCH(1,1)																	
$T = 200$	GK		28.3	16.9	10.5	26.1	18.5	9.2	36.7	27.0	13.8	32.6	22.6	11.2	31.3	20.1	10.3
	KRS		9.2	4.6	0.5	10.2	4.8	1.1	9.6	3.9	0.5	7.7	3.5	0.6	7.0	3.3	0.7
	RWB	$N(0, 1)$	9.6	4.2	0.7	9.2	3.9	0.8	7.4	3.4	0.6	8.7	3.2	0.5	7.4	3.9	0.4
		Bernoulli	9.4	4.4	0.7	8.9	4.5	0.9	7.6	3.6	0.6	9.4	4.0	0.5	8.7	4.3	0.5
$T = 800$	GK		46.2	39.9	22.6	49.5	40.3	25.8	65.7	55.4	38.2	56.3	45.8	30.9	54.2	47.1	33.1
	KRS		8.9	4.2	0.5	8.9	4.5	1.0	8.9	4.5	0.6	8.1	4.5	0.8	7.6	4.1	0.7
	RWB	$N(0, 1)$	10.6	5.0	1.2	10.7	5.2	1.1	8.4	4.0	0.7	8.7	3.3	0.8	7.6	3.9	0.5
		Bernoulli	11.0	4.9	0.8	11.2	5.2	1.5	9.1	4.4	0.7	9.0	3.6	0.9	8.9	4.4	0.5
(c) Fbilinear(1,0.3)																	
$T = 200$	GK		29.4	17.0	5.9	27.3	19.5	6.2	31.1	21.0	7.1	27.9	16.1	6.9	30.5	18.3	5.7
	KRS		12.3	7.0	2.5	13.2	6.9	1.8	13.8	6.8	1.8	11.9	6.9	1.35	11.4	6.6	1.4
	BRWB	$N(0, 1)$	10.7	4.8	1.2	11.5	5.5	0.5	9.2	4.4	0.9	9.7	4.2	0.8	8.8	4.2	0.4
	$(b_T = 3)$	Bernoulli	11.4	5.3	1.1	10.4	5.4	1.3	9.3	4.9	0.9	10.5	5.3	0.8	7.4	4.5	0.9
	BRWB	$N(0, 1)$	11.2	5.9	0.8	11.4	5.3	0.4	8.9	3.7	0.8	9.1	3.5	0.3	7.3	2.7	0.2
	$(b_T = 6)$	Bernoulli	11.2	5.7	1.2	10.2	4.8	0.7	9.								

still observe some upward size distortion for certain K values in DGP (c). However, the proposed test outperforms the two alternative methods in most cases. In particular, empirical sizes are close to nominal ones, even for large values of K , suggesting that they can effectively overcome the inherent shortcomings associated with portmanteau tests.

(v) The proposed RWB/BRWB tests are robust to both the choice of block size and the distribution of the random weighting variable, demonstrating insensitivity to user-specified parameters.

Next, we use DGP (d) FAR(1, S)-BM process to evaluate the empirical power of the (B)RWB test and the other comparative methods. Here S represents the Hilbert-Schmidt norm of the autoregressive operator, which quantifies the degree of departure from the null hypothesis ($S = 0$). To evaluate powers, S is set to 0.2, 0.4, 0.6, and 0.8. For simplicity, we take $K = 6$ in $V_{T,K}$, $G_{T,K,p}$ and $R_{T,K}$, considering the sample sizes of $T = 100, 300, 500$. The block size in our method is fixed at $b_T = 3$ across all sample sizes, and the random weights are generated as IID standard normal random variables.

The empirical rejection rates of the three test statistics based on 1000 replications are shown in Figure 2. The rejection rates across the three tests for FAR(1, S)-BM are almost identical, although the proposed test exhibits a notable power advantage as the sample size increases. As expected, the empirical power of all tests increases with the norm S . Furthermore, the empirical power converges more rapidly to 1 as the sample size T increases, demonstrating the effectiveness of the tests in detecting deviations from the null hypothesis in larger samples.

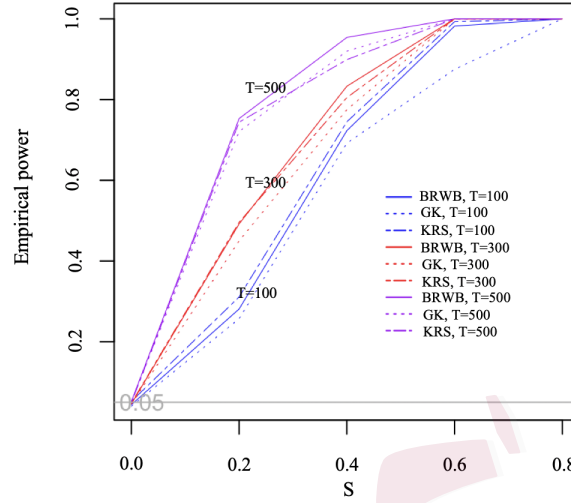


Figure 2. Empirical powers of the three tests for DGP (d) FAR(1, S)-BM. The nominal significance level is $\alpha = 0.05$.

5.2 Diagnostic checking for FAR models

In this subsection, we illustrate the practical use of our proposed test for diagnostic checking of functional autoregressive (FAR) models, specifically focusing on the FAR(1) model. Although conceptually similar to the univariate AR(1) model, the FAR(1) model provides greater flexibility by allowing for varying degrees of nonlinearity in its autoregressive operators. This adaptability has made it widely applicable, as shown in studies like Kargin and Onatski (2008) and Chen et al. (2021), among others.

The FAR(1) model is defined as follows:

$$X_t(\tau_1) = \int \phi_c(\tau_1, \tau_2) X_{t-1}(\tau_2) d\tau_2 + \varepsilon_t(\tau_1), \quad \tau_1 \in [0, 1], \quad (5.3)$$

where $\phi_c(\tau_1, \tau_2)$ represents a kernel function and the innovations $\{\varepsilon_t\}$ constitute a mean-zero $\mathbb{H}$ -valued white noise sequence.

We aim to test the following hypotheses:

$\mathcal{H}_0 : \{X_t(\tau)\}$ satisfies an FAR(1) representation,

against the alternative

$\mathcal{H}_A : \{X_t(\tau)\}$ does not admit an FAR(1) representation or admits an

FAR(p) representation with $p > 1$.

To evaluate model adequacy, the above hypothesis test can be transformed into a white noise test for the error sequence. To achieve this, we first perform the estimation. The kernel function $\phi_c(\tau_1, \tau_2)$ can be expressed as a double expansion in terms of K_0 basis functions η_k for $k = 1, \dots, K_0$ as follows:

$$\phi_c(\tau_1, \tau_2) = \sum_{k,l=1}^{K_0} b_{kl} \eta_k(\tau_1) \eta_l(\tau_2) = \boldsymbol{\eta}(\tau_1)^T \mathbf{B} \boldsymbol{\eta}(\tau_2),$$

where $\boldsymbol{\eta} = (\eta_1, \dots, \eta_{K_0})^T$, and $\mathbf{B}$ is a $K_0 \times K_0$ matrix of coefficients b_{kl} . Consequently, the coefficient matrix $\mathbf{B}$ can be estimated by minimizing the following objective function:

$$SSE(\mathbf{B}) = \int \sum_{t=1}^T \left\{ X_t(\tau_1) - \int \boldsymbol{\eta}^T(\tau_1) \mathbf{B} \boldsymbol{\eta}(\tau_2) X_{t-1}(\tau_2) d\tau_2 \right\}^2 d\tau_1. \quad (5.4)$$

We then compute the test statistic based on the residuals $\{\hat{\varepsilon}_t\}$ as follows:

$$V_{T,K,\hat{\varepsilon}} = T \sum_{h=1}^K \hat{\rho}_{h,\hat{\varepsilon}}^2, \quad \text{where} \quad \hat{\rho}_{h,\hat{\varepsilon}} = \frac{\|\hat{\gamma}_{h,\hat{\varepsilon}}\|}{\int \hat{\gamma}_{0,\hat{\varepsilon}}(\tau, \tau) d\tau},$$

and $\hat{\gamma}_{h,\hat{\varepsilon}}(\tau_1, \tau_2) = \frac{1}{T} \sum_{t=1+h}^T \{\hat{\varepsilon}_t(\tau_1) \hat{\varepsilon}_{t-h}(\tau_2)\}$.

Since the estimation effect is involved in the test statistic and is not asymptotically negligible, it must be adequately accounted for. To address this issue, we introduce a new blockwise random weighting bootstrap method, which differs slightly from the approach discussed in Section 4.1, to obtain the estimated critical values for the test. The steps of the proposed bootstrap procedure are as follows:

Step 1'. The same as in Step 1.

Step 2'. Generate IID random draws $\delta_s, s = 1, 2, \dots, L_T$, independent of the data, from a common distribution V , where $E(V) = 1$, $E(V^2) = 1$ and $E(V^4) < \infty$. Define the auxiliary variables $v_t = \delta_s$ if $t \in B_s$, for $t = 1, \dots, T$. Then, calculate the bootstrapped estimation $\phi_c^*(\tau_1, \tau_2)$ by minimizing the weighted least squares objective function:

$$\hat{\mathbf{B}}^* := \arg \min_{\mathbf{B}} \int \sum_{t=1}^T v_t \left\{ X_t(\tau_1) - \int \boldsymbol{\eta}^T(\tau_1) \mathbf{B} \boldsymbol{\eta}(\tau_2) X_{t-1}(\tau_2) d\tau_2 \right\}^2 d\tau_1.$$

Step 3'. Let $\hat{\varepsilon}_t^*(\tau_1) = X_t(\tau_1) - \int \boldsymbol{\eta}^T(\tau_1) \hat{\mathbf{B}}^* \boldsymbol{\eta}(\tau_2) X_{t-1}(\tau_2) d\tau_2$. Then define

$$\hat{\gamma}_{h,\hat{\varepsilon}}^*(\tau_1, \tau_2) := \hat{\gamma}_{h,\hat{\varepsilon}^*}(\tau_1, \tau_2) = \frac{1}{T} \sum_{t=1+h}^T v_t \left\{ \hat{\varepsilon}_t^*(\tau_1) \hat{\varepsilon}_{t-h}^*(\tau_2) - \hat{\gamma}_{h,\hat{\varepsilon}}(\tau_1, \tau_2) \right\},$$

$$\text{for } h = 1, \dots, K. \text{ Calculate } \hat{\rho}_{h,\hat{\varepsilon}}^* = \frac{\|\hat{\gamma}_{h,\hat{\varepsilon}}^*\|}{\int \hat{\gamma}_{0,\hat{\varepsilon}}(\tau, \tau) d\tau}.$$

Step 4'. Compute the bootstrapped test statistic $V_{T,K,\hat{\varepsilon}}^* = T \sum_{h=1}^K (\hat{\rho}_{h,\hat{\varepsilon}}^*)^2$.

Step 5'. Repeat Steps 2'-4' for B times and denote the empirical $100(1 - \alpha)\%$ sample quantile of $V_{T,K,\hat{\varepsilon}}^*$ as $V_{T,\hat{\varepsilon},\alpha}^*$. Reject the null hypothesis at the significance level α

if $V_{T,K,\hat{\varepsilon}} > V_{T,\hat{\varepsilon},\alpha}^*$.

To illustrate the empirical sizes and powers of our test procedure on model checking, we consider the following DGPs (e) and (f):

- (e) FAR(1, S)-Fbilinear(1, U): $X_t(\tau_1) = \int \phi_c(\tau_1, \tau_2)X_{t-1}(\tau_2)d\tau_2 + \varepsilon_t(\tau_1)$, $\tau_1 \in [0, 1]$, where $\{\varepsilon_t(\tau_1)\}$ follows Fbilinear(1, U). The parameter c is chosen so that $\|\phi_c\| = S \in (0, 1)$.
- (f) FAR(2, S_1, S_2)-Fbilinear(1, U): $X_t(\tau_1) = \int \phi_c(\tau_1, \tau_2)X_{t-1}(\tau_2)d\tau_2 + \int \phi_{c'}(\tau_1, \tau_2)X_{t-2}(\tau_2)d\tau_2 + \varepsilon_t(\tau_1)$, $\tau_1 \in [0, 1]$, where $\{\varepsilon_t(\tau_1)\}$ follows Fbilinear(1, U). The parameters c and c' are chosen so that $\|\phi_c\| = S_1$, $\|\phi_{c'}\| = S_2$, and $S_1 + S_2 < 1$ guarantee the stationarity of the process (Kokoszka and Reimherr, 2013).

We generate 1000 replications from DGPs (e) and (f) with sample sizes 200 and 800, respectively, and fit the data by the FAR(1) model specified in Equation (5.3). Following Step 2', we consider two types of distributions: the standard Exponential distribution and a Bernoulli distribution defined as follows (see Zhu (2016), Li and Zhang (2022)):

$$v_t = \begin{cases} \frac{3 - \sqrt{5}}{2} & \text{w.p. } \frac{\sqrt{5} + 1}{2\sqrt{5}}, \\ \frac{3 + \sqrt{5}}{2} & \text{w.p. } \frac{\sqrt{5} - 1}{2\sqrt{5}}. \end{cases} \quad (5.5)$$

For each replication, we perform $B = 500$ bootstrap iterations with block sizes $b_T = 3, 6$ for $T = 200$ and $b_T = 4, 8$ for $T = 800$, to compute empirical critical values. The rejection rates are summarized in Table 2.

Table 2 highlights several key findings:

- (i) The BRWB test demonstrates satisfactory empirical sizes, even with smaller sample sizes. However, the test tends to be undersized when K increases, a limitation that can

Table 2. Empirical sizes and powers of BRWB on model checking of DGP (e) FAR(1,0.3)-Fbilinear(1,0.3) and (f) FAR(2,0.3,0.5)-Fbilinear(1,0.3).

			<i>K</i> = 1			<i>K</i> = 3			<i>K</i> = 5			<i>K</i> = 8			<i>K</i> = 10		
Test	<i>V</i>		10%	5%	1%	10%	5%	1%	10%	5%	1%	10%	5%	1%	10%	5%	1%
(e) FAR(1,0.3)-Fbilinear(1,0.3)																	
<i>T</i> = 200	BRWB	Std.Exp.	10.6	5.3	0.8	10.2	4.4	0.9	9.0	2.8	0.2	6.2	3.9	0.3	6.3	4.0	0.3
	(<i>b_T</i> = 3)	Bernoulli	11.8	5.9	1.3	11.8	5.6	1.1	11.6	6.1	1.0	8.7	3.8	0.8	8.1	4.1	0.7
	BRWB	Std.Exp.	10.9	4.5	0.7	10.3	4.5	0.7	8.4	2.2	0.3	7.9	3.3	0.4	7.7	3.0	0.2
	(<i>b_T</i> = 6)	Bernoulli	12.7	7.5	1.4	11.3	6.2	1.2	11.2	6.5	0.9	8.9	4.0	0.8	8.7	4.4	0.5
<i>T</i> = 800	BRWB	Std.Exp.	9.6	5.1	0.9	8.3	4.3	0.6	9.7	5.0	0.7	8.0	3.7	0.4	8.0	4.3	0.4
	(<i>b_T</i> = 4)	Bernoulli	11.0	6.0	1.5	11.0	5.8	1.4	11.2	6.4	1.7	9.1	4.7	1.1	10.7	5.2	1.1
	BRWB	Std.Exp.	9.7	4.3	0.6	8.5	4.1	0.4	9.8	4.4	0.8	6.9	3.9	0.4	7.6	3.6	0.3
	(<i>b_T</i> = 8)	Bernoulli	11.2	5.2	1.5	11.4	5.0	1.7	11.2	6.2	1.5	9.0	4.8	1.2	11.7	4.8	1.0
(f) FAR(2,0.3,0.5)-Fbilinear(1,0.3)																	
<i>T</i> = 200	BRWB	Std.Exp.	99.9	98.3	85.9	99.9	99.6	92.6	99.8	97.0	79.4	97.9	92.9	56.4	95.7	85.5	44.0
	(<i>b_T</i> = 3)	Bernoulli	99.8	98.9	91.6	99.9	99.8	97.1	99.9	97.2	90.4	99.0	95.2	83.9	97.9	92.5	75.9
	BRWB	Std.Exp.	99.7	97.8	81.2	99.9	99.1	86.7	99.5	96.4	68.7	96.8	86.4	36.2	94.1	76.1	31.8
	(<i>b_T</i> = 6)	Bernoulli	99.9	98.9	90.9	99.9	99.7	96.3	99.9	97.3	89.9	98.8	94.7	78.7	97.7	91.3	68.7
<i>T</i> = 800	BRWB	Std.Exp.	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	(<i>b_T</i> = 4)	Bernoulli	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	BRWB	Std.Exp.	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	(<i>b_T</i> = 8)	Bernoulli	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

be effectively mitigated by increasing the sample size. In particular, its size performance remains consistent across different block sizes and random weight distributions.

(ii) A comparison between bootstrapped tests using standard exponential random weights and those using Bernoulli random weights shows that the former performs better in this setting. In contrast, the latter tends to produce slightly oversized results.

(iii) The BRWB test exhibits reasonable power. Although its empirical power may decrease as *K* increases - an inherent characteristic of portmanteau tests - it improves with larger sample sizes, achieving values close to 1 when *T* = 800. Furthermore, the power performance is robust to variations in random weights and block sizes.

In general, these simulation results underscore the effectiveness of the proposed test in evaluating the adequacy of FAR models.

6. Real data analysis

In this section, we apply the proposed testing procedure to analyze electricity price data, which plays a pivotal role in shaping the dynamics of the energy market, economic stability, and the formulation of government policies. Given the substantial influence of electricity prices on various sectors, a precise analysis of this data is crucial for stakeholders, including policymakers, energy providers, and investors. Such an analysis enables informed decision-making that promotes economic efficiency and ensures a reliable electricity supply.

The dataset under examination consists of hourly electricity prices from the Spanish Electricity Market, provided by the Spanish Electricity Market Operator (www.omie.es), covering the period from January 1, 2014, to December 31, 2014. This dataset has been previously analyzed in Mestre et al. (2021), where the functional autocorrelation and partial autocorrelation functions were investigated. However, their analysis did not include a comprehensive modeling procedure within the context of our proposed framework, highlighting the need for a thorough evaluation of model adequacy in this important application.

We represent the original price curves, after linear interpolation, as $\{X_t(\tau), t = 1, 2, \dots, 365, \tau \in [1, 24]\}$. The left panel of Figure 3 displays the average daily electricity price curves for each day of the week, with each day differentiated by color. As shown in the left panel of Figure 3, the curves exhibit a distinct intraday pattern, with

prices being lower during the early morning hours and higher during peak demand periods. Additionally, average electricity prices on weekends are notably lower than on weekdays, reflecting typical market behavior influenced by reduced industrial demand. To account for the strong weekly effect present in the data, we define the differenced series $\{Y_t(\tau)\}$, where $Y_t(\tau) = X_t(\tau) - X_{t-7}(\tau)$ and present this series in the right panel of Figure 3. We then assess the stationarity of $\{Y_t(\tau)\}$ using the test proposed by Horváth et al. (2014), yielding a p -value of 0.785. Based on this result, we conclude that the differenced sequence $\{Y_t(\tau)\}$ can be considered stationary at the 5% significance level.

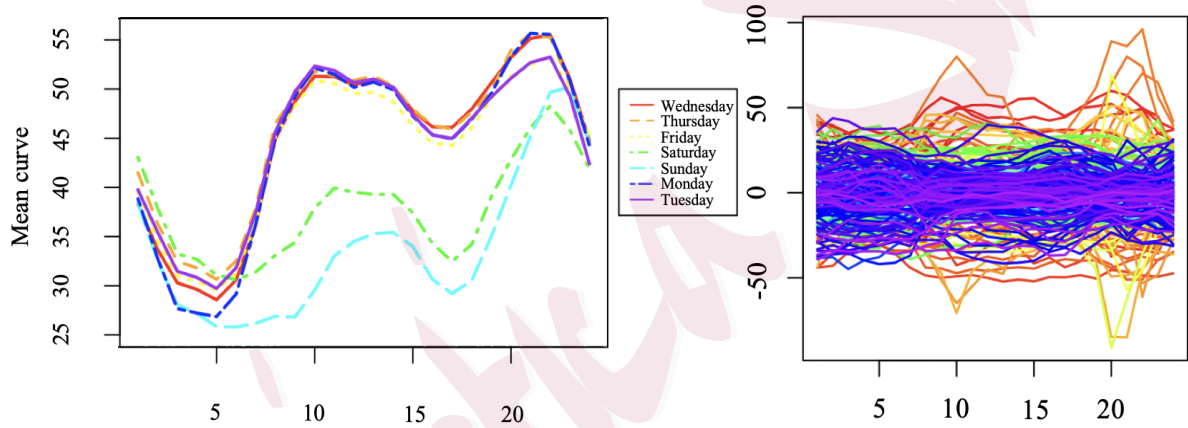


Figure 3. Left: Average daily electricity price curves $\{X_t(\tau)\}$ for each day of the week in the Spanish market in 2014. Right: The corresponding weekly differenced curves $\{Y_t(\tau)\}$.

Next, we apply our test to the differenced series $\{Y_t(\tau)\}$. As demonstrated in the simulation study (Section 5), the test statistic exhibits robustness with respect to the choices of tuning parameters. This robustness is also reflected in the real data application, where consistent conclusions are obtained across various settings. Specifically, we apply our BRWB bootstrap test to $\{Y_t(\tau)\}$ with $b_T = 1, 3, 6$, $K = 1, 3, 5, 8, 10$ and random weights drawn from both $N(0, 1)$ and Bernoulli distribution as specified in (5.2). In all cases, our test strongly rejects the null hypothesis of serial uncorrelatedness, with

extremely small p -values below 10^{-4} . Furthermore, The left panel of Figure 5 presents the sample autocorrelogram along with the 95% upper confidence bounds under the WWN assumption (using our BRWB test with $b_T = 3$ and Bernoulli weights), the MDS bound from the KRS test, and the IID bound proposed by Mestre et al. (2021). Significant correlations are observed at several lags across all tests, with a particularly large value at lag 7, suggesting the presence of serial correlations and a potential seasonal effect.

In light of this, we propose to use the Hyndman-Ullah method (Hyndman and Ullah, 2007), referred to as the “HU” model, to fit the data. This approach involves three key steps: (i) By utilizing the idea of truncated Karhunen–Loève expansion, we project each of the functional time series onto the first J eigenfunctions of the sample covariance operator, thereby obtaining what is known as the functional principal component (FPC) scores; (ii) for each FPC score series, we fit a seasonal autoregressive integrated moving average (SARIMA) model. In particular, different models may be employed for different score series; (iii) using the models from step (ii), we compute fitted values for the score series and subsequently recover the corresponding fitted curves using the FPCs. This entire process then results in residual curves.

In our analysis, we select a truncation parameter of $J = 3$, applying the corresponding HU(3) model to the process $\{Y_t(\tau)\}$. Each score series is fitted with a SARIMA model specified as follows: $(1, 0, 1) \times (1, 0, 0)_7$, $(2, 1, 0) \times (1, 0, 0)_7$, and $(0, 0, 0) \times (1, 0, 0)_7$. Figure 4 displays the fitted curves (purple dashed lines) for four selected days in 2014. A visual comparison of these fitted curves with the actual curves (blue lines) demonstrates that the HU(3) model provides an excellent fit. Furthermore, the right panel of Figure 5 shows the sample autocorrelogram of the residuals from the fitted HU(3) model, along with

the 95% WWN upper bound with $b_T = 3$ and standard exponential random weightings, 95% MDS upper bound and 95% IID upper bound. We observe that $\hat{\rho}_h$ at several lags exceed the 95% IID upper bound, and the sample autocorrelation at lag 7 exceeds the MDS upper bound, while all values lie within the WWN upper bound. This discrepancy suggests that the IID and MDS upper bounds reject the hypothesis of model adequacy. However, our test, which accounts for more general dependencies, provides more robust upper bounds that support the adequacy of the fitted model. Therefore, this approach not only confirms the suitability of the HU(3) model but also helps mitigate the risk of overfitting, ensuring a more reliable evaluation of model performance.

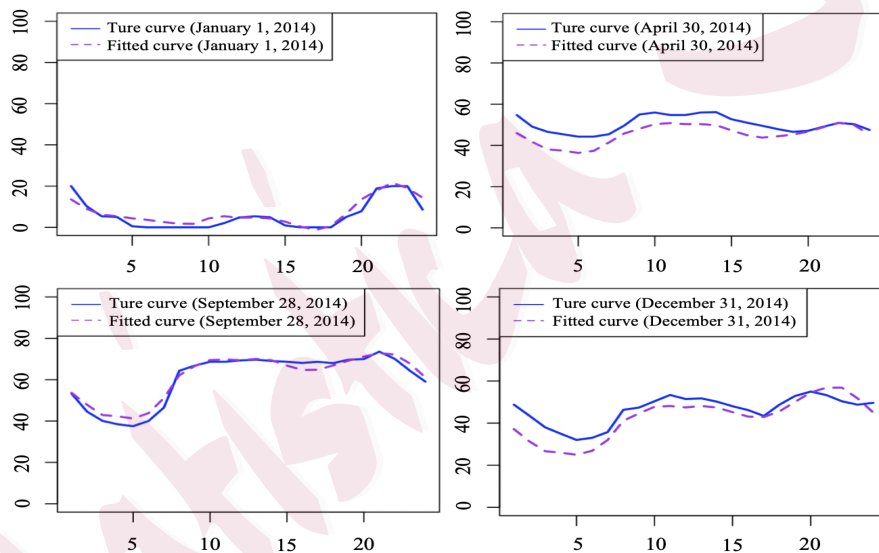


Figure 4. Fitted electricity price curves using HU(3) model for January 1, April 30, September 28, and December 31 in 2014 (dashed lines), together with true curves for those days (solid lines).

7. Conclusion

In this paper, we propose a blockwise random weighting bootstrap (BRWB) procedure for general functional white noise checks and establish its asymptotic validity using

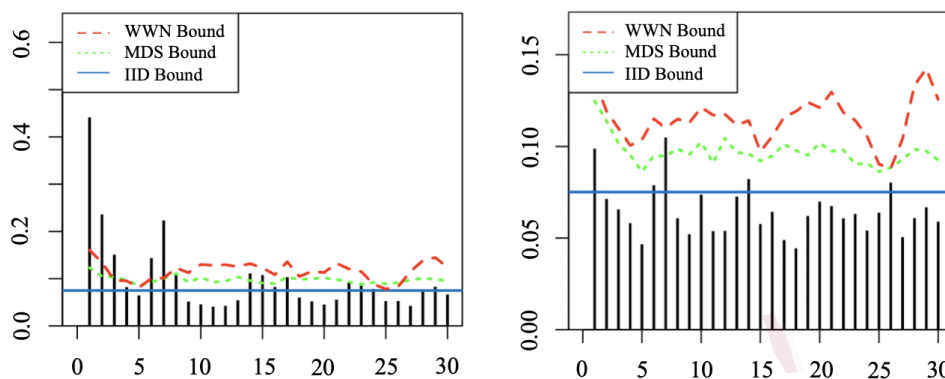


Figure 5. Left: The autocorrelogram of $\{Y_t(\tau)\}$ with 95% WWN, MDS and IID upper bounds. Right: The autocorrelogram of residuals with 95% WWN, MDS and IID upper bounds.

empirical process theory. Our approach offers several notable advantages. First, our test accommodates a broader white noise structure, which effectively addresses potential over-rejections that can arise from restrictive IID or MDS assumptions commonly used in white noise testing. Second, the test is constructed based on squared functional autocorrelation functions, which differs from methods that rely on the functional principal components. This design eliminates the need to pre-select the number of principal components, enhancing robustness in respect to variations in autocovariance magnitude. Third, the proposed BRWB procedure is robust to the selection of block sizes and random weights and can be readily extended to adequacy checks for weak functional autoregressive models.

In summary, our approach builds upon the classical Box-Pierce-type portmanteau test and extends it to the Hilbert space, providing valuable theoretical insights for modeling functional time series. Moreover, we offer practical support through an R package to facilitate the implementation of general functional white noise checks.

Despite these contributions, there are certain limitations. Notably, while we extend the proposed test to adequacy checks for weak functional autoregressive models and in-

introduce a corresponding bootstrap procedure, as discussed in Section 5.2, the theoretical investigation of how estimation affects the asymptotic distribution of the test remains incomplete. Additionally, our analysis focuses on global power, while local power, which warrants further exploration, is left for future study.

Supplementary Material

All technical details and additional numerical studies can be found in the Supplementary Material.

Acknowledgments

Muyi Li acknowledges the financial support of NSFC (72073112, 72033008, 71988101), Natural Science Foundation of Fujian Province (2025J01036), and Key program of National Bureau of Statistics (2023LZ019). Wai Keung Li would like to thank the partial support of EdUHK grant RG44/2019-2020R. Xingbai Xu acknowledges the financial support of NSFC (72073110, 72473118). The authors thank for Yuhan Chi (Xiamen University) providing technical support.

References

- Bagchi, P., V. Characiejus, and H. Dette (2018). A simple test for white noise in functional time series. Journal of Time Series Analysis 39(1), 54–74.
- Bosq, D. (2000). Linear processes in function spaces: theory and applications, Volume 149. Springer Science & Business Media.

- Cerovecki, C., C. Francq, S. Hörmann, and J.-M. Zakoïan (2019). Functional garch models: The quasi-likelihood approach and its applications. Journal of Econometrics 209(2), 353–375.
- Characiejus, V. and G. Rice (2020). A general white noise test based on kernel lag-window estimates of the spectral density operator. Econometrics and Statistics 13, 175–196.
- Chen, Y., T. Koch, K. G. Lim, X. Xu, and N. Zakiyeva (2021). A review study of functional autoregressive models with application to energy forecasting. Wiley Interdisciplinary Reviews: Computational Statistics 13(3), e1525.
- Dehling, H., O. S. Sharipov, and M. Wendler (2015). Bootstrap for dependent hilbert space-valued random variables with application to von mises statistics. Journal of Multivariate analysis 133, 200–215.
- Francq, C. and H. Raïssi (2007). Multivariate portmanteau test for autoregressive models with uncorrelated but nonindependent errors. Journal of Time Series Analysis 28(3), 454–470.
- Gabrys, R. and P. Kokoszka (2007). Portmanteau test of independence for functional observations. Journal of the American Statistical Association 102(480), 1338–1348.
- Hörmann, S., L. Kidziński, and M. Hallin (2015). Dynamic functional principal components. Journal of the Royal Statistical Society Series B: Statistical Methodology 77(2), 319–348.
- Hörmann, S. and P. Kokoszka (2010). Weakly dependent functional data. The Annals of Statistics 38(3), 1845–1884.
- Horváth, L., M. Hušková, and G. Rice (2013). Test of independence for functional data. Journal of Multivariate Analysis 117, 100–119.
- Horváth, L. and P. Kokoszka (2012). Inference for functional data with applications, Volume 200. Springer Science & Business Media.
- Horváth, L., P. Kokoszka, and G. Rice (2014). Testing stationarity of functional time series. Journal of

Econometrics 179(1), 66–82.

Horváth, L., G. Rice, and S. Whipple (2016). Adaptive bandwidth selection in the long run covariance estimator of functional time series. Computational Statistics & Data Analysis 100, 676–693.

Hyndman, R. J. and M. S. Ullah (2007). Robust forecasting of mortality and fertility rates: A functional data approach. Computational Statistics & Data Analysis 51(10), 4942–4956.

Kargin, V. and A. Onatski (2008). Curve forecasting by functional autoregression. Journal of Multivariate Analysis 99(10), 2508–2526.

Kim, M., P. Kokoszka, and G. Rice (2023). White noise testing for functional time series. Statistic Surveys 17, 119–168.

Kokoszka, P. and N. Mohammadi Jouzdani (2020). Frequency domain theory for functional time series: Variance decomposition and an invariance principle. Bernoulli 26(3), 2383–2399.

Kokoszka, P. and M. Reimherr (2013). Determining the order of the functional autoregressive model. Journal of Time Series Analysis 34(1), 116–129.

Kokoszka, P., G. Rice, and H. L. Shang (2017). Inference for the autocovariance of a functional time series under conditional heteroscedasticity. Journal of Multivariate Analysis 162, 32–50.

Li, M. and Y. Zhang (2022). Bootstrapping multivariate portmanteau tests for vector autoregressive models with weak assumptions on errors. Computational Statistics & Data Analysis 165, 107321.

Li, Q., C. Hsiao, and J. Zinn (2003). Consistent specification tests for semiparametric/nonparametric models based on series estimation methods. Journal of Econometrics 112(2), 295–325.

Liu, R. Y. (1988). Bootstrap procedures under some non-iid models. The Annals of Statistics 16(4), 1696–1708.

Lobato, I. N., J. C. Nankervis, and N. Savin (2002). Testing for zero autocorrelation in the presence of statistical dependence. Econometric Theory 18(3), 730–743.

- Mainassara, Y. B. (2011). Multivariate portmanteau test for structural varma models with uncorrelated but non-independent error terms. Journal of Statistical Planning and Inference 141(8), 2961–2975.
- Mammen, E. (1993). Bootstrap and wild bootstrap for high dimensional linear models. The Annals of Statistics 21(1), 255–285.
- Mestre, G., J. Portela, G. Rice, A. M. San Roque, and E. Alonso (2021). Functional time series model identification and diagnosis by means of auto-and partial autocorrelation analysis. Computational Statistics & Data Analysis 155, 107108.
- Panaretos, V. M. and S. Tavakoli (2013). Fourier analysis of stationary time series in function space. The Annals of Statistics 41(2), 568–603.
- Patilea, V., C. Sánchez-Sellero, and M. Saumard (2016). Testing the predictor effect on a functional response. Journal of the American Statistical Association 111(516), 1684–1695.
- Pilavakis, D., E. Paparoditis, and T. Sapatinas (2020). Testing equality of autocovariance operators for functional time series. Journal of Time Series Analysis 41(4), 571–589.
- Politis, D. N. and J. P. Romano (1994a). Limit theorems for weakly dependent hilbert space valued random variables with application to the stationary bootstrap. Statistica Sinica 4(2), 461–476.
- Politis, D. N. and J. P. Romano (1994b). The stationary bootstrap. Journal of the American Statistical association 89(428), 1303–1313.
- Priestly, M. (1981). Spectral Analysis and Time Series. Academic Press, London.
- Ramsay, J. O. and G. Hooker (2009). Functional data analysis with R and Matlab. Springer science & business media.
- Ramsay, J. O. and B. W. Silverman (2005). Functional Data Analysis. Springer, New York.
- Shao, X. (2011). A bootstrap-assisted spectral test of white noise under unknown dependence. Journal of

Econometrics 162(2), 213–224.

Wu, C.-F. J. (1986). Jackknife, bootstrap and other resampling methods in regression analysis. the Annals of Statistics 14(4), 1261–1295.

Wu, W. and X. Shao (2004). Limit theorems for iterated random functions. Journal of Applied Probability 41(2), 425–436.

Yeh, C.-K., G. Rice, and J. A. Dubin (2023). Functional spherical autocorrelation: A robust estimate of the autocorrelation of a functional time series. Electronic Journal of Statistics 17(1), 650–687.

Zhang, X. (2016). White noise testing and model diagnostic checking for functional time series. Journal of Econometrics 194(1), 76–95.

Zhu, K. (2016). Bootstrapping the portmanteau tests in weak auto-regressive moving average models. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 78(2), 463–485.

Department of Statistics and Data Science, School of Economics, Xiamen University, China

E-mail: (miaoyu@stu.xmu.edu.cn)

Wang Yanan Institute for Studies in Economics (WISE) and Department of Statistics and Data Science, School of Economics, Xiamen University, China

E-mail: (limuyi@xmu.edu.cn)

Department of Mathematics and Information Technology, The Education University of Hong Kong, Hong Kong

E-mail: (waikeungli@eduhk.hk)

Wang Yanan Institute for Studies in Economics (WISE) and Department of Statistics and Data Science, School of Economics, Xiamen University, China

E-mail: (xuxingbai@xmu.edu.cn)