Test of independence for functional data

Lajos Horváth a, Marie Hušková b, Gregory Rice a,∗

a Department of Mathematics, University of Utah, Salt Lake City, UT 84112-0090, USA
b Department of Statistics, MFF UK, Sokolovská 83, CZ-18600 Praha, Czech Republic

ABSTRACT

We wish to test the null hypothesis that a collection of functional observations are independent and identically distributed. Our procedure is based on the sum of the $L^2$ norms of the empirical correlation functions. The limit distribution of the proposed test statistic is established under the null hypothesis. Under the alternative the sample exhibits serial correlation, and consistency is shown when the sample size as well as the number of lags used in the test statistic tend to $\infty$. A Monte Carlo study illustrates the small sample behavior of the test and the procedure is applied to data sets, Eurodollar futures and magnetogram records.

1. Introduction and results

The majority of statistical methods in the book of Ramsay and Silverman [36] are based on the assumption that the functional observations are independent and identically distributed. This assumption is ensured in designed experiments (cf. [31]) or reasonably justified by data generating processes (cf. [1, 11, 4]) in some cases. However, it is not always obvious that this assumption holds if data is generated by observing a time series, as is often the case with physical phenomena (cf. [30]) or economic activities (cf. [23]). It has been observed by Horváth et al. [21] that neglecting dependence in functional observations reduces the power of statistical tests and may cause misleading results. Hence it is crucial when using these procedures to check for the validity of the independence assumption in the data.

Tests for the independence of real and vector valued observations have been developed in the time series literature (cf. [9, 27]). Due to the popularity of the Box–Ljung–Pierce approach (cf. [7, 26]), the majority of tests used to check the independence assumption verify that all autocovariances and/or autocorrelations up to lag $H$ are suitably close to 0. The univariate results of Box and Pierce [7] and Ljung and Box [26] were extended to multivariate time series by Chitturi [12], Hosking [22] and Li and McLeod [25]. It is noted in [17] that the methods used for univariate and vector valued time series cannot be used in the case of functional observations.

In this paper we follow the Box–Ljung–Pierce approach in the case of functional observations and provide autocovariance based procedures to test if a given functional time series is a white noise sequence. We assume that we have observations $X_1(t), \ldots, X_n(t)$ which are square integrable random functions on $[0, 1]$. Let $\|f\| = (\int f^2(t) dt)^{1/2}$, where $f = f_0$.
We wish to test the null hypothesis

\[ H_0 : X_1, \ldots, X_n \text{ are independent and identically distributed random functions}, \]

against the alternative

\[ H_A : X_1, \ldots, X_n \text{ is a stationary and ergodic sequence such that for some } h_0 \geq 1 \]

\[ \int C_{h_0}^2(t, s) ds > 0, \text{ where } C_{h_0}(t, s) = \text{cov}(X_1(t), X_{1+h_0}(s)). \]

Such tests are typically referred to as “portmanteau” tests; the null hypothesis \( H_0 \) is well specified, however any test of \( H_0 \) cannot have power against all possible alternatives. The reason for defining \( H_A \) as we have done is quite simple. Since this test is to be applied to a functional time series, it should have power to detect whether the data is a white noise sequence, or if instead it follows one of the available models for dependent functional time series, such as the functional autoregressive model of order one (FAR(1)) (cf. [23,2]), or functional linear process (cf. [6]). Such models exhibit serial correlation by construction and thus will satisfy \( H_A \). A similar argument is behind the definition of the alternative hypothesis for the analogous tests for univariate and multivariate data, in which case the data follows an ARIMA model under the alternative.

To motivate the definition of the test statistic we note that under \( H_0 \), \( \text{cov}(X_1(t), X_{1+h}(s)) = 0 \) for all \( h \geq 1 \) and for almost all \( (s, t) \in [0, 1]^2 \), hence all sample autocovariance functions

\[ \hat{C}_{n,h}(t, s) = \frac{1}{n} \sum_{i=1}^{n-h} (X_i(t) - \bar{X}_n(t)) (X_{i+h}(s) - \bar{X}_n(s)), \quad h \geq 0, \]

should be close to 0 for all \( h \geq 1 \), where

\[ \bar{X}_n(t) = \frac{1}{n} \sum_{i=1}^{n} X_i(t) \]

denotes the sample mean. The function \( \hat{C}_{n,0}(t, s) \) is an estimator for the covariance function

\[ C_0(t, s) = E[(X_1(t) - EX_1(t))(X_1(s) - EX_1(s))], \]

assuming that \( X_i, i \geq 1 \) is a stationary sequence. The weak convergence of the covariance operator given by \( \hat{C}_{n,0} \) was studied by Mas [29]. Panaretos et al. [32] and Fremdt et al. [15] used the estimated covariance function \( \hat{C}_{n,0} \) in statistical inference. The testing procedures for multivariate observations in [12,22,25] are based on quadratic forms of the sample correlation or covariance matrices. In our approach these quadratic forms are replaced with the square integrals of the sample covariance functions \( \hat{C}_{n,h}(t, s) \). Under \( H_A \) at least one of the functions \( \hat{C}_{n,h}(t, s) \) is significantly different from zero, and so we reject \( H_0 \) if

\[ \hat{V}_{n,H} = \sum_{h=1}^{H} \int \hat{C}_{n,h}^2(t, s) dtds \]

is large.

In the case of univariate and vector valued observations, the number of correlations used in the testing procedure has been a fixed number. The limit distributions of these portmanteau tests are \( \chi^2 \), where the degree of freedom depends on the number of lags used. However, it has been observed that the \( \chi^2 \) approximation might be poor and several modifications have been suggested (cf. [34,24,14,28]). The approximation tends to work well if both the sample size and the number of lags used to define the statistic are large. Hence, in this paper we consider the case when \( H \), the number of lags used in the definition of \( \hat{V}_{n,H} \), depends on the sample and tends to \( \infty \) as the sample size increases:

**Assumption 1.1.** \( H = H(n) \to \infty \) and \( H = O((\log n)^{\alpha}) \) with some \( \alpha > 0 \), as \( n \to \infty \).

First we consider the asymptotic behavior of \( \hat{V}_{n,H} \) under the null hypothesis. We assume

**Assumption 1.2.** \( E\|X_i\|^4 < \infty \), for all \( i \geq 1 \),

which implies that \( C_0(t, s) \) is square integrable, and there are eigenvalues \( \lambda_1 \geq \lambda_2 \geq \cdots \geq 0 \) and orthonormal eigenfunctions \( \varphi_1, \varphi_2, \ldots \) satisfying

\[ \lambda_i \varphi_i(t) = \int C_0(t, s) \varphi_i(s) ds. \quad (1.1) \]

We also note that

\[ C_0(t, s) = \sum_{i=1}^{\infty} \lambda_i \varphi_i(t) \varphi_i(s) \quad (1.2) \]
It is a simple consequence of the Cauchy–Schwarz inequality that under Assumption 1.2
\[ \sum_{\ell=1}^{\infty} \lambda_{\ell} < \infty \quad \text{and therefore} \quad \sum_{\ell=1}^{\infty} \lambda_{\ell}^2 < \infty. \tag{1.3} \]

**Theorem 1.1.** If $H_0$, Assumptions 1.1 and 1.2 are satisfied, then
\[ \frac{1}{(2H\sigma^2)^{1/2}} \left\{ \hat{n} \hat{\nu}_{n,H} - H \hat{\mu} \right\} \overset{D}{\to} N(0, 1), \]
where $N(0, 1)$ stands for a standard normal random variable.

\[ \mu = \left( \sum_{\ell=1}^{\infty} \lambda_{\ell} \right)^2 \quad \text{and} \quad \sigma^2 = \left( \sum_{\ell=1}^{\infty} \lambda_{\ell}^2 \right)^2. \]

It follows from (1.3) that both $\mu$ and $\sigma^2$ are finite. Since the eigenvalues are unknown in practice we need to estimate $\mu$ and $\sigma^2$ from the sample. It follows from (1.2) that
\[ \int C_0(t, t) dt = \sum_{\ell=1}^{\infty} \lambda_{\ell} \quad \text{and} \quad \iint C_0(t, s) dt ds = \sum_{\ell=1}^{\infty} \lambda_{\ell}^2. \]

Since $\hat{C}_{n,H}(t, s)$, the sample covariance, is a well established and studied estimator for $C_0(t, s)$, we use
\[ \hat{\mu}_n = \left( \int \hat{C}_{n,0}(t, t) dt \right)^2 \quad \text{and} \quad \hat{\sigma}_n^2 = \left( \iint \hat{C}_{n,0}^2(t, s) dt ds \right)^2 \]
to estimate $\mu$ and $\sigma^2$. Interestingly, Theorem 1.1 remains true when the theoretical $\mu$ and $\sigma^2$ are replaced with the corresponding estimators.

**Theorem 1.2.** If $H_0$, Assumptions 1.1 and 1.2 are satisfied, then
\[ \frac{1}{(2H\sigma^2)^{1/2}} \left\{ n \hat{\nu}_{n,H} - H \hat{\mu} \hat{n} \right\} \overset{D}{\to} N(0, 1), \]
where $N(0, 1)$ stands for a standard normal random variable.

Gabrys and Kokoszka [17] and Gabrys et al. [16] provide a different approach to test $H_0$. By the Karhunen–Loève expansion (cf. [20]) we have
\[ X_i(t) = \sum_{\ell=1}^{\infty} \xi_{i,\ell} \varphi_{\ell}(t), \tag{1.4} \]
where $\xi_{i,\ell} \equiv \langle X_i, \varphi_{\ell} \rangle$ and $\langle f, g \rangle = \int f(t)g(t) dt$ denotes the inner product in the Hilbert space $L^2$. Gabrys and Kokoszka [17] suggest using the first $p$ coefficients in expansion (1.4) and test if the vectors $(\xi_{i,1}, \ldots, \xi_{i,p})^T \in \mathbb{R}^p$ are independent using the sum of the sample correlation matrices up to lag $H$. The portmanteau test in [17] is extended to test for independence in the residuals of a functional linear model by Gabrys et al. [16]. In the papers of Gabrys and Kokoszka [17] and Gabrys et al. [16] it is assumed that the number of lags $H$ and the number of projections used $p$ are fixed and do not depend on the sample size. Furthermore, since the eigenfunctions $\varphi_{\ell}$ are unknown, they are estimated with $\hat{\varphi}_{\ell}$ from the sample and the empirical projections $(\hat{X}_i, \hat{\varphi}_{\ell})$ are used in the statistical analysis. They showed under the independence null hypothesis that the limit distribution of the portmanteau test statistic is $\chi^2$ with $pH$ degrees of freedom. For large $p$ and $H$ the $\chi^2$ is suitable. The choice of $p$ is widely discussed in the literature (cf. [36]), but there is no optimal choice. The estimation of the $\varphi_{\ell}$’s uses the fda package, where the observations are replaced with smoothed curves. The extra smoothing step can introduce bias and might mask important features of the original data.

In our approach we do not use projections so we do not need to fix the value of $p$, and we allow $H$ to increase with the sample size. Also, our method has the advantage that the empirical eigenfunctions $\hat{\varphi}_{\ell}$ need not be computed.

Next we consider the consistency of the testing procedure based on Theorem 1.2 under $H_0$. We only require under the alternative $H_A$ that

**Assumption 1.3.** $E\|X_i\|^2 < \infty$

and

**Assumption 1.4.** $H = H(n), n/H^{3/2} \rightarrow \infty$. 

Theorem 1.3. If $H_A$, Assumptions 1.3 and 1.4 are satisfied, then

$$\frac{1}{(2H\hat{\sigma}_n^2)^{1/2}} \left\{ n\hat{V}_{n,H} - H\hat{\mu}_n \right\} \overset{p}{\rightarrow} \infty.$$  

Next we show an example when the conditions of Theorem 1.3 are satisfied.

Example 1.1. The FAR(1) process satisfies the recursion

$$X_n(t) = \int \psi(t, u)X_{n-1}(u)du + \varepsilon_n(t), \quad 0 \leq t \leq 1, \quad -\infty < n < \infty,$$  

(1.5)

where $\varepsilon_n(t), -\infty < n < \infty$ are independent and identically distributed random processes with $E\varepsilon_n(t) = 0$ and $E\|\varepsilon_n\| < \infty$. Bosq [6] proved, if $\|\psi\| < 1$, then (1.5) has a unique stationary and ergodic solution. Since $C_0(t, s)$ is a positive definite function we have (1.1) and (1.2), and we can assume that $\varphi_{\ell}, 1 \leq \ell < \infty$ is a basis or it can be extended into a basis. The square integrability of $\psi$ yields that

$$\psi(s, u) = \sum_{i,j=1}^{\infty} \mu_{ij} \varphi_i(s) \varphi_j(u).$$  

(1.6)

Multiplying both sides of (1.5) with $X_{n-1}(t)$ and then taking expected values we obtain the equation

$$C_1(t, s) = \int \psi(s, u)C_0(u, t)du.$$  

Now the expansions in (1.2) and (1.6) with the orthonormality of the $\varphi_i$’s imply

$$C_1(t, s) = \sum_{i,j=1}^{\infty} \mu_{ij} \lambda_i \lambda_j \varphi_i(s) \varphi_j(t)$$  

and therefore

$$\int \int C_1^2(t, s)dtds = \sum_{i,j=1}^{\infty} \mu_{ij}^2 \lambda_i^2.$$  

If all $\lambda_i$’s are positive (i.e. $X_n$ is not spanned by finitely many functions) and $\|\psi\| > 0$, then $H_A$ holds with $h_0 = 1$.

Remark 1.1. One can show that under the conditions of Theorem 1.1 for every fixed $h_0$

$$\left( \int \int n^2_{\varphi_n}^2(t, s)dtds, \quad 1 \leq h \leq h_0 \right) \overset{D}{\rightarrow} (\varphi_n, 1 \leq h \leq h_0),$$  

where $\varphi_n, 1 \leq h \leq h_0$ are independent identically distributed random variables which can be represented as an infinite sum of weighted $\chi^2(1)$ random variables. However, the weights depend on the unknown eigenvalues $\lambda_1, \lambda_2, \ldots$, so the implementation of the Bonferroni inequality is not immediate when finitely many lags are used.

Remark 1.2. To have a model for a local alternative in the functional setting one would consider

$$X_i(t) = X_{in}(t) = Y_i(t) + \gamma_nZ_i(t), \quad 0 \leq t \leq 1, \quad 1 \leq i \leq n,$$  

where $\{Y_i\}$ is an independent, identically distributed sequence, and $\{Z_i\}$ is a Bernoulli shift, $\{4 + \delta\}$-decomposable for some $\delta > 0$ in the sense of Hörmann and Kokoszka [19]. We assume that $\{Y_i\}$ and $\{Z_i\}$ are independent of each other, $EY_i(t) = EZ_i(t) = 0$, and $E|Y_i|^4 < \infty$. The deviation from the null hypothesis is asymptotically small in the sense that $H^{1/2}\gamma_n \rightarrow 0$ as $n \rightarrow \infty$. The sequence $\{Z_i\}$ is correlated at some lag, which means that there is $h_0 \geq 1$ such that $\int \int A_{h_0}^2(t, s)dtds > 0$, where

$$A_{h_0}(t, s) = EZ_i(t)Z_{i+h_0}(s).$$  

(1.7)

We also assume that $H$ satisfies Assumption 1.1. We illustrate at the end of Section 4 that if $\gamma_n^4nH^{1/2} \rightarrow 0$, then Theorem 1.2 remains true. If $\gamma_n^4nH^{1/2} \rightarrow \infty$, then we reject $H_0$ with probability tending to 1, since $(n\hat{V}_{n,H} - H\hat{\mu}_n)/(2H\hat{\sigma}_n^2)^{1/2} \rightarrow \infty$ in probability. In the borderline case when $\gamma_n^4nH^{1/2} \rightarrow c > 0$, then $(n\hat{V}_{n,H} - H\hat{\mu}_n)/(2H\hat{\sigma}_n^2)^{1/2}$ converges to a normal random variable with mean $c(2\hat{\sigma}_n^2)^{-1/2} \sum_{i=1}^{\infty} \int A_{h_0}^2(t, s)dtds$. In this case the rejection probability does not go to zero as $n \rightarrow \infty$. The asymptotic variance $\hat{\sigma}_n^2$ is defined as the limit of the sum of the squares of the eigenvalues of $E \left[ X_{1,n}(t)X_{1,n}(s) \right]$ as $n \rightarrow \infty$. We note that $\hat{\sigma}_n^2$ is also the sum of the squares of the eigenvalues of $E \left[ Y_{1}(t)Y_{1}(s) \right]$. 

Author's personal copy
2. Simulations and applications

In the first part of this section we investigate the finite-sample properties of the test in Theorem 1.2. The estimator $\hat{c}_{n,h}$ is biased, so to improve the finite sample performance we replaced $H\hat{\mu}_n$ with $H^+\bar{\mu}_n$ with $H^+ = H - H(H + 1)/(2n)$. Let

$$V_{n,H} = (2H\hat{\sigma}_n^2)^{-1/2} \left[ n\hat{V}_{n,H} - H^+\hat{\mu}_n \right].$$

To investigate the size of the test in the case of finite sample sizes, we simulated sample paths of independent Wiener processes (standard Brownian motions) and Brownian bridges. To simulate Wiener processes, cumulative sums of independent normal random variables were computed on a grid of 100 equi-spaced points in the interval $[0, 1]$ and linearly connected.

Table 2.1 shows the percentage of $V_{n,H} \geq \Phi^{-1}(1-\alpha)$ ($\alpha = 0.1, 0.05, 0.01$) based on 1000 repetitions. Table 2.2 provides the same information when the $X_i$'s are independent and identically distributed Brownian bridges.

As one can see from the empirical sizes in Tables 2.1 and 2.2, the convergence of the distribution of the test statistic to that of the standard normal appears to be somewhat slow in the tails. This is not surprising, since for every $h$ the limit of $n\hat{V}_{n,H}$ is a weighted sum of independent $\chi^2(1)$ random variables, so the limit has an exponential tail, where $\chi^2(v)$ denotes a $\chi^2$ random variable with $v$ degrees of freedom. It has been known since [13] that the normal approximation to the $\chi^2$ when the degree of freedom is large does not work well on the tails. To get better finite sample performance, transformations of the $\chi^2$ variables have been suggested (cf. [13,33,10]). By a simple manipulation, Theorem 1.2 implies that

$$\left(\frac{2H\hat{\mu}_n}{\hat{\sigma}_n^2}\right)^{-1/2} \left( n\hat{V}_{n,H}\hat{\mu}_n - H\hat{\mu}_n^2 \right) \xrightarrow{D} \mathcal{N}(0, 1).$$

This can be interpreted as $n\hat{V}_{n,H}\hat{\sigma}_n^2/\mu$ being distributed approximately as a $\chi^2(nH\mu^2/\sigma^2)$ random variable. It is shown in [10] that if

$$m(v) = \frac{5}{6} - \frac{1}{9v} - \frac{7}{648v^2} + \frac{25}{2187v^3},$$

and

$$v(v) = \frac{1}{18v} + \frac{1}{162v^2} - \frac{37}{11664v^3},$$

then the power transformation

$$\mathcal{T}(\chi^2(v)) = v(v)^{-1/2} \left[ \left( \frac{\chi^2(v)}{v} \right)^{1/6} - \frac{1}{2} \left( \frac{\chi^2(v)}{v} \right)^{1/3} + \frac{1}{3} \left( \frac{\chi^2(v)}{v} \right)^{1/2} - m(v) \right]$$

is closer to being distributed as the standard normal in terms of the maximum absolute error than the ordinary normal approximation of a $\chi^2(v)$ random variable. Applying this transformation to $n\hat{V}_{n,H}\hat{\mu}_n/\hat{\sigma}_n^2$ and using degrees of freedom $H\hat{\mu}_n^2/\hat{\sigma}_n^2$ provides some improvement to the empirical size of the test. Fig. 2.1 shows the cumulative distribution function of the standard normal along with the empirical distribution functions of $V_{n,H}$ and $\mathcal{T}(n\hat{V}_{n,H}\hat{\mu}_n/\hat{\sigma}_n^2)$ computed from 1000 simulations using $n = 100$ and $H = 20$.

The study of the power of the test is based on the alternative when $X_i(t)$ is a Hilbert–space valued autoregressive(1) process (HAR(1)) of Example 1.1. In order to compare the power of the test in Theorem 1.2 with Table 2 in [17] we used the kernel

$$\psi_c(t,s) = c \exp \left( \frac{t^2 + s^2}{2} \right).$$

With the choice of $c = 0.3416$, we have that $\|\psi_c\| \approx 0.5$. 

### Table 2.1

<table>
<thead>
<tr>
<th>n</th>
<th>H</th>
<th>10%</th>
<th>5%</th>
<th>1%</th>
<th>H</th>
<th>10%</th>
<th>5%</th>
<th>1%</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>5</td>
<td>8.3</td>
<td>5.3</td>
<td>2.9</td>
<td>100</td>
<td>8.9</td>
<td>5.2</td>
<td>2.3</td>
</tr>
<tr>
<td>10</td>
<td>9.2</td>
<td>5.4</td>
<td>2.7</td>
<td>20</td>
<td>10.1</td>
<td>6.6</td>
<td>3.5</td>
<td>2.3</td>
</tr>
<tr>
<td>20</td>
<td>9.0</td>
<td>5.8</td>
<td>2.9</td>
<td>30</td>
<td>10.4</td>
<td>5.6</td>
<td>2.1</td>
<td>2.4</td>
</tr>
<tr>
<td>25</td>
<td>9.1</td>
<td>5.6</td>
<td>2.8</td>
<td>40</td>
<td>9.9</td>
<td>6.2</td>
<td>2.4</td>
<td>2.4</td>
</tr>
<tr>
<td>200</td>
<td>10</td>
<td>10.2</td>
<td>5.9</td>
<td>2.1</td>
<td>250</td>
<td>10.7</td>
<td>5.8</td>
<td>2.5</td>
</tr>
<tr>
<td>20</td>
<td>10.9</td>
<td>5.5</td>
<td>2.1</td>
<td>20</td>
<td>9.4</td>
<td>4.8</td>
<td>2.0</td>
<td>1.6</td>
</tr>
<tr>
<td>30</td>
<td>9.5</td>
<td>6.1</td>
<td>2.1</td>
<td>30</td>
<td>10.9</td>
<td>5.7</td>
<td>2.0</td>
<td>2.0</td>
</tr>
<tr>
<td>40</td>
<td>9.5</td>
<td>5.6</td>
<td>2.6</td>
<td>40</td>
<td>9.3</td>
<td>5.1</td>
<td>2.5</td>
<td>1.8</td>
</tr>
</tbody>
</table>

$\chi_n$ is a weighted sum of independent $\chi^2(1)$ random variables. By a simple manipulation, Theorem 1.2 implies that $\mathcal{T}(\chi^2(v)) = v(v)^{-1/2} \left[ \left( \frac{\chi^2(v)}{v} \right)^{1/6} - \frac{1}{2} \left( \frac{\chi^2(v)}{v} \right)^{1/3} + \frac{1}{3} \left( \frac{\chi^2(v)}{v} \right)^{1/2} - m(v) \right]$ is closer to being distributed as the standard normal in terms of the maximum absolute error than the ordinary normal approximation of a $\chi^2(v)$ random variable. Applying this transformation to $n\hat{V}_{n,H}\hat{\mu}_n/\hat{\sigma}_n^2$ and using degrees of freedom $H\hat{\mu}_n^2/\hat{\sigma}_n^2$ provides some improvement to the empirical size of the test. Fig. 2.1 shows the cumulative distribution function of the standard normal along with the empirical distribution functions of $V_{n,H}$ and $\mathcal{T}(n\hat{V}_{n,H}\hat{\mu}_n/\hat{\sigma}_n^2)$ computed from 1000 simulations using $n = 100$ and $H = 20$.
Table 2.2

Empirical size of the test when the observations are independent Brownian bridges.

<table>
<thead>
<tr>
<th>n</th>
<th>H</th>
<th>10%</th>
<th>5%</th>
<th>1%</th>
<th>n</th>
<th>H</th>
<th>10%</th>
<th>5%</th>
<th>1%</th>
<th>n</th>
<th>H</th>
<th>10%</th>
<th>5%</th>
<th>1%</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>5</td>
<td>9.4</td>
<td>5.8</td>
<td>3.2</td>
<td>100</td>
<td>10</td>
<td>9.3</td>
<td>6.2</td>
<td>2.2</td>
<td>150</td>
<td>9.6</td>
<td>4.8</td>
<td>2.3</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>8.9</td>
<td>6.1</td>
<td>2.6</td>
<td></td>
<td>20</td>
<td>10.2</td>
<td>5.7</td>
<td>2.3</td>
<td></td>
<td>20</td>
<td>10.3</td>
<td>6.2</td>
<td>2.2</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>5.5</td>
<td>5.9</td>
<td>2.7</td>
<td></td>
<td>30</td>
<td>10.0</td>
<td>7.0</td>
<td>2.9</td>
<td></td>
<td>30</td>
<td>11.1</td>
<td>6.3</td>
<td>3.2</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>5.4</td>
<td>2.5</td>
<td></td>
<td>40</td>
<td>10.5</td>
<td>7.0</td>
<td>2.9</td>
<td></td>
<td>40</td>
<td>10.8</td>
<td>7.0</td>
<td>3.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>10</td>
<td>11.4</td>
<td>6.3</td>
<td>2.4</td>
<td>250</td>
<td>11.8</td>
<td>6.7</td>
<td>2.4</td>
<td></td>
<td>300</td>
<td>9.6</td>
<td>5.9</td>
<td>1.7</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>10.6</td>
<td>6.3</td>
<td>2.1</td>
<td></td>
<td>20</td>
<td>9.4</td>
<td>5.7</td>
<td>2.3</td>
<td></td>
<td>20</td>
<td>9.8</td>
<td>5.6</td>
<td>1.8</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>10.0</td>
<td>6.4</td>
<td>2.1</td>
<td></td>
<td>30</td>
<td>10.1</td>
<td>6.2</td>
<td>1.9</td>
<td></td>
<td>30</td>
<td>9.5</td>
<td>5.2</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>9.5</td>
<td>5.7</td>
<td>2.1</td>
<td></td>
<td>40</td>
<td>10.2</td>
<td>5.8</td>
<td>1.7</td>
<td></td>
<td>40</td>
<td>10.0</td>
<td>6.3</td>
<td>1.5</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 2.1. The graphs of the standard normal distribution function (solid) and the simulated distribution functions of $V_{100,20} (\circ \circ \circ)$ and $\mathcal{F}(\hat{\nu}_{100,20}/\hat{\sigma}^2)$ ($+++$).

Table 2.3

Empirical power of test (in %).

<table>
<thead>
<tr>
<th>H</th>
<th>n = 50</th>
<th>10%</th>
<th>5%</th>
<th>1%</th>
<th>H</th>
<th>n = 100</th>
<th>10%</th>
<th>5%</th>
<th>1%</th>
<th>H</th>
<th>n = 200</th>
<th>10%</th>
<th>5%</th>
<th>1%</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>79.7</td>
<td>74.1</td>
<td>61.1</td>
<td>10</td>
<td>96.1</td>
<td>95.5</td>
<td>90.6</td>
<td></td>
<td>20</td>
<td>100</td>
<td>100</td>
<td>99.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>70.1</td>
<td>62.9</td>
<td>49.7</td>
<td>20</td>
<td>93.3</td>
<td>89.4</td>
<td>81.1</td>
<td>40</td>
<td>99.4</td>
<td>99.0</td>
<td>97.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>65.8</td>
<td>58.6</td>
<td>44.9</td>
<td>40</td>
<td>88.2</td>
<td>84.0</td>
<td>73.6</td>
<td>60</td>
<td>99.2</td>
<td>98.8</td>
<td>96.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>64.7</td>
<td>56.6</td>
<td>41.4</td>
<td>50</td>
<td>86.3</td>
<td>81.2</td>
<td>70.1</td>
<td>100</td>
<td>98.8</td>
<td>97.1</td>
<td>93.5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2.3 shows the percentage of $V_{n,H} \geq \Phi^{-1}(1 - \alpha)$ for $\alpha = 0.1, 0.05$ and 0.01 based on 1000 repetitions under the alternative when $X_n$ follows a HAR(1) process.

To illustrate the applicability of our test we consider two data sets which have been widely studied using functional data analysis, the first of which consists of Eurodollar futures rate data. Such data was considered in [23], where it is argued that the curves constructed from the prices of Eurodollar futures contracts with decreasing expiration dates follow a HAR(1) model. In this case the curves are constructed from 114 points per day; point $i$ of $X_n$ corresponds to the price of a contract with closing date $i$ months from day $n$. The sample we consider consists of 100 days of data taken from January 24 to June 17, 1997. Fig. 2.2 gives the first seven rate curves from this sample, normalized to the $[0, 1]$ interval, and Fig. 2.3 shows the same curves after centering with the sample mean. As expected, the test rejects the null hypothesis at the 1% level for all lags $H$ between 10 and 40.

We also applied our test to ground-based magnetogram records taken from Honolulu in 2001. In particular, we focused on the horizontal intensity, which is the component of the magnetic field tangent to the Earth’s surface and pointing toward the magnetic north pole. In this case each daily curve is constructed from measurements of the horizontal intensity taken every minute, giving 1440 total measurements per day. We considered 365 of such curves taken from January 1 to December 31, 2001. Seven of these curves are illustrated in Fig. 2.4. Following Xu and Kamide [39] and Gabrys and Kokoszka [17], we subtracted the linear change over each day to obtain the sample to which we applied our test. Fig. 2.5 shows this
Fig. 2.2. The graphs of seven Eurodollar futures rate curves taken from January 24 to February 1, 1997 normalized to the [0, 1] interval.

Fig. 2.3. The graphs of seven Eurodollar futures rate curves taken from January 24 to February 1, 1997 after subtracting the sample mean and normalizing to the [0, 1] interval.

transformation applied to the curves in Fig. 2.4. Using $H = 40$ we get a $p$-value less than $10^{-4}$ and therefore we reject the i.i.d. null hypothesis at the 1% level.

3. Proofs of Theorems 1.1 and 1.2

Throughout this section we assume that $H_0$ is satisfied. It follows from the definition of $\hat{V}_{n,h}$ that it does not depend on the value of $EX_1(t)$ and therefore in this section we assume $EX_1(t) = 0$ holds. First we show that $\hat{V}_{n,h}$ and $V_{n,H}$ have the same limit distribution, where

$$V_{n,h} = \sum_{h=1}^{H} \int \int C_{n,h}^2(t, s) dt ds$$

and

$$C_{n,h}(t, s) = \frac{1}{n} \sum_{i=1}^{n} X_i(t) X_{i+h}(s).$$
Lemma 3.1. If $H_0$, Assumptions 1.1 and 1.2 hold, then

$$\frac{n}{H^{1/2}} \left| \hat{V}_{n,H} - V_{n,H} \right| = o_P(1), \quad \text{as } n \to \infty,$$

(3.1)

Proof. First we show that

$$\frac{n}{H^{1/2}} \left| \hat{V}_{n,H} - \tilde{V}_{n,H} \right| = o_P(1), \quad \text{as } n \to \infty,$$

(3.1)

where

$$\tilde{V}_{n,H} = \sum_{h=1}^{H} \int \int \tilde{C}_{n,h}^2(t,s)dt ds$$
with
\[ \tilde{c}_{n,h}(t, s) = \frac{1}{n} \sum_{i=1}^{n-h} X_i(t)X_{i+h}(s). \]

It is easy to see that
\[ \left| \tilde{c}_{n,h}^2(t, s) - c_{n,h}^2(t, s) \right| \leq 2|\tilde{c}_{n,h}(t, s)a_{n,h}(t, s)| + a_{n,h}^2(t, s), \]

where \( a_{n,h}(t, s) = -a_{n,h}^{(1)}(t, s) + a_{n,h}^{(2)}(t, s) + a_{n,h}^{(3)}(t, s) \) with
\[
\begin{align*}
a_{n,h}^{(1)}(t, s) &= (1 + h/n)\tilde{X}_n(t)\tilde{X}_n(s), \\
a_{n,h}^{(2)}(t, s) &= \tilde{X}_n(t) - \frac{1}{n} \sum_{i=n-h+1}^{n} X_i(t) \quad \text{and} \quad a_{n,h}^{(3)}(t, s) = \tilde{X}_n(t) - \frac{1}{n} \sum_{i=1}^{h} X_i(t).
\end{align*}
\]

Using \( H_0 \) we get that
\[
E(\tilde{X}_n(t)\tilde{X}_n(s))^2 \leq \frac{1}{n^2} \left\{ nEX_1^2(t)X_1^2(s) + 3n^2EX_1^4(t)EX_1^2(s) \right\}
\] (3.2)

and
\[
\begin{align*}
\iint EX_1^2(t)X_1^2(s)dtds &= E \left( \int X_1^2(t)dt \right)^2 = E\|X_1\|^4, \\
\iint EX_1^2(t)EX_1^2(s)dtds &= \left( E\|X_1\|^2 \right)^2 \leq E\|X_1\|^4,
\end{align*}
\] (3.3)

resulting in
\[
E \iint (\tilde{X}_n(t)\tilde{X}_n(t))^2dtds \leq \frac{4}{n^2}E\|X_1\|^4.
\] (3.4)

Repeating the arguments leading to (3.4) we obtain that
\[
\begin{align*}
E(\tilde{c}_{n,h}^{(2)}(t, s)) &\leq \frac{4}{n^2} (EX_1(t)^2X_1^2(s) + EX_1^2(t)EX_1^2(s)), \\
E(\tilde{c}_{n,h}^{(3)}(t, s)) &\leq \frac{4}{n^2} (EX_1(t)^2X_1^2(s) + EX_1^2(t)EX_1^2(s))
\end{align*}
\] (3.5) (3.6)

and therefore
\[
E \iint a_{n,h}^2(t, s)dtds \leq \frac{C_1}{n^2}E\|X_1\|^4
\] (3.7)

with some constant \( C_1 \). Elementary arguments give
\[
E\tilde{c}_{n,h}^2(t, s) \leq \frac{1}{n}EX_1^2(t)EX_1^2(s).
\] (3.8)

By the Cauchy–Schwarz inequality for expected values, (3.5), (3.6) and (3.8) we have with some constant \( C_2 \) that
\[
E|\tilde{c}_{n,h}(t, s)a_{n,h}(t, s)| \leq (E\tilde{c}_{n,h}^2(t, s)EX_{n,h}^2(t, s))^{1/2} \leq C_2n^{1/2}(EX_1^2(t)EX_1^2(s))^{1/2}(EX_1^2(t)X_1^2(s))^{1/2} + (EX_1^2(t)EX_1^2(s))^{1/2}
\]

and by the Cauchy–Schwarz inequality for integrals and then for expected values we obtain that
\[
\begin{align*}
\iint (EX_1^2(t)EX_1^2(s))^{1/2}(EX_1^2(t)X_1^2(s))^{1/2}dtds &\leq \left\{ \iint EX_1^2(t)EX_1^2(s)dtds \iint EX_1^2(t)X_1^2(s)dtds \right\}^{1/2} \\
&= \int EX_1^2(t)dt \left\{ E \left( \int X_1^2(t)dt \right)^2 \right\}^{1/2} \\
&= E\|X_1\|^2(E\|X_1\|^4)^{1/2} \leq E\|X_1\|^4.
\end{align*}
\]
where the last line is obtained by Hölder's inequality. Hence by (3.3) we conclude
\[ \int \int E|\tilde{C}_{n,h}(t, s)\alpha_{n,h}(t, s)| \leq \frac{2c_2}{n^{\gamma/2}} E\|X_1\|^4. \]

Thus we showed that
\[ E \left| \tilde{V}_{n,H} - \bar{V}_{n,H} \right| \leq c_3 \frac{1}{(nH)^{1/2}}, \]
and therefore (3.1) follows from Assumption 1.1 via Markov's inequality.

Next we prove that
\[ \frac{n}{H^{1/2}} \left| \bar{V}_{n,H} - V_{n,H} \right| = o_P(1), \quad \text{as } n \to \infty. \]

(3.9)

As in the proof of (3.1) we have
\[ \left| \tilde{C}_{n,h}(t, s) - C_{n,h}(t, s) \right| \leq 2|C_{n,h}(t, s)\alpha_{n,h}(t, s)| + \tilde{a}_{n,h}^2(t, s) \]
with
\[ \tilde{a}_{n,h}(t, s) = \frac{1}{n} \sum_{i=n-h+1}^{n} X_i(t)X_i(t+h(s)). \]

It is easy to see that
\[ E\tilde{a}_{n,h}^2(t, s) \leq \frac{h}{n^2} EX_1^2(t)X_1^2(s) \]
and
\[ EC_{n,h}^2(t, s) \leq \frac{1}{n} EX_1^2(t)X_1^2(s). \]

Thus we get
\[ E \left| \tilde{V}_{n,H} - V_{n,H} \right| \leq c_1 \sum_{h=1}^{H} \frac{H^{1/2}}{n^{3/2}}, \]
so Assumption 1.1 and Markov's inequality imply (3.9). Lemma 3.1 is an immediate consequence of (3.1) and (3.9).

Next we note that for all \( i = 1, 2, \ldots \)
\[ E\xi_{i,\ell} = 0, \quad E\xi_{i,\ell}^2 = \lambda_\ell \quad \text{and} \quad E\xi_{i,\ell} \xi_{i,k} = 0 \quad \text{if } k \neq \ell, \]
where \( \xi_{i,\ell} \) is defined in (1.4). Using (1.4) we obtain that
\[ \int \int C_{n,h}^2(t, s)dt ds = \frac{1}{n^2} \sum_{i,j=1}^{n} \sum_{\ell,k=1}^{\infty} \xi_{i,\ell} \xi_{i+h,k} \xi_{i,\ell} \xi_{i+h,k}. \]

Let \( T > 1 \) be an integer and define
\[ Q_{n}(T) = \frac{1}{H^{1/2}} \sum_{h=1}^{H} \sum_{i=1}^{n} \sum_{j=1}^{T} (\xi_{i,\ell} \xi_{i+h,k} \xi_{j,\ell} \xi_{j+h,k} - E\xi_{i,\ell} \xi_{i+h,k} \xi_{j,\ell} \xi_{j+h,k}) \]
\[ Q_{n,1}(T) = \frac{1}{H^{1/2}} \sum_{h=1}^{H} \sum_{i=1}^{n} \sum_{j=1}^{T} \sum_{\ell=1}^{\infty} (\xi_{i,\ell} \xi_{i+h,k} \xi_{j,\ell} \xi_{j+h,k} - E\xi_{i,\ell} \xi_{i+h,k} \xi_{j,\ell} \xi_{j+h,k}) \]
\[ Q_{n,2}(T) = \frac{1}{H^{1/2}} \sum_{h=1}^{H} \sum_{i=1}^{n} \sum_{j=1}^{T} \sum_{k=1}^{\infty} (\xi_{i,\ell} \xi_{i+h,k} \xi_{j,\ell} \xi_{j+h,k} - E\xi_{i,\ell} \xi_{i+h,k} \xi_{j,\ell} \xi_{j+h,k}) \]
\[ Q_{n,3}(T) = \frac{1}{H^{1/2}} \sum_{h=1}^{H} \sum_{i=1}^{n} \sum_{j=1}^{T} \sum_{\ell=1}^{\infty} (\xi_{i,\ell} \xi_{i+h,k} \xi_{j,\ell} \xi_{j+h,k} - E\xi_{i,\ell} \xi_{i+h,k} \xi_{j,\ell} \xi_{j+h,k}). \]

If there is a \( T_0 \) such that \( \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{T_0} > \lambda_{T_0+1} = 0 \), then \( X_i(t) \) in (1.4) is a finite sum and \( Q_{n,i}(T) = 0 \) for all \( i = 1, 2, 3 \). So in this case Lemma 3.2 automatically holds, and the proofs are mathematically much simpler. Hence from now on we assume that \( \lambda_j > 0 \) for all \( j = 1, 2, \ldots \).
Lemma 3.2. If $H_0$, Assumptions 1.1 and 1.2 hold, then for every $\varepsilon > 0$

$$\lim_{T \to \infty} \lim_{n \to \infty} \sup \left\{ \left| Q_{n,i}(T) \right| > \varepsilon \right\} = 0, \quad i = 1, 2, 3.$$  \hspace{1cm} (3.11)

Proof. It follows from (3.10) that

$$Q_{n,1}(T) = \frac{1}{H^{1/2}n} \sum_{h=1}^{H} \sum_{\ell=1}^{T} \sum_{k=1}^{n} \left( \sum_{i=1}^{n} \xi_{i,\ell}^2 \xi_{i+h,k}^2 \lambda_{i,k} - \lambda_{i,k} \right) + \frac{2}{H^{1/2}n} \sum_{h=1}^{H} \sum_{\ell=1}^{T} \sum_{k=1}^{H} \sum_{1 \leq i, j \leq n} \xi_{i,\ell} \xi_{i+h,k} \xi_{j,\ell} \xi_{j+h,k}$$

$$= Q_{n,1,1}(T) + 2Q_{n,1,2}(T).$$

Introducing

$$\zeta_{i+h} = \xi_{i+h} + \sum_{k=1}^{T} \xi_{i+h,k}^2, \quad \eta_{i} = \eta_{i}(T) = \sum_{\ell=1}^{T} \xi_{i,\ell}^2 \quad \text{and} \quad w_{i,h} = \eta_{i} \xi_{i+h} - E\eta_{i} \xi_{i+h},$$

we can write

$$Q_{n,1,1}(T) = \frac{1}{H^{1/2}n} \sum_{h=1}^{H} \sum_{i=1}^{n} w_{i,h}.$$ 

Clearly, $E Q_{n,1,1}(T) = 0$ and

$$E Q_{n,1,1}^2(T) = \frac{1}{H n^2} \sum_{i,j=1}^{H} \sum_{h,k} E w_{i,h} w_{j,k}.$$ 

Using the independence of the $\chi_i$'s we observe that $E w_{i,h} w_{j,k} = 0$, except in the following cases:

(i) $i \neq j$ and $i = j + k$: The number of such terms is bounded by $n^2 H$ and $|E w_{i,h} w_{j,k}| \leq \left( |E (\eta_{i} \xi_{i}) \eta_{j} \xi_{j}^2 + E (\eta_{i} \xi_{i})^2 \right)$

(ii) $i \neq j$ and $j = k + h$: As in (i) the number of such terms is less than $n^2 H$ and $|E w_{i,h} w_{j,k}| \leq \left( |E (\eta_{i} \xi_{j}) \eta_{j} \xi_{j}^2 + E (\eta_{j} \xi_{j})^2 \right)$

(iii) $i \neq j$ and $i + h = j + k$: The number of such terms is again less than $n^2 H$ and $|E w_{i,h} w_{j,k}| \leq \left( |E (\eta_{i} \xi_{i}) \eta_{j} \xi_{j}^2 + E (\eta_{j} \xi_{j})^2 \right)$

(iv) $i = j$: The number of such terms is less than $n H^2$ and $|E w_{i,h} w_{j,k}| \leq \left( |E (\eta_{i} \xi_{i}) \xi_{j}^2 + E (\eta_{j} \xi_{j})^2 \right)^2$.

Thus we get with some constant $c_1$

$$E Q_{n,1,1}^2(T) \leq c_1 Q^*(T), \hspace{1cm} (3.12)$$

where

$$Q^*(T) = |E (\eta_{i} \xi_{i}) \eta_{j} \xi_{j}^2 + E (\eta_{j} \xi_{j})^2 \xi_{j}^2 + E (\eta_{j} \xi_{j})^2 \xi_{j}^2 + E (\eta_{j} \xi_{j})^2 \xi_{j}^2.$$ 

With

$$\tilde{\xi}_{i,j,h} = \xi_{i,j,h}(T) = \sum_{k=1}^{T} \xi_{i+h,k} \xi_{j+h,k} \quad \text{and} \quad \tilde{\eta}_{i,j} = \eta_{i,j}(T) = \sum_{\ell=1}^{T} \xi_{i,\ell} \xi_{j,\ell}$$

we can write

$$Q_{n,1,2}(T) = \frac{1}{H^{1/2}n} \sum_{h=1}^{H} \sum_{1 \leq i \leq j \leq n} \tilde{\xi}_{i,j,h} \tilde{\eta}_{i,j}.$$ 

By the independence of the $\chi_i$'s we have $E Q_{n,1,2}(T) = 0$ and

$$E Q_{n,1,2}^2(T) = \frac{1}{H n^2} \sum_{h,k=1}^{H} \sum_{1 \leq i \leq j \leq n} \sum_{1 \leq i \leq j \leq n} E[\tilde{\xi}_{i,j,h} \tilde{\eta}_{i,j} \tilde{\xi}_{i,j,k} \tilde{\eta}_{i,j,k}].$$

Due to the independence of the $\chi_i$'s, $E[\tilde{\xi}_{i,j,h} \tilde{\eta}_{i,j} \tilde{\xi}_{i,j,k} \tilde{\eta}_{i,j,k}] = 0$ except in the case when $i_1 = i_2, j_1 = j_2$ and $h = k$. In this case

$$E[\tilde{\xi}_{i,j,h} \tilde{\eta}_{i,j} \tilde{\xi}_{i,j,k} \tilde{\eta}_{i,j,k}] = E \tilde{\xi}_{i,j,h}^2 E \tilde{\eta}_{i,j}^2.$$ 

Thus we have that

$$E Q_{n,1,2}^2(T) \leq E \tilde{\xi}_{i,j,h}^2 E \tilde{\eta}_{i,j}^2.$$
Next we observe that by (3.10)
\[ E\overline{\xi}_{1,2,0}^2 = \sum_{k_1,k_2 = T+1}^{\infty} E\xi_{1,k_1}\xi_{2,k_2} = \left( \sum_{k = T+1}^{\infty} E\xi_{1,k}^2 \right)^2 = \left( \sum_{k = T+1}^{\infty} \lambda_i \right)^2 \]
and similarly
\[ E\overline{\eta}_{1,2}^2 = \left( \sum_{k = 1}^{T} \lambda_i \right)^2. \]

It follows that
\[ \lim_{T \to \infty} \limsup_{n \to \infty} EQ^2_{n,1,2}(T) = 0. \tag{3.13} \]

Using (1.4) and (3.10) we get
\[
E\|X\|^4 = E \left( \int \left( \sum_{\ell = 1}^{\infty} \xi_{\ell,\ell}\varphi(t) \right)^2 dt \right)^2
= E \left( \sum_{\ell = 1}^{\infty} \xi_{\ell,\ell}^2 \right)^2
= \sum_{\ell = 1}^{\infty} E\xi_{\ell,\ell}^4 + \sum_{1 \leq \ell \neq k < \infty} \lambda_{\ell}\lambda_{k}, \tag{3.14}
\]
and therefore Assumption 1.2 yields that
\[ \sum_{\ell = 1}^{\infty} E\xi_{\ell,\ell}^4 < \infty. \tag{3.15} \]

Using the definition of \( Q^*(T) \) and (3.15) we conclude that \( Q^*(T) \to 0 \), as \( T \to \infty \), and therefore by (3.12)
\[ \lim_{T \to \infty} \limsup_{n \to \infty} EQ^2_{n,1,1}(T) = 0. \tag{3.16} \]

Now (3.11) follows from (3.13) and (3.16) via Chebyshev’s inequality if \( i = 1 \). The other two cases can be proved in the same way, so the details are omitted. \( \square \)

To study the asymptotic behavior of \( Q_n(T) \) we note that
\[ Q_n(T) = \frac{1}{H^{1/2}} \sum_{h = 1}^{H} \sum_{\ell,k = 1}^{T} \left( \frac{1}{n^{1/2}} \sum_{i = 1}^{n} \xi_{i,\ell}\xi_{i+h,k} \right)^2 - \lambda_{\ell}\lambda_{k}. \]

The goal of the next lemmas is to provide an approximation with rates for the vector of partial sums \( n^{-1/2} \sum_{i = 1}^{n} \xi_{i,\ell}\xi_{i+h,k} \), \( 1 \leq \ell, k \leq T, 1 \leq h \leq H \). Next we introduce the standardized variables
\[ \chi_{i,h,\ell,k} = \frac{\xi_{i,\ell}\xi_{i+h,k}}{(\lambda_{\ell}\lambda_{k})^{1/2}} \]
and the column vector \( \chi_i \in \mathbb{R}^d \), \( d = HT^2 \), whose coordinates are \( \chi_{i,h,\ell,k} \) ordered lexicographically in \( h, \ell, k \).

**Lemma 3.3.** If \( H_0 \), Assumptions 1.1 and 1.2 hold, then
\[ E\chi_i = 0, \tag{3.17} \]
\[ E\chi_i\chi_i^T = I_d, \tag{3.18} \]
where \( I_d \) is the \( d \times d \) identity matrix,
\[ E\left( \sum_{i = 1}^{m} \chi_i \right)^T = ml_d, \tag{3.19} \]
and for all $m \geq H$

$$
E \left( \sum_{i=1}^{m} \chi_{i,h,c,k} \right)^4 \leq \frac{c}{\lambda_f^2} H^2 m^2,
$$

(3.20)

where $c$ only depends on $E\|X\|^4$.

**Proof.** The independence of $\xi_{i,c}$ and $\xi_{i+h,c}$ gives (3.17), since $E\xi_{i,c} = 0$. It follows from (3.10) and independence that $E\chi_{i,h,c,k}^4 = 1$ and the coordinates of $\chi_i$ are uncorrelated, proving (3.18). Also, $X_1, X_2, \ldots$ are uncorrelated random vectors, (3.18) implies (3.19). Let

$$
Q(r) = Q(r; h) = \{i : 1 \leq i \leq m, \text{ mod}(i, h+1) = r\}, \quad r = 0, 1, \ldots, h
$$

and decompose

$$
\sum_{i=1}^{m} \chi_{i,h,c,k} = \sum_{r=0}^{h} \sum_{i \in Q(r)} \chi_{i,h,c,k}.
$$

Since for every $r$, $\sum_{i \in Q(r)} \chi_{i,h,c,k}$ is a sum of independent and identically distributed random variables, the Rosenthal inequality (cf. [35, p. 59]) yields

$$
E \left( \sum_{i \in Q(r)} \chi_{i,h,c,k} \right)^4 \leq c_1 \left\{ \frac{m}{h+1} E\chi_{i,h,c,k}^4 + \left( \frac{m}{h+1} \right)^2 \right\}.
$$

Also, by the independence of the $X_i$’s we obtain from (3.14) that

$$
E\chi_{i,h,c,k}^4 = \frac{1}{\lambda_f^2}\lambda_k^2 E\xi_{i,c}^4 E\xi_{i,k}^4 \leq \frac{(E\|X\|^4)^2}{\lambda_f^2 \lambda_k^2},
$$

and therefore

$$
E \left( \sum_{i \in Q(r)} \chi_{i,h,c,k} \right)^4 \leq c_2 \frac{1}{\lambda_f^2} \frac{m^2}{(h+1)^2}.
$$

By Hölder’s inequality

$$
E \left( \sum_{i=1}^{m} \chi_{i,h,c,k} \right)^4 \leq (h+1)^3 \sum_{r=0}^{h} E \left( \sum_{i \in Q(r)} \chi_{i,h,c,k} \right)^4 \leq c_3 \frac{1}{\lambda_f^2} h^2 m^2,
$$

and therefore we obtain immediately (3.20). □

Let $|x|$ denote the Euclidean norm of vectors and matrices.

**Lemma 3.4.** Let $\delta_1, \delta_2, \ldots, \delta_m$ be independent and identically distributed random vectors in $\mathbb{R}^d$ with $E\delta_1 = 0, E\delta_1^T = I_d$ and $E|\delta_1|^3 < \infty$. Then for all $m$ we can define $y_m$, a standard normal vector in $\mathbb{R}^d$ such that

$$
P \left\{ \left| m^{-1/2} \sum_{i=1}^{m} \delta_i - y_m \right| \geq cm^{-1/8} d^{1/4} (E|\delta_1|^3 + E|y_m|^3)^{1/4} \right\} \leq cm^{-1/8} d^{1/4} (E|\delta_1|^3 + E|y_m|^3)^{1/4},
$$

where $c$ is an absolute constant.

**Proof.** The result is an immediate consequence of Theorem 6.4.1 in [37, p. 207] and the corollary to Theorem 11 in [38]. For further results on the dependence of the rate of convergence of the dimension of the summands in the central limit theorem we refer to Bentkus [3]. □

We note that

$$
(E|\delta_1|^3 + E|y_m|^3)^{1/4} \leq (E|\delta_1|^3)^{1/4} + (E|y_m|^3)^{1/4} \leq (E|\delta_1|^4)^{3/16} + (E|y_m|^4)^{3/16}
$$

(3.21)

and since $|y_m|^2$ has a $\chi^2$ distribution with $d$ degrees of freedom,

$$
E|y_m|^4 \leq c_1 d^2.
$$

(3.22)
Lemma 3.5. If $H_0$, Assumptions 1.1 and 1.2 hold, then for every $n$ we can define a standard normal vector $\mathbf{y}_n \in \mathbb{R}^d$ such that

$$
P \left\{ n^{-1/2} \left| \sum_{i=1}^{n} X_i - \mathbf{y}_n \right| \geq c n^{-2/21} H^{11/8} T / \lambda_T^{3/4} \right\} \leq c n^{-2/21} H^{11/8} T / \lambda_T^{3/4},$$

where $c$ only depends on $E[\|X_1\|^4]$.

**Proof.** Since $X_1, \ldots, X_n$ are not independent random variables, Lemma 3.4 cannot be used immediately. However, the sequence $X_i$, $i \geq 1$, is $H$-dependent. Let $N > H$ and define the vectors

$$
\delta_i = \sum_{j \in U(i)} X_j \quad \text{and} \quad \delta_{i,H} = \sum_{j \in J(i)} X_j,
$$

where

$$
U(i) = \{ j : 1 \leq j \leq n, (N + H)(i - 1) + 1 \leq j \leq (N + H)(i + 1) \}
$$

and

$$
J(i) = \{ j : 1 \leq j \leq n, (N + H)(i - 1) + N + 1 \leq j \leq (N + H)i \}.
$$

It is easy to see that the number of random vectors defined in (3.23) is not more then $K + 1$, where $K = \lfloor n/(N + H) \rfloor$. Also, $\delta_i, 1 \leq i \leq K + 1$ are independent and $\delta_i, 1 \leq i \leq K$ are identically distributed. Similarly, $\delta_{i,H}, 1 \leq i \leq K + 1$ are independent and, with the possible exception of $\delta_{K+1,H}$, are identically distributed. Clearly,

$$
\sum_{i=1}^{n} X_i = \sum_{i=1}^{K} \delta_i + \sum_{i=1}^{K+1} \delta_{i,H} + \delta_{K+1}.
$$

For every $x > 0$ we have by Markov’s inequality that

$$
P \left\{ n^{-1/2} \left| \sum_{i=1}^{K+1} \delta_{i,H} + \delta_{K+1} \right| > x \right\} \leq \frac{1}{x^4 H^2} \left( \sum_{i=1}^{K+1} \delta_{i,H} + \delta_{K+1} \right)^4 \leq \frac{2^4}{x^4 H^2} \left( \sum_{i=1}^{K+1} \delta_{i,H} \right)^4 + E |\delta_{K+1}|^4.
$$

It follows from (3.20) that

$$
E |\delta_{K+1}|^4 \leq c_1 d^4 H^2 N^2 / \lambda_T^4.
$$

We use the notation $\delta_{i,H}(\ell)$ for the $\ell$th coordinate of $\delta_{i,H}$. It is easy to see that

$$
E \left( \sum_{i=1}^{K+1} \delta_{i,H}(\ell) \right)^4 \leq d^4 \left( \sum_{i=1}^{K+1} \delta_{i,H}(j) \right)^4 \leq \left( \sum_{i=1}^{K+1} \delta_{i,H}(j) \right)^4,
$$

Using Rosenthal’s inequality (cf. [35, p. 59]) with Lemma 3.3 we obtain that

$$
E \left( \sum_{i=1}^{K+1} \delta_{i,H}(j) \right)^4 \leq c_2 \left( KH^4 / \lambda_T^4 + (KH)^2 \right) \leq c_2 K^2 H^4 / \lambda_T^4.
$$

Hence (3.25) and (3.26) imply

$$
P \left\{ n^{-1/2} \left| \sum_{i=1}^{K+1} \delta_{i,H} + \delta_{K+1} \right| > x \right\} \leq \frac{c_1}{x^4 H^2} \frac{d^4}{\lambda_T^4} \left( H^2 N^2 + K^2 H^4 \right).
$$

The vectors $N^{-1/2} \delta_i \in \mathbb{R}^d$, $1 \leq i \leq K$ satisfy all the conditions of Lemma 3.4, since they are independent and identically distributed and according to Lemma 3.3, $E N^{-1/2} \delta_i = 0$, $E (N^{-1/2} \delta_i)(N^{-1/2} \delta_i)^T = I$. So in light of (3.22) we only need to compute $E |N^{-1/2} \delta_i|^4$. It follows from (3.20) that

$$
E |N^{-1/2} \delta_i|^4 \leq c_4 d^4 H^2 / \lambda_T^4,
$$

and therefore for any standard normal random vector $\mathbf{y}_m \in \mathbb{R}^d$

$$
(E |N^{-1/2} \delta_i|^4)^{3/16} + (E |\mathbf{y}_m|^4)^{3/16} \leq c_5 d^{12/16} H^{3/8} / \lambda_T^{3/4}.
$$
Applying Lemma 3.4, we can find a standard normal random vector $\tilde{y}_K \in \mathbb{R}^d$ such that
\[
P \left\{ K^{-1/2} \sum_{i=1}^T N^{-1/2} \delta_i - \tilde{y}_K \geq c_6 K^{-1/8} d H^{3/8} \lambda_T^{3/4} / \sqrt{T} \right\} \leq c_6 K^{-1/8} d H^{3/8} \lambda_T^{3/4}. \tag{3.28}
\]
Let $N = \lfloor n^{5/21} \rfloor$ in the definitions of $\delta_i$ and $\delta_i H$. By Assumption 1.1, $K \approx n^{16/21}$. So using (3.24), (3.27) with $x = n^{-2/21} d H^{3/8} / \lambda_T^{3/4}$ and (3.28), we conclude
\[
P \left\{ n^{-1/2} \sum_{i=1}^n X_i - \left( \frac{NK}{n} \right)^{1/2} \tilde{y}_K \geq c_7 n^{-2/21} d H^{3/8} / \lambda_T^{3/4} \right\} \leq c_7 d^{-1/2} \log n \leq c_8 d / n^2.
\]
Elementary arguments show that
\[
\left| \left( \frac{NK}{n} \right)^{1/2} \right| = O(H n^{-5/21})
\]
and since $|\tilde{y}_K|$ is a $\chi^2$ random variable with $d$ degrees of freedom, there is a constant $c_8$ such that
\[
P \{|\tilde{y}_K| > c_8 d^{1/2} \log n\} \leq c_8 d / n^2.
\]
Hence the lemma follows from (3.29) with the choice of $y_n = \tilde{y}_K$ and using $d = HT^2$. □

**Lemma 3.6.** If $H_0$, Assumptions 1.1 and 1.2 hold, then for every $n$ we can define independent standard normal random variables $\gamma_n(h; \ell, k), 1 \leq h \leq H, 1 \leq k, \ell \leq T$ such that for every $T > 0$
\[
|Q_n(T) - Z^*_n(T)| = o_p(1), \quad \text{as } n \to \infty,
\]
where
\[
Z^*_n(T) = \frac{1}{H^{1/2}} \sum_{h=1}^H \sum_{\ell, k=1}^T (\gamma^2_n(h; \ell, k) - 1) \lambda_\ell \lambda_k.
\]

**Proof.** The result is an immediate consequence of Lemma 3.5 and Assumption 1.1. □

Let $\gamma(h; \ell, k), 1 \leq k, \ell, h < \infty$ be independent standard normal random variables and define
\[
\tilde{Z}_n(T) = \frac{1}{H^{1/2}} \sum_{h=1}^H \sum_{\ell, k=1}^T (\gamma^2(h; \ell, k) - 1) \lambda_\ell \lambda_k
\]
and
\[
Z_n = \frac{1}{H^{1/2}} \sum_{h=1}^H \sum_{\ell, k=1}^\infty (\gamma^2(h; \ell, k) - 1) \lambda_\ell \lambda_k.
\]
(We note that Assumption 1.1 implies that $Z_n$ is a well defined random variable.) It is clear that for every $n$ and $T$
\[
Z^*_n(T) \overset{D}{=} \tilde{Z}_n(T).
\]

**Lemma 3.7.** If $H_0$, Assumptions 1.1 and 1.2 hold, then for every $\varepsilon > 0$ we have
\[
\lim_{T \to \infty} \limsup_{n \to \infty} P \{|\tilde{Z}_n(T) - Z_n| > \varepsilon\} = 0.
\]

**Proof.** The result can be established along the lines of the proof of Lemma 3.2, but the arguments can be simplified since $\tilde{Z}_n(T)$ as well as $Z_n$ are weighted sums of independent $\chi^2$ random variables with one degree of freedom. We omit the details. □

**Lemma 3.8.** If $H_0$, Assumptions 1.1 and 1.2 hold, then
\[
\frac{Z_n}{(2\sigma^2)^{1/2}} \overset{D}{\to} N(0, 1),
\]
where $N(0, 1)$ stands for a standard normal random variable and $\sigma^2$ is defined in Theorem 1.1.
Proof. The variables $\sum_{i,k=1}^{\infty} (\gamma^2(h; \ell, k) - 1) \lambda_i \lambda_k$, $1 \leq h \lesssim H$, are independent and identically distributed with zero mean and variance $2\sigma^2$, so the result follows immediately from the central limit theorem. □

Proof of Theorem 1.2. The proof is done in several steps. It follows from Lemma 3.1 that it is enough to consider $V_{n,H}$. Next we use the Karhunen–Loéve expansion of (1.4) and we express $V_{n,H}$ as an infinite sum of the projection coefficients $\xi_i = \langle X_i, \psi_i \rangle$. Lemma 3.2 proves that it is sufficient to establish the normality of $Q_n(T)$, i.e. the infinite series representation of $V_{n,H}$ we need to keep finitely many terms. Lemmas 3.6 and 3.7 approximate $Q_n(T)$ with a weighted average of independent $\chi^2$ random variables with one degree of freedom. The final step of the proof is the central limit theorem given in Lemma 3.8. □

Proof of Theorem 1.3. It follows from the definition of $C_{n,0}$ that
\[
\int \hat{C}_{n,0}(t, t) dt = \frac{1}{n} \sum_{i=1}^{n} \int X_i^2(t) dt - \int \bar{X}_n^2(t) dt.
\]
Since by Assumption 1.2 we have that $E(\int X_i^2(t) dt)^2 < \infty$, the central limit theorems yields that
\[
\left| \int \hat{C}_{n,0}(t, t) dt - \int C_0(t, t) dt \right| = O_P(n^{-1/2}).
\]
Elementary calculations give
\[
E \int \bar{X}_n^2(t) dt = \frac{1}{n} \int C_0(t, t) dt,
\]
so by Markov’s inequality we have
\[
\int \bar{X}_n^2(t) dt = O_P(1/n).
\]
Thus we get immediately that
\[
|\hat{\mu}_n - \mu| = O_P(n^{-1/2}).
\]
Next we observe that $X_i(t)X_i(s), i \geq 1$ is a sequence of independent and identically distributed random variables in $L^2([0, 1]^2)$ and by Assumption 1.2
\[
E \iint (\hat{C}_{n,0}(t, s) - C_0(t, s))^2 dtds = O(1/n).
\]
Hence
\[
\hat{\sigma}_n^2 \xrightarrow{P} \sigma^2
\]
follows from Markov’s inequality. By Assumption 1.1 we have that $H/n \to 0$, so Theorem 1.1 implies the result. □

4. Proof of Theorem 1.3 and the outline of the proof of Remark 1.2

Proof of Theorem 1.3. We can assume that $H \geq h_0$ and clearly
\[
\frac{1}{(2H\hat{\sigma}_n^2)^{1/2}} \left\{ n\hat{Y}_{n,H} - H\hat{\mu}_n \right\} \geq \frac{1}{(2H\hat{\sigma}_n^2)^{1/2}} \left\{ n \iint \hat{C}_{n,h_0}^2 dtds - H\hat{\mu}_n \right\}.
\]
We can assume without loss of generality that $EX_1(t) = 0$. Elementary arguments yield
\[
\hat{\mu}_n = \frac{1}{n} \sum_{i=1}^{n} \int X_i^2(t) dt - \left( \int \bar{X}_n^2(t) dt \right)^2.
\]
Under the assumptions of Theorem 1.3, $\int X_i^2(t) dt$ is a stationary and ergodic sequence, so by the ergodic theorem (cf. [8, p. 118])
\[
\frac{1}{n} \sum_{i=1}^{n} \int X_i^2(t) dt \to \int EX_i^2(t) dt \quad \text{a.s.}
\]
Next we use Theorem A.1 with $Y_i(t) = X_i(t)$ to conclude
\[
\int \bar{X}_n^2(t) dt \to 0 \quad \text{a.s.} \quad (4.1)
\]
and therefore
\[ \hat{\mu}_n \to \mu \quad \text{a.s.} \] (4.2)

Similarly,
\[ \hat{\sigma}^2_n = \iint \left( \frac{1}{n} \sum_{i=1}^{n} X_i(t)X_i(s) \right)^2 dt ds - 2 \iint \left( \frac{1}{n} \sum_{i=1}^{n} X_i(t)X_i(s) \right) \bar{X}_n(t)\bar{X}_n(s) dt ds + \iint (\bar{X}_n(t)\bar{X}_n(s))^2 dt ds. \]

Next we use Theorem A.1 with \( Y_i(t, s) = X_i(t)X_i(s) \) to conclude
\[ \iint \left( \frac{1}{n} \sum_{i=1}^{n} X_i(t)X_i(s) \right)^2 dt ds \to \iint (EX_i(t)X_i(s))^2 dt ds \quad \text{a.s.} \]

Hence the Cauchy–Schwarz inequality via (4.1) implies
\[ \left| \iint \left( \frac{1}{n} \sum_{i=1}^{n} X_i(t)X_i(s) \right) \bar{X}_n(t)\bar{X}_n(s) dt ds \right| \leq \left( \iint \left( \frac{1}{n} \sum_{i=1}^{n} X_i(t)X_i(s) \right)^2 dt ds \right)^{1/2} \int \bar{X}_n^2(t) dt \to 0 \quad \text{a.s.} \]

Hence we get
\[ \hat{\sigma}^2_n \to \sigma^2 \quad \text{a.s.} \] (4.3)

Repeating the arguments leading to (4.3), one can easily verify that
\[ \iint \bar{C}_{n,h}^2 dt ds \to \iint C_{h}^2 dt ds > 0 \quad \text{a.s.} \]

Now Theorem 1.3 follows from Assumption 1.4. \( \square \)

Outline of the proof of Remark 1.2. One can verify that \( \bar{C}_{n,h}(t, s) \) can be approximated with
\[ \bar{C}_{n,h}(t, s) = \frac{1}{n} \sum_{i=1}^{n} Y_i(t)Y_{i+h}(s) + \frac{\gamma_n}{n} \sum_{i=1}^{n} Y_i(t)Z_{i+h}(s) + \frac{\gamma_n}{n} \sum_{i=1}^{n} Z_i(t)Y_{i+h}(s) + \frac{\gamma_n}{n} \sum_{i=1}^{n} Z_i(t)Z_{i+h}(s). \]

So
\[ n\bar{\nu}_{n,H} = \sum_{h=1}^{H} \iint \left( \frac{1}{n^{1/2}} \sum_{i=1}^{n} Y_i(t)Y_{i+h}(s) + \frac{\gamma_n}{n^{1/2}} \sum_{i=1}^{n} Y_i(t)Z_{i+h}(s) \right. \]
\[ + \left. \frac{\gamma_n}{n} \sum_{i=1}^{n} Z_i(t)Y_{i+h}(s) + \frac{\gamma_n^2 n^{1/2}}{1} \frac{1}{n} \sum_{i=1}^{n} Z_i(t)Z_{i+h}(s) \right)^2 dt ds \]
\[ = \sum_{h=1}^{H} \iint \left( \frac{1}{n^{1/2}} \sum_{i=1}^{n} Y_i(t)Y_{i+h}(s) + \gamma_n^2 n^{1/2} \frac{1}{n} \sum_{i=1}^{n} Z_i(t)Z_{i+h}(s) \right)^2 dt ds + o_p(H^{1/2}) \]

since \( H^{1/2} \gamma_n \to 0 \). For every \( h \{Z_{i+h}\} \) satisfies the central limit theorem in \( L^2 \) according to Berkes et al. [5], thus
\[ \left| \frac{1}{n} \sum_{i=1}^{n} Z_i(t)Z_{i+h}(s) - A_h(t, s) \right| = O_p(n^{-1/2}), \]

where the square integrable function \( A_h(t, s) \) is defined in (1.7). By the Cauchy–Schwarz inequality we have
\[ \sum_{h=1}^{H} \iint \left( \frac{1}{n^{1/2}} \sum_{i=1}^{n} Y_i(t)Y_{i+h}(s) + \gamma_n^2 n^{1/2} \frac{1}{n} \sum_{i=1}^{n} Z_i(t)Z_{i+h}(s) \right)^2 dt ds \]
\[ = \sum_{h=1}^{H} \iint \left( \frac{1}{n^{1/2}} \sum_{i=1}^{n} Y_i(t)Y_{i+h}(s) + \gamma_n^2 n^{1/2} A_h(t, s) \right)^2 dt ds + O_p(\gamma_n^4 H n^{1/2}). \]

Due to the \( (4 + \delta)\)-decomposability of \( \{Z_i\} \)
\[ \sum_{h=1}^{\infty} \left( \int A_h^2(t, s) dt ds \right)^{1/2} < \infty \]
holds, therefore
\[
\begin{align*}
\sum_{h=1}^{H} & \iint \left( \frac{1}{n^{1/2}} \sum_{i=1}^{n} Y_i(t) Y_{i+h}(s) + \gamma_n^2 n^{1/2} A_n(t, s) \right)^2 dtds = \sum_{h=1}^{H} \iint \left( \frac{1}{n^{1/2}} \sum_{i=1}^{n} Y_i(t) Y_{i+h}(s) \right)^2 dtds \\
& + 2\gamma_n^2 n^{1/2} \sum_{h=1}^{H} \iint \left( \frac{1}{n^{1/2}} \sum_{i=1}^{n} Y_i(t) Y_{i+h}(s) \right) A_n(t, s) dtds + \gamma_n^2 n \sum_{h=1}^{H} \iint A_n^2(t, s) dtds \\
& = \sum_{h=1}^{H} \iint \left( \frac{1}{n^{1/2}} \sum_{i=1}^{n} Y_i(t) Y_{i+h}(s) \right)^2 dtds + \gamma_n^2 n \sum_{h=1}^{H} \iint A_n^2(t, s) dtds + O_p(\gamma_n^2 n^{1/2}).
\end{align*}
\]

The remark follows from Theorem 1.1. □

Appendix. Ergodic theorem in Hilbert spaces

In this section we assume \( Y_1(t), Y_2(t), \ldots \) are random functions defined on a measurable set \( \mathcal{C} \subseteq \mathbb{R}^d \) satisfying the following conditions:

Assumption A.1. \( Y_1, Y_2, \ldots \) is a stationary and ergodic sequence of square integrable random functions and

Assumption A.2. \( E\|Y_1\| < \infty \).

Theorem A.1. If Assumptions A.1 and A.2 hold, then, as \( n \to \infty \),

\[
\left\| \frac{1}{n} \sum_{i=1}^{n} Y_i \right\| \to E\|Y_1\| \quad \text{a.s.}
\]

Proof. Since \( \left\| (1/n) \sum_{i=1}^{n} Y_i - EY_1 \right\| \leq \left\| (1/n) \sum_{i=1}^{n} (Y_i - EY_1) \right\| \), so we can assume without loss of generality that \( EY_1 = 0 \).

Let \( v_\ell, 1 \leq \ell < \infty \) be a basis for the square integrable functions defined on \( \mathcal{C} \). Then we have

\[
Y_i(t) = \sum_{\ell=1}^{\infty} \langle Y_i, v_\ell \rangle v_\ell(t) \quad \text{and} \quad \|Y_i\|^2 = \sum_{\ell=1}^{\infty} \langle Y_i, v_\ell \rangle^2.
\]

According to Assumption A.2

\[
E \left( \sum_{\ell=1}^{\infty} \langle Y_i, v_\ell \rangle^2 \right)^{1/2} < \infty
\]

and therefore

\[
\lim_{K \to \infty} E \left( \sum_{\ell=K+1}^{\infty} \langle Y_i, v_\ell \rangle^2 \right)^{1/2} = 0. \tag{A.1}
\]

For every fixed \( K \geq 1 \) we define

\[
Y_{i,K}(t) = \sum_{\ell=1}^{K} \langle Y_i, v_\ell \rangle v_\ell(t), \quad 1 \leq i < \infty.
\]

By the triangle inequality we have

\[
\left\| \sum_{i=1}^{n} Y_i \right\| \leq \left\| \sum_{i=1}^{n} (Y_i - Y_{i,K}) \right\| + \left\| \sum_{i=1}^{n} Y_{i,K} \right\| \leq \left\| \sum_{i=1}^{n} Y_i - Y_{i,K} \right\| + \left\| \sum_{i=1}^{n} Y_{i,K} \right\|.
\]

It follows from Assumption A.1 that \( \|Y_i - Y_{i,K}\|, 1 \leq i < \infty \) is a stationary and ergodic sequence and by Assumption A.1 \( E\|Y_1 - Y_{1,K}\| < \infty \). So using the ergodic theorem for random variables (cf. [8, p. 118]) we conclude that

\[
\frac{1}{n} \sum_{i=1}^{n} \|Y_i - Y_{i,K}\| \to E\|Y_1 - Y_{1,K}\| \quad \text{a.s.} \tag{A.2}
\]
Let $\varepsilon > 0$. Putting together (A.1) and (A.2) we can find an integer $K = K(\varepsilon)$ such that

$$\lim_{n \to \infty} \frac{1}{n} \left\| Y_i - Y_{i,K} \right\| \leq \varepsilon \quad \text{a.s.} \quad (A.3)$$

For every fixed $K$ we can write

$$\sum_{i=1}^{n} Y_{i,K} = \sum_{\ell=1}^{K} \sum_{i=1}^{n} \langle Y_i, v_{\ell} \rangle v_{\ell}(t)$$

and therefore

$$\left\| \sum_{i=1}^{n} Y_{i,K} \right\|^2 = \left( \sum_{\ell=1}^{K} \sum_{i=1}^{n} \langle Y_i, v_{\ell} \rangle \right)^2.$$

For every $\ell \geq 1$, the sequence $\langle Y_i, v_{\ell} \rangle$, $1 \leq i < \infty$ is stationary and ergodic and by Assumption A.2

$$E|\langle Y_1, v_{\ell} \rangle| \leq E\left\| Y_1 \right\| < \infty.$$

Hence $EY_1(t) = 0$ implies $E\langle Y_1, v_{\ell} \rangle = \int_{c}^{e} EY_1(t) v_{\ell}(t) dt = 0$. So using again the ergodic theorem in [8, p. 118] we get for any $\ell$

$$\frac{1}{n} \sum_{i=1}^{n} \langle Y_i, v_{\ell} \rangle \to 0 \quad \text{a.s.}$$

resulting in

$$\left\| \sum_{i=1}^{n} Y_i \right\| \to 0 \quad \text{a.s.} \quad (A.4)$$

for every $K \geq 1$. It follows from (A.2) and (A.4) that

$$\limsup_{n \to \infty} \left\| \frac{1}{n} \sum_{i=1}^{n} Y_i \right\| \leq \varepsilon \quad \text{a.s.}$$

for any $\varepsilon > 0$, which completes the proof of Theorem A.1. □

References