KPSS test for functional time series

Piotr Kokoszka*  
Colorado State University

Gabriel Young  
Colorado State University

Abstract

Econometric and financial data often take the form of a collection of curves observed consecutively over time. Examples include, intraday price curves, term structure curves, and intraday volatility curves. Such curves can be viewed as a time series of functions. A fundamental issue that must be addressed, before an attempt is made to statistically model or predict such series, is whether they contain a random walk component. This paper extends the KPSS test to the setting of functional time series. We develop the form of the test statistic, and propose two testing procedures: Monte Carlo and asymptotic. The limit distributions are derived, the procedures are algorithmically described and illustrated by an application to yield curves and a simulation study.

JEL Classification: C12, C32.

Keywords: Functional data, Integrated time series, Random walk, Trend stationarity, Yield curve.

1 Introduction

Many econometric and financial data sets take the form of a time series of curves, or functions. The best known and most extensively studied data of this form are yield curves. Even though they are observed at discrete maturities, in financial theory they are viewed as continuous functions, one function per month or per day. The yield curves can thus be viewed as a time series of curves, a functional time series. Other examples include intraday price, volatility or volume curves. Intraday price curves are smooth, volatility and volume

*Address for correspondence: Piotr Kokoszka, Department of Statistics, Colorado State University, Fort Collins, CO 80522, USA. E-mail: Piotr.Kokoszka@colostate.edu
curves are noisy and must be smoothed before they can be effectively treated as curves. As with scalar and vector valued time series, it is important to describe the random structure of a functional time series. A fundamental question, which has received a great deal of attention in econometric research, is whether the time series has a random walk, or unit root, component. The present paper addresses this issue in the context of functional time series by extending the KPSS test of Kwiatkowski et al. (1992) and developing the required asymptotic theory.

The work of Kwiatkowski et al. (1992) was motivated by the fact that unit root tests developed by Dickey and Fuller (1979), Dickey and Fuller (1981), and Said and Dickey (1984) indicated that most aggregate economic series had a unit root. In these tests, the null hypothesis is that the series has a unit root. Since such tests have low power in samples of sizes occurring in many applications, Kwiatkowski et al. (1992) proposed that trend stationarity should be considered as the null hypothesis, and the unit root should be the alternative. Rejection of the null of trend stationarity could then be viewed as a convincing evidence in favor of a unit root. It was soon realized that the KPSS test of Kwiatkowski et al. (1992) has a much broader utility. For example, Lee and Schmidt (1996) and Giraitis et al. (2003) used it to detect long memory, with short memory as the null hypothesis; de Jong et al. (1997) developed a robust version of the KPSS test. The work of Lo (1991) is crucial because he observed that under temporal dependence, to obtain parameter free limit null distributions, statistics similar to the KPSS statistic must be normalized by the long run variance rather than by the sample variance. Likewise, there have been dozens of contributions that enhance unit root testing, see e.g. Cavaliere and Xu (2014) and Chambers et al. (2014), and references therein.

In the functional setting, the null hypothesis of trend stationarity is stated as follows:

\[(1.1) \quad H_0 : \quad X_n(t) = \mu(t) + n\xi(t) + \eta_n(t).\]

Note that the time index is \(n\), it refers to day, month, quarter or year. The index \(t\) refers to “time” for each function. For example, for intraday price curves, \(t\) is time within a trading day, measured in minutes or at an even finer resolution. For yield curves, \(t\) does not correspond to physical time but to time until expiration, the maturity horizon of a bond. The functions \(\mu\) and \(\xi\) correspond, respectively, to the intercept and slope. The errors \(\eta_n\) are also functions which model random departures of the observed functions \(X_n\) form a deterministic model. Under the alternative, the model contains a random walk component:

\[(1.2) \quad H_A : \quad X_n(t) = \mu(t) + n\xi(t) + \sum_{i=1}^{n} u_i(t) + \eta_n(t),\]

where \(u_1, u_2, \ldots\) are mean zero identically distributed random functions.
Our approach to testing exploits the ideas of functional data analysis (FDA), mostly those related to functional principal component expansions; the monographs of Bosq (2000), Ramsay and Silverman (2002) and Horváth and Kokoszka (2012) explain them in detail. Section 2 contains the prerequisites needed to understand our theory. Application of FDA methodology in an econometric context is not new. Among others, Kargin and Onatski (2008) studied prediction of yield curves, Müller et al. (2011) considered functional modeling of volatility, Kokoszka et al. (2014) used a regression type model to explains the shapes of price curves. A contribution most closely related to the present work is Horváth et al. (2014a) who developed a test of level stationarity. Incorporating a possible trend, changes the structure of functional residuals and leads to different limit distributions. It also requires the asymptotic analysis of the long run variance of these residuals.

The remainder of the paper is organized as follows. Section 2 introduces the requisite background in FDA and states and explains the assumptions. Test statistics and their asymptotic distributions are described in Section 3. Section 4 contains algorithmic descriptions of the test procedures, their application to yield curves, and a small simulation study. Proofs of the theorems stated in Section 3 are developed in Section 5.

2 Definitions and assumptions

All random functions and deterministic functional parameters $\mu$ and $\xi$ are assumed to be elements of the Hilbert space $L^2 = L^2([0, 1])$ with the inner product $\langle f, g \rangle = \int_0^1 f(t)g(t)dt$. This means that the domain of all functional observations, e.g. of the daily price or yield curves, has been normalized to be the unit interval. If the limits of integration are omitted, integration is over the interval $[0, 1]$. All random functions are assumed to be square integrable, i.e. $E||\eta_n||^2 < \infty$, $E||\eta_n||^2 < N$, $E||\eta_n||^2 < \infty$. Further background on random elements of $L^2$ is given in Chapter 2 of Horváth and Kokoszka (2012); a more extensive theoretical treatment is presented in Bosq (2000).

The partial sum process of the curves $X_1(t), X_2(t), \ldots, X_N(t)$ is defined by

\begin{equation}
S_N(x, t) = \frac{1}{\sqrt{N}} \sum_{n=1}^{[Nx]} X_n(t), \quad 0 \leq t, x \leq 1,
\end{equation}

and the partial sum process of the unobservable errors by

\begin{equation}
V_N(x, t) = \frac{1}{\sqrt{N}} \sum_{n=1}^{[Nx]} \eta_n(t), \quad 0 \leq t, x \leq 1.
\end{equation}

Kwiatkowski et al. (1992) assumed that the errors $\eta_n$ are iid, but subsequent research extended their work to errors which form a stationary time series, see e.g. Giraitis et al.
(2003) and references therein. In the case of scalar observations, temporal dependence can be quantified in many ways, e.g. via structural, mixing or cumulant conditions, and a large number of asymptotic results established under such assumptions can be used. For functional time series, the corresponding results are much fewer and fall into two categories: 1) those derived assuming a linear, ARMA type, structure, Bosq (2000); 2) those assuming a nonlinear moving average representation (Bernoulli shifts) with the decay of dependence specified by a moment condition. We will use a general nonparametric specification of the dependence structure which falls to the second category. It is quantified by the following assumption.

**Assumption 2.1** The errors $\eta_j$ are Bernoulli shifts, i.e. $\eta_j = g(\epsilon_j, \epsilon_{j-1}, \ldots)$ for some measurable function $g : S^{\infty} \to L^2$ and iid functions $\epsilon_j$, $-\infty \leq j \leq \infty$, with values in a measurable space $S$. The functions $(t,\omega) \mapsto \eta_j(t,\omega)$ are product measurable.

$E\eta_0 = 0$ in $L^2$ and $E\|\eta_0\|^{2+\delta} < \infty$ for some $0 < \delta < 1$.

The sequence $\{\eta_n\}_{n=-\infty}^{\infty}$ can be approximated by $\ell$-dependent sequences $\{\eta_{n,\ell}\}_{n=-\infty}^{\infty}$ in the sense that

$$\sum_{\ell=1}^{\infty} (E\|\eta_n - \eta_{n,\ell}\|^{2+\delta})^{1/\kappa} < \infty \text{ for some } \kappa > 2 + \delta,$$

where $\eta_{n,\ell}$ is defined by

$$\eta_{n,\ell} = g(\epsilon_n, \epsilon_{n-1}, \ldots, \epsilon_{n-\ell}, \epsilon_n^{*}, \epsilon_{n-\ell-1}, \ldots),$$

where the $\epsilon_i^*$ are independent copies of $\epsilon_0$, independent of $\{\epsilon_i, -\infty < i < \infty\}$.

Assumption 2.1 has been shown to hold for all known models for temporally dependent functions, assuming the parameters of these models satisfy nonrestrictive conditions, see H"ormann and Kokoszka (2010, 2012), or Chapter 16 of Horváth and Kokoszka (2012). Its gist is that the dependence of the function $g$ on the iid innovations $\epsilon_j$ far in the past decays so fast that these innovations can be replaced by their independent copies. Such a replacement is asymptotically negligible in the sense quantified by (2.3). For scalar time series, conditions similar in spirit were used by P"otscher and Prucha (1997), Wu (2005), Shao and Wu (2007) and Berkes et al. (2011), to name just a few references. In this paper, Assumption 2.1 is needed to ensure that the partial sums (2.2) can be approximated by a two-parameter Gaussian process. In particular, (2.3) is not used directly; it is a condition used by Berkes et al. (2013) to prove Theorem 5.2. To establish the results of Section 3, one can, in fact, replace Assumption 2.1 by the conclusions of Theorem 5.2 and the existence of an estimator $\hat{c}(t,s)$ such that

$$\int \int [\hat{c}(t,s) - c(t,s)]^2 dt ds \to 0, \quad \text{as } N \to \infty,$$
with the kernel $c$ defined by (2.5). Assumption 2.1 is a general weak dependence condition under which these conclusions hold. To establish the consistency of the tests under $H_A$ (1.2), exactly the same assumptions must be imposed on the random walk errors $u_i$. The error sequences $\eta_n$ and $u_i$ thus need not be iid, but merely stationary and weakly dependent. In summary, the conditions stated in Assumption 2.1 provide a nonrestrictive, nonparametric quantification of weak dependence of a sequence of functions which implies all technical properties needed in the proofs.

We now define the bivariate functions appearing in (2.4). The long-run covariance function of the errors $\eta_n$ is defined as

$$c(t, s) = E\eta_0(t)\eta_0(s) + \sum_{i=1}^{\infty} (E\eta_0(t)\eta_i(s) + E\eta_0(s)\eta_i(t)).$$

The series defining the function $c(t, s)$ converges in $L^2([0,1] \times [0,1])$, see Horváth et al. (2013). The function $c(t, s)$ is positive definite. Therefore there exist eigenvalues $\lambda_1 \geq \lambda_2 \geq ... \geq 0$, and orthonormal eigenfunctions $\phi_i(t)$, $0 \leq t \leq 1$, satisfying

$$\lambda_i \phi_i(t) = \int c(t, s)\phi_i(s)ds, \quad 0 \leq i \leq \infty.$$

To ensure that the $\phi_i$ corresponding to the $d$ largest eigenvalues are uniquely defined (up to a sign), we assume that

$$\lambda_1 > \lambda_2 > \cdots > \lambda_d > \lambda_{d+1} > 0.$$

The eigenvalues $\lambda_i$ play a crucial role in our tests. They are estimated by the sample, or empirical, eigenvalues defined by

$$\hat{\lambda}_i \hat{\phi}_i(t) = \int \hat{c}(t, s)\hat{\phi}_i(s)ds, \quad 0 \leq i \leq N,$$

where $\hat{c}(\cdot, \cdot)$ is an estimator of (2.5). We use a kernel estimator similar to that introduced by Horváth et al. (2013), but with suitably defined residuals in place of the centered observations $X_n$. To define model residuals, consider the least squares estimators of the functional parameters $\xi(t)$ and $\mu(t)$ in model (1.1):

$$\hat{\xi}(t) = \frac{1}{s_N} \sum_{n=1}^{N} \left( n - \frac{N + 1}{2} \right) X_n(t)$$

with

$$s_N = \sum_{n=1}^{N} \left( n - \frac{N + 1}{2} \right)^2.$$
and

\begin{equation}
\hat{\mu}(t) = \bar{X}(t) - \hat{\xi}(t) \left( \frac{N+1}{2} \right).
\end{equation}

The functional residuals are therefore

\begin{equation}
e_n(t) = (X_n(t) - \bar{X}(t)) - \hat{\xi}(t) \left( n - \frac{N+1}{2} \right), \quad 1 \leq n \leq N.
\end{equation}

Defining their empirical autocovariances by

\begin{equation}
\hat{\gamma}_i(t, s) = \frac{1}{N} \sum_{j=i+1}^{N} e_j(t)e_{j-i}(s), \quad 0 \leq i \leq N-1,
\end{equation}

leads to the kernel estimator

\begin{equation}
\hat{c}(t, s) = \hat{\gamma}_0(t, s) + \sum_{i=1}^{N-1} K\left( \frac{i}{h} \right) (\hat{\gamma}_i(t, s) + \hat{\gamma}_i(s, t)).
\end{equation}

The following assumption is imposed on kernel function \( K \) and the bandwidth \( h \).

**Assumption 2.2** The function \( K \) is continuous, bounded, \( K(0) = 1 \) and \( K(u) = 0 \) if \(|u| > c\), for some \( c > 0 \). The smoothing bandwidth \( h = h(N) \) satisfies

\begin{equation}
 h(N) \to \infty, \quad \frac{h(N)}{N} \to 0, \quad \text{as} \quad N \to \infty.
\end{equation}

The assumption that \( K \) vanishes outside a compact interval is not crucial to establish (2.4). It is a simplifying condition which could be replaced by a sufficiently fast decay condition, at the cost of technical complications in the proof of (2.4).

Recall that if \( \{W(x), 0 \leq x \leq 1\} \) is a standard Brownian motion (Wiener process), then the Brownian bridge is defined by \( B(x) = W(x) - xW(1), \quad 0 \leq x \leq 1 \). The second–level Brownian bridge is defined by

\begin{equation}
V(x) = W(x) + \left(2x - 3x^2\right)W(1) + \left(-6x + 6x^2\right) \int_0^1 W(y)dy, \quad 0 \leq x \leq 1.
\end{equation}

Both the Brownian bridge and the second–level Brownian bridge are special cases of the generalized Brownian bridge introduced by MacNeill (1978) who studied the asymptotic behavior of partial sums of polynomial regression residuals. Process (2.16) appears as the null limit of the KPSS statistic of Kwiatkowski et al. (1992). We will see in Section 3 that for functional data the limit involves an infinite sequence of independent and identically distributed second-level Brownian bridges \( V_1(x), V_2(x), \ldots \).
3 Test statistics and their limit distributions

Horváth et al. (2014a) developed tests of level–stationarity of a functional time series, i.e. of the null hypothesis \( X_n(t) = \mu(t) + \eta_n(t) \), using the partial sum process

\[
U_N(x,t) = \frac{1}{\sqrt{N}} \sum_{n=1}^{[Nx]} (X_n(t) - \bar{X}(t)) = S_N(x,t) - \frac{[Nx]}{N}S_N(1,t),
\]

where \( S_N(x,t) \) is defined in (2.1). The process \( U_N(x,t) \) has the form of a functional Brownian bridge. Their main statistic

\[
T_N = \int \int U_N^2(x,t) \, dt \, dx = \int \|U_N(x, \cdot)\|^2 \, dx, \quad 0 \leq t, x \leq 1,
\]

is asymptotically distributed, under the null, as \( \sum_{i=1}^{\infty} \lambda_i \int B_i^2(x) \, dx \), where \( \lambda_1, \lambda_2, \ldots \) are eigenvalues of the long–run covariance function of the observations \( X_n \), and \( B_1, B_2, \ldots \) are iid Brownian bridges. In the case of trend stationarity, a different distribution arises; the \( B_i \) must be replaced by second level Brownian bridges, and the \( \lambda_i \) are defined differently. The remainder of this section explains the details.

The test statistic for the trend-stationary case is based on the partial sum process of residuals (2.12). Observe that

\[
\frac{1}{\sqrt{N}} \sum_{n=1}^{[Nx]} e_n(t) = \frac{1}{\sqrt{N}} \sum_{n=1}^{[Nx]} ((X_n(t) - \bar{X}(t)) - \hat{\xi}(t)(n - (N + 1)/2))
\]

\[
= \frac{1}{\sqrt{N}} \sum_{n=1}^{[Nx]} X_n(t) - \frac{1}{\sqrt{N}} \sum_{n=1}^{[Nx]} \bar{X}(t) - \frac{\hat{\xi}(t)}{\sqrt{N}} \sum_{n=1}^{[Nx]} (n - (N + 1)/2)
\]

\[
= \frac{1}{\sqrt{N}} \sum_{n=1}^{[Nx]} X_n(t) - \frac{[Nx]}{N} \frac{1}{\sqrt{N}} \sum_{n=1}^{N} X_n(t) - \frac{\hat{\xi}(t)}{2\sqrt{N}} ([xN]/([xN] - N))
\]

\[
= S_N(x,t) - \frac{[Nx]}{N}S_N(1,t) - \frac{1}{2} N^{3/2} \hat{\xi}(t) \left( \frac{[xN]}{N} \left( \frac{[xN]}{N} - 1 \right) \right).
\]

A suitable test statistic is therefore given by

\[
R_N = \int \int Z_N^2(x,t) \, dt \, dx = \int \|Z_N(x, \cdot)\|^2 \, dx, \quad 0 \leq t, x \leq 1,
\]

where

\[
Z_N(x,t) = S_N(x,t) - \frac{[Nx]}{N}S_N(1,t) - \frac{1}{2} N^{3/2} \hat{\xi}(t) \left( \frac{[Nx]}{N} \left( \frac{[Nx]}{N} - 1 \right) \right)
\]

and \( S_N(x,t) \) and \( \hat{\xi}(t) \) are, respectively, defined in equations (2.1) and (2.9). The null limit distribution of test statistic (3.1) is given in the following theorem.
Theorem 3.1 If Assumption 2.1 holds, then under null model (1.1),

\[ R_N \xrightarrow{d} \sum_{i=1}^{\infty} \lambda_i \int V_i^2(x)dx, \]

where \( \lambda_1, \lambda_2, \ldots \) are eigenvalues of the long–run covariance function (2.5), and \( V_1, V_2, \ldots \) are iid second–level Brownian bridges.

The proof of Theorem 3.1 is given in Section 5. We now explain the issues arising in the functional case by comparing our result to that obtained by Kwiatkowski et al. (1992). If all curves are constant functions \( (X(t) = X_i \text{ for } t \in [0,1]) \), the statistic \( R_N \) given by (3.1) is the numerator of the KPSS test statistic of Kwiatkowski et al. (1992), which is given by

\[ KPSS_N = \frac{1}{N^2 \hat{\sigma}^2} \sum_{n=1}^{N} S_n^2 = \frac{R_N}{\hat{\sigma}^2}, \]

where \( \hat{\sigma}^2 \) is a consistent estimator of the long-run variance \( \sigma^2 \) of the residuals. In the scalar case, Theorem 3.1 reduces to \( R_N \xrightarrow{d} \sigma^2 \int_0^1 V^2(x)dx \), where \( V(x) \) is a second–level Brownian bridge. If \( \hat{\sigma}^2 \) is a consistent estimator of \( \sigma^2 \), the result of Kwiatkowski et al. (1992) is recovered, i.e. \( KPSS_N \xrightarrow{d} \int_0^1 V^2(x)dx \). In the functional case, the eigenvalues \( \lambda_i \) can be viewed as long–run variances of the residual curves along the principal directions determined by the eigenfunctions of the kernel \( c(\cdot, \cdot) \) defined by (2.5). To obtain a test analogous to the scalar KPSS test, with a parameter free limit null distribution, we must construct a statistic which involves a division by consistent estimators of the \( \lambda_i \). We use only \( d \) largest eigenvalues in order not to increase the variability of the statistic caused by division by small empirical eigenvalues. A suitable statistic is

\[ R_0^N = \sum_{i=1}^{d} \frac{1}{\hat{\lambda}_i} \int_0^1 (Z_N(x, \cdot), \hat{\phi}_i)^2 dx, \]

where the sample eigenvalues \( \hat{\lambda}_i \) and eigenfunctions \( \hat{\phi}_i \) are defined by (2.8). Statistic (3.3) extends the statistic \( KPSS_N \). Its limit distribution is given in the next theorem.

Theorem 3.2 If Assumptions 2.1, 2.2 and (2.7) hold, then under null model (1.1),

\[ R_0^N \xrightarrow{d} \sum_{i=1}^{d} \int_0^1 V_i^2(x)dx, \]

with the \( V_i, 1 \leq i \leq d \), the same as in Theorem 3.1.
Theorem 3.2 is proven in Section 5. Here we only note that the additional Assumption 2.2 is needed to ensure that (2.4) holds which is known to imply \( \hat{\lambda}_i \overset{P}{\to} \lambda_i, \ 1 \leq i \leq d. \)

We conclude this section by discussing the consistency of the tests based on the above theorems. Theorem 3.3 implies that under \( H_A \) statistic \( R_N \) of Theorem 3.1 increases like \( N^2 \). The critical values increase at the rate not greater than \( N \). The test based on Theorem 3.1 is thus consistent. The exact asymptotic behavior under \( H_A \) of the normalized statistic \( R_N^0 \) appearing in Theorem 3.2 is more difficult to study due to almost intractable asymptotics (under \( H_A \)) of the empirical eigenvalues and eigenfunctions of the kernel \( \hat{c}(\cdot, \cdot) \). The precise asymptotic behavior under \( H_A \) is not known even in the scalar case, i.e. for the statistic \( KPSS_N \). We therefore focus on the asymptotic limit under \( H_A \) of the statistic \( R_N \) whose derivation is already quite complex. This limit involves iid copies of the process

\[
(3.4) \quad \Delta(x) = \int_0^x W(y)dy + (3x^2 - 4x) \int_0^1 W(y)dy + (-6x^2 + 6x) \int_0^1 yW(y)dy, \quad 0 \leq x \leq 1, 
\]

where \( W(\cdot) \) is a standard Brownian motion.

**Theorem 3.3**  
If the errors \( u_i \) satisfy Assumption 2.1, then under the alternative (1.2),

\[
\frac{1}{N^2} R_N \overset{D}{\to} \sum_{i=1}^{\infty} \tau_i \int_0^1 \Delta_i^2(x)dx, 
\]

where \( R_N \) is the test statistic defined in (3.1) and \( \Delta_1, \Delta_2(x), \ldots \) are iid processes defined by (3.4). The weights \( \tau_i \) are the eigenvalues of the long-run covariance kernel of the errors \( u_i \), defined analogously to (2.5) by

\[
(3.5) \quad c_u(t, s) = E[u_0(t)u_0(s)] + \sum_{l=1}^{\infty} Eu_0(t)u_l(s) + \sum_{l=1}^{\infty} Eu_0(s)u_l(t). 
\]

The proof of Theorem 3.3 is given in Section 5.

### 4 Implementation and application to yield curves

We begin this section with algorithmic descriptions of tests based on Theorems 3.1 and 3.2.

**Algorithm 4.1** [Monte Carlo test based on Theorem 3.1]

1. Estimate the null model (1.1) and compute the residuals defined in equation (2.12).
2. Select kernel \( K \) and a bandwidth \( h \) in (2.14) and compute the eigenvalues \( \hat{\lambda}_i \hat{\phi}_i, \ 1 \leq i \leq N \), defined by (2.8).
Table 1: Critical values of the distribution of the variable $R^0(d)$ given by (4.1).

<table>
<thead>
<tr>
<th>$d$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>10%</td>
<td>0.1201</td>
<td>0.2111</td>
<td>0.2965</td>
<td>0.3789</td>
<td>0.4576</td>
</tr>
<tr>
<td>Size</td>
<td>5%</td>
<td>0.1494</td>
<td>0.2454</td>
<td>0.3401</td>
<td>0.4186</td>
</tr>
<tr>
<td>1%</td>
<td>0.2138</td>
<td>0.3253</td>
<td>0.4257</td>
<td>0.5149</td>
<td>0.6131</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>10%</td>
<td>0.5347</td>
<td>0.6150</td>
<td>0.6892</td>
<td>0.7646</td>
<td>0.8416</td>
</tr>
<tr>
<td>Size</td>
<td>5%</td>
<td>0.5909</td>
<td>0.6687</td>
<td>0.7482</td>
<td>0.8252</td>
</tr>
<tr>
<td>1%</td>
<td>0.6960</td>
<td>0.7799</td>
<td>0.8574</td>
<td>0.9487</td>
<td>1.0326</td>
</tr>
</tbody>
</table>

3. Simulate a large number, say $G = 10,000$, of vectors $[V_1, V_2, \ldots, V_N]$ consisting of independent second level Brownian bridge processes $V_i$ defined in (2.16). Find the 95th percentile, $R_{\text{critical}}$, of the $G$ replications of $R^*$ given by

$$R^*_N = \sum_{i=1}^{N} \hat{\lambda}_i \int_0^1 V_i^2(x)dx.$$ 

4. Compute the test statistic $R_N$ defined in (3.1). If $R_N \geq R_{\text{critical}}$, reject $H_0$ at the 5% significance level.

In most applications, the $\hat{\lambda}_i$ decay very quickly to zero, so if $N$ is large, it can be replaced in Algorithm 4.1 by a smaller number, e.g by $d = 20$, and the empirical distribution of the $R^*_N$ can be replaced by that of the $R^*_d$. In Algorithm 4.1 the critical value must be obtained via Monte Carlo simulations for each data set. In Algorithm 4.2, tabulated critical values can be used. These depend on the number $d$ of the functional principal components used to construct statistic $R_N^d$. Typically $d$ is a small, single digit, number. Table 1 lists selected critical values. They have been obtained by simulating $G = 10,000$ vectors $[V_1, V_2, \ldots, V_d]$ and finding the percentiles of the $G$ replications of

$$R^0(d) = \sum_{i=1}^{d} \int_0^1 V_i^2(x)dx.$$ 

**Algorithm 4.2** [Asymptotic test based on Theorem 3.2]

1. Perform steps 1 and 2 of Algorithm 4.1.
2. Choose the smallest $d$ such that $\sum_{i \leq d} \hat{\lambda}_i / \sum_{i \leq N} \hat{\lambda}_i > 0.85$. 

10
3. Calculate the statistic $R^0_N$ given by (3.3) and reject $H_0$ if $R^0_N > R^0_{\text{critical}}$, with the critical value given in Table 1.

In some applications, Step 2 of Algorithm 4.2 may be replaced by a selection of $d$ based on a visual fit of the truncated principal component expansion

$$X^{(d)}_n(t) = \hat{\mu}(t) + \sum_{j=1}^{d} \langle X_n, \hat{\phi}_j \rangle \hat{\phi}_j(t)$$

to the observed curves $X_n(t)$. In other applications, existing theory or experience may support certain choices of $d$. This is the case for the yield curves, which we use to illustrate the application of our tests. In both algorithms, testing at a fixed level can be replaced by the calculation of P–values based on the Monte Carlo distributions of $R^*_N$ or $R^0(d)$.

An important step is the choice of $h$ needed to estimate the long run covariance function. A great deal of research in this direction has been done for scalar and vector time series. For functional time series, the method proposed by Horváth et al. (2014b) often gives good results. It uses the flat top kernel

$$K(t) = \begin{cases} 1, & 0 \leq t < 0.1 \\ 1.1 - |t|, & 0.1 \leq t < 1.1 \\ 0, & |t| \geq 1.1 \end{cases}$$

(4.2)

advocated by Politis and Romano (1996, 1999) and Politis (2011), and a data driven selection of $h$. This method performs well if the series length $N$ is larger than several hundred, longer than the series we consider. In our data example, a deterministic bandwidth $h = N^{2/5}$ (combined with the flat top kernel) produced good size and power. The optimal selection of $h$ is not a focus of this paper, this complex issue must be investigated in a separate work.

Intended merely as an illustration rather than an extensive empirical study, we apply both tests to yield curves. We consider a time series of daily United States Federal Reserve yield curves constructed from discrete rates at maturities of 1, 3, 6, 12, 24, 36, 60, 84, 120 and 360 months. Yield curves are discussed in many finance textbooks, see e.g. Chapter 10 of Campbell et al. (1997) or Diebold and Rudebusch (2013). Figure 1 shows five consecutive yield curves. Following the usual practice, each yield curve is treated as a single functional observation, and so the yield curves observed over a period of many days form a functional time series. Figure 2 shows the sample period we study, which covers 100 consecutive trading days. It shows a downward trend in interest rates, and we want to test if these curves also contain a random walk component. The tests were performed using $d = 2$. The first two principal components of $\hat{c}$ explain over 95% of variance and provide excellent visual fit. Our selection thus uses three principal shapes to describe the yield curves, the mean function
and the first two principal components. It is in agreement with recent approaches to modeling the yield curve, cf. Hays et al. (2012) and Diebold and Rudebusch (2013), which are based on the three component Nelson–Siegel model.

We first apply both tests to the time series of yield curves shown in Figure 2, \( N = 100 \). The test based on statistic \( R_N \), Algorithm 4.1, yields the P–value of 0.0282 and the test based on \( R_0^N \), Algorithm 4.2, 0.0483, indicating the presence of random walk in addition to a downward trend. Extending the sample by adding the next 150 business days, so that \( N = 250 \), yields the respective P–values 0.0005 and 0.0013. In all computation the bandwidth \( h = N^{2/5} \) has been used. Examination of different periods shows that trend stationarity does not hold if the period under observation is sufficiently long. This agrees with the empirical finding of Chen and Niu (2014) whose method of yield curve prediction, based on utilizing periods of approximate stationarity, performs better than predictions based on the whole sample; random walk is not predictable. Even though our tests are motivated by the alternative of a random walk component, they reject any serious violation of trend stationarity. Broadly speaking, our analysis shows that daily yield curves can be treated as a trend stationary functional time series only over short periods of time, generally not longer than a calendar quarter.

We complement our data example with a small simulation study. There is a multitude of data generating process that could be used. The following quantities could vary: shapes of the mean and the principal components functions, the magnitude of the eigenvalues, the distribution of the scores and their dependence structure. In this paper, concerned chiefly with theory, we merely want to present a very limited simulation study that validates the
conclusions of the data example. We therefore attempt to simulate curves whose shapes resemble those of the real data, and for which either the null or the alternative holds. The artificial data is therefore generated according to the following algorithm.

**Algorithm 4.3** [Yield curves simulation under \( H_0 \)]

1. Using real yield curves, calculate the estimates \( \hat{\xi}(t) \) and \( \hat{\mu}(t) \) defined, respectively, by (2.9) and (2.11). Then compute the residuals \( e_n(t) \) defined in (2.12).
2. Calculate the first two empirical principal components \( \hat{\phi}_1(t) \) and \( \hat{\phi}_2(t) \) using the empirical covariance function

\[
\hat{\gamma}_0(s,t) = \frac{1}{N} \sum_{n=1}^{N} (e_n(t) - \bar{e}(t))(e_n(s) - \bar{e}(s)).
\]

This step leads to the approximation

\[
e_n(t) \approx a_{1,n}\hat{\phi}_1(t) + a_{2,n}\hat{\phi}_2(t), \quad n = 1, 2, \ldots, N,
\]

where \( a_{1,n} \) and \( a_{2,n} \) are the first two functional scores. The functions \( \hat{\phi}_1(t) \) and \( \hat{\phi}_2(t) \) are treated as deterministic, while the scores \( a_{1,n} \) and \( a_{2,n} \) form random sequences indexed by \( n \).
3. To simulate temporally independent residuals \( e_n \), generate independent in \( n \) scores \( a'_{1,n} \sim N(0, \sigma_{a_1}^2) \) and \( a'_{2,n} \sim N(0, \sigma_{a_2}^2) \), where \( \sigma_{a_1}^2 \) and \( \sigma_{a_2}^2 \) are the sample variances of
the real scores, and set

\[ e'_n(t) = a'_{1,n}\hat{\phi}_1(t) + a'_{2,n}\hat{\phi}_2(t), \quad n = 1, 2, \ldots, N. \]

To simulate dependent residual curves, generate scores \( a'_{1,n}, a'_{2,n} \sim \text{AR}(1) \), where each autoregressive process has parameter 0.5.

4. Using the estimated functional parameters \( \hat{\mu}(t), \hat{\xi}(t) \) and the simulated residuals \( e'_n(t) \), construct the simulated data set

\[
X'_n(t) = \hat{\mu}(t) + \hat{\xi}(t)n + e'_n(t), \quad n = 1, 2, \ldots, N.
\]

Table 2 shows empirical sizes based on 1000 replication of the data generating process described in Algorithm 4.3. We use two ways of estimating the eigenvalues and eigenfunctions. The first one uses the function \( \hat{\gamma}_0 \) defined by (4.3) (in the scalar case this corresponds to using the usual sample variance rather than estimating the long-run variance). The second uses the estimated long-run covariance function (2.14) with the bandwidth \( h \) specified in Table 2. The covariance kernel \( \hat{\gamma}_0(t,s) \) is appropriate for independent error curves. The bandwidth \( h = N^{1/3} \) is too small, not enough temporal dependence is absorbed. The bandwidth \( h = N^{2/5} \) gives fairly consistent empirical size, typically within one percent of the empirical size. The bandwidth \( h \) is not relevant when the kernel \( \hat{\gamma}_0 \) is used. The different empirical sizes reflect random variability due to three different sets of 1000 replications being used. This indicates that with 1000 replications, a difference of one percent in empirical sizes is not significant.

Table 2: Empirical sizes for functional time series generated using Algorithm 4.3.

<table>
<thead>
<tr>
<th>N</th>
<th>Test statistic</th>
<th>( R_N )</th>
<th>( \hat{\gamma}_0(t,s) )</th>
<th>( \hat{\xi}(t,s) )</th>
<th>( \hat{\mu}(t,s) )</th>
<th>( R'_N )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DGP</td>
<td>iid normal</td>
<td>iid normal</td>
<td>AR(1)</td>
<td>iid normal</td>
<td>iid normal</td>
</tr>
<tr>
<td></td>
<td>Cov-kernel</td>
<td>( \hat{\gamma}_0(t,s) )</td>
<td>( \hat{\xi}(t,s) )</td>
<td>( \hat{\mu}(t,s) )</td>
<td>( \hat{\gamma}_0(t,s) )</td>
<td>( \hat{\xi}(t,s) )</td>
</tr>
<tr>
<td>100</td>
<td>( h = N^{1/3} )</td>
<td>6.3</td>
<td>5.6</td>
<td>9.4</td>
<td>5.9</td>
<td>5.2</td>
</tr>
<tr>
<td></td>
<td>( h = N^{2/5} )</td>
<td>5.6</td>
<td>4.4</td>
<td>6.6</td>
<td>5.8</td>
<td>3.6</td>
</tr>
<tr>
<td></td>
<td>( h = N^{1/2} )</td>
<td>5.1</td>
<td>4.8</td>
<td>3.5</td>
<td>4.5</td>
<td>5.1</td>
</tr>
<tr>
<td>250</td>
<td>( h = N^{1/3} )</td>
<td>5.0</td>
<td>4.3</td>
<td>10.2</td>
<td>5.8</td>
<td>5.2</td>
</tr>
<tr>
<td></td>
<td>( h = N^{2/5} )</td>
<td>5.5</td>
<td>4.9</td>
<td>7.2</td>
<td>4.5</td>
<td>4.1</td>
</tr>
<tr>
<td></td>
<td>( h = N^{1/2} )</td>
<td>5.5</td>
<td>5.9</td>
<td>4.3</td>
<td>4.8</td>
<td>3.4</td>
</tr>
<tr>
<td>1000</td>
<td>( h = N^{1/3} )</td>
<td>4.8</td>
<td>4.2</td>
<td>7.0</td>
<td>5.9</td>
<td>5.6</td>
</tr>
<tr>
<td></td>
<td>( h = N^{2/5} )</td>
<td>6.1</td>
<td>6.3</td>
<td>6.3</td>
<td>6.0</td>
<td>5.1</td>
</tr>
<tr>
<td></td>
<td>( h = N^{1/2} )</td>
<td>5.8</td>
<td>4.9</td>
<td>4.6</td>
<td>5.6</td>
<td>4.7</td>
</tr>
</tbody>
</table>
To evaluate power, instead of (4.4), the data generating process is

\begin{equation}
X_n(t) = \hat{\mu}(t) + \hat{\xi}(t)n + \sum_{i=1}^{n} u_i(t) + \epsilon_n(t), \quad n = 1, 2, \ldots, N,
\end{equation}

where the increments \( u_i \) are defined by

\[ u_i(t) = aN_{i1} \sin(\pi t) + aN_{i2} \sin(2\pi t), \]

with standard normal \( N_{ij}, \; j = 1, 2, \; 1 \leq i \leq N \), totally independent of each other. The scalar \( a \) quantifies the distance from \( H_0; \; ia = 0 \), corresponds to \( H_0 \). For all empirical power simulations, we use a 5% size critical value and \( h = N^{2/5} \). The empirical power reported in Table 3 increases as the sample size and the distance from \( H_0 \) increase. It is visibly higher for iid curves as compared to dependent curves.

\begin{table}[h]
\centering
\begin{tabular}{|l|c|c|c|c|c|}
\hline
\multicolumn{2}{|c|}{Test statistic} & \multicolumn{2}{|c|}{\( R_N \)} & \multicolumn{2}{|c|}{\( R_N^0 \)} \\
DGP & Cov-kernel & iid normal & iid normal & AR(1) & iid normal & iid normal & AR(1) \\
\hline
\hline
a = 0.1 & \( N = 125 \) & 100.0 & 89.9 & 10.1 & 100.0 & 87.9 & 10.4 \\
& \( N = 250 \) & 100.0 & 97.0 & 27.7 & 100.0 & 96.0 & 21.9 \\
\hline
a = 0.5 & \( N = 125 \) & 100.0 & 91.5 & 83.1 & 100.0 & 89.7 & 71.2 \\
& \( N = 250 \) & 100.0 & 97.3 & 96.4 & 100.0 & 97.4 & 92.4 \\
\hline
\end{tabular}
\caption{Empirical power based on the DGP (4.5) and \( h = N^{2/5} \).}
\end{table}

5 Proofs of the results of Section 3

5.1 Preliminary results

For ease of reference, we state in this section two theorems which are used in the proofs of the results of Section 3. Theorem 5.1 is well-known, see Theorem 4.1 in Billingsley (1968). Theorem 5.2 was recently established in Berkes et al. (2013).

\begin{theorem}
Suppose \( Z_N, Y_N, Y \) are random variables taking values in a separable metric space with the distance function \( \rho \). If \( Y_N \xrightarrow{P} Y \) and \( \rho(Z_N, Y_N) \xrightarrow{P} 0 \), then \( Z_N \xrightarrow{P} Y \).
\end{theorem}
In our setting, we work in the metric space $D([0, 1], L^2)$ which is the space of right-continuous functions with left limits taking values in $L^2([0, 1])$. A generic element of $D([0, 1], L^2)$ is

$$z = \{z(x, t), 0 \leq x \leq 1, 0 \leq t \leq 1\}.$$ 

For each fixed $x$, $z(x, \cdot) \in L^2$, so $\|z(x, \cdot)\|^2 = \int z^2(x, t)dt < \infty$. The uniform distance between $z_1, z_2 \in D([0, 1], L^2)$ is

$$\rho(z_1, z_2) = \sup_{0 \leq x \leq 1} |z_1(x, \cdot) - z_2(x, \cdot)| = \sup_{0 \leq x \leq 1} \left\{ \int (z_1(x, t) - z_2(x, t))^2 dt \right\}^{1/2}.$$ 

In the following, we work with the space $D([0, 1], L^2)$ equipped with the uniform distance.

**Theorem 5.2** If Assumption 2.1 holds, then $\sum_{i=1}^{\infty} \lambda_i < \infty$, and we can construct a sequence of Gaussian processes $\Gamma_N(x, t)$ such that for every $N$

$$\{\Gamma_N(x, t), 0 \leq x, t \leq 1\} \overset{D}{\rightarrow} \{\Gamma(x, t), 0 \leq x, t \leq 1\},$$

(5.1)

where $\Gamma(x, t) = \sum_{i=1}^{\infty} \lambda_i^{1/2} W_i(x) \phi_i(t)$,

and

(5.2)

$$\kappa_N = \sup_{0 \leq x \leq 1} \|V_N(x, \cdot) - \Gamma_N(x, \cdot)\| = o_p(1).$$

Recall that the $W_i$ are independent standard Wiener processes, $\lambda_i$ and $\phi_i$ are defined in (2.6) and $V_N(x, t)$ is defined in (2.2).

### 5.2 Proof of Theorem 3.1

The proof of Theorem 3.1 is constructed from several lemmas decomposing the statistic $R_N$ into a form suitable for the application of the results of Section 5.1, i.e. to leading and asymptotically negligible terms. Throughout this section, we assume that the null model (1.1) and Assumption 2.1 hold.

**Lemma 5.1** For the $s_N$ is defined in (2.10),

$$\frac{s_N}{N^3} \rightarrow \frac{1}{12}, \quad \text{as } N \rightarrow \infty.$$ 

**Proof:** Identify the left–hand side with a Riemann sum. \(\square\)

**Lemma 5.2** For the functional slope estimate $\hat{\xi}$ defined by (2.9),

$$\hat{\xi}(t) - \xi(t) = \frac{1}{s_N} \sum_{n=1}^{N} \left( n - \frac{N + 1}{2} \right) \eta_n(t).$$

16
Proof: Use the identities
\[ \sum_{n=1}^{N} \left( n - \frac{N + 1}{2} \right) = 0, \quad \sum_{n=1}^{N} \left( n - \frac{N + 1}{2} \right)^2 = s_N. \]

Lemma 5.3 The following identity holds
\[ \frac{N^{3/2}}{s_N} \sum_{n=1}^{N} \left( n - \frac{N + 1}{2} \right) \eta_n(t) = \frac{1}{N^{-3} s_N} \left\{ \left( \frac{N-1}{2N} \right) V_N(1, t) - \frac{1}{N} \sum_{k=1}^{N-1} V_N \left( \frac{k}{N}, t \right) \right\}, \]
where \( V_N(x, t) \) is the partial sum process of the errors defined in (2.2).

Proof: Notice that
\[ \frac{N^{3/2}}{s_N} \sum_{n=1}^{N} \left( n - \frac{N + 1}{2} \right) \eta_n(t) = \frac{1}{N^{3/2} s_N} \left\{ \frac{1}{N} \sum_{n=1}^{N} n \eta_n(t) - \left( \frac{N+1}{2N} \right) \frac{1}{\sqrt{N}} \sum_{n=1}^{N} \eta_n(t) \right\}. \]

Using the relation
\[ \sum_{n=1}^{N} n \eta_n(t) = N \sum_{n=1}^{N} \eta_n(t) - \sum_{k=1}^{N-1} \sum_{n=1}^{k} \eta_n(t), \]
we have
\[ \frac{N^{3/2}}{s_N} \sum_{n=1}^{N} \left( n - \frac{N + 1}{2} \right) \eta_n(t) = \frac{1}{N^{3/2} s_N} \left\{ \frac{1}{N} \sum_{n=1}^{N} \eta_n(t) - \sum_{k=1}^{N-1} \sum_{n=1}^{k} \eta_n(t) \right\} = \frac{1}{N^{3/2} s_N} \left\{ \left( \frac{N-1}{2N} \right) V_N(1, t) - \frac{1}{N} \sum_{k=1}^{N-1} V_N \left( \frac{k}{N}, t \right) \right\}. \]

\[ \square \]
Lemma 5.4. The process $Z_N(x,t)$ defined by (3.2) admits the decomposition

$$Z_N(x,t) = V_N(x,t) - \left\lfloor \frac{N x}{N} \right\rfloor V_N(1,t)$$

$$- \frac{1}{2} N^{-3} s_N \left\{ \left( \frac{N - 1}{2N} \right) V_N(1,t) - \frac{1}{N} \sum_{k=1}^{N-1} V_N\left( \frac{k}{N}, t \right) \right\} \left\lfloor \frac{N x}{N} \right\rfloor \left( \frac{N x}{N} - 1 \right).$$

Proof: Notice that

$$Z_N(x,t) = S_N(x,t) - \left\lfloor \frac{N x}{N} \right\rfloor S_N(1,t) - \frac{1}{N} N^{3/2} \hat{\xi}(t) \left( \frac{\left\lfloor \frac{N x}{N} \right\rfloor}{N} \left( \frac{N x}{N} - 1 \right) \right)$$

$$= \frac{1}{\sqrt{N}} \sum_{n=1}^{\left\lfloor \frac{N x}{N} \right\rfloor} X_n(t) - \left\lfloor \frac{N x}{N} \right\rfloor \frac{1}{\sqrt{N}} \sum_{n=1}^{N} X_n(t) - \frac{1}{N} N^{3/2} \hat{\xi}(t) \left( \frac{\left\lfloor \frac{N x}{N} \right\rfloor}{N} \left( \frac{N x}{N} - 1 \right) \right)$$

$$= \frac{1}{\sqrt{N}} \sum_{n=1}^{\left\lfloor \frac{N x}{N} \right\rfloor} (\mu(t) + \xi(t)n + \eta_n(t)) - \frac{\left\lfloor \frac{N x}{N} \right\rfloor}{N} \frac{1}{\sqrt{N}} \sum_{n=1}^{N} (\mu(t) + \xi(t)n + \eta_n(t))$$

$$- \frac{1}{2} N^{3/2} \hat{\xi}(t) \left( \frac{\left\lfloor \frac{N x}{N} \right\rfloor}{N} \left( \frac{N x}{N} - 1 \right) \right).$$

Therefore,

$$Z_N(x,t) = \mu(t) \frac{1}{\sqrt{N}} \left\lfloor N x \right\rfloor + \xi(t) \frac{1}{\sqrt{N}} \sum_{n=1}^{\left\lfloor N x \right\rfloor} n + \frac{1}{\sqrt{N}} \sum_{n=1}^{\left\lfloor N x \right\rfloor} \eta_n(t)$$

$$- \mu(t) \frac{1}{\sqrt{N}} \left\lfloor N x \right\rfloor \frac{1}{N} - \frac{\left\lfloor N x \right\rfloor}{N} \frac{1}{\sqrt{N}} \xi(t) \sum_{n=1}^{N} n - \frac{\left\lfloor N x \right\rfloor}{N} \sum_{n=1}^{N} \eta_n(t)$$

$$- \frac{1}{2} N^{3/2} \hat{\xi}(t) \left( \frac{\left\lfloor N x \right\rfloor}{N} \left( \frac{N x}{N} - 1 \right) \right)$$

$$= \frac{1}{\sqrt{N}} \sum_{n=1}^{\left\lfloor N x \right\rfloor} \eta_n(t) - \frac{\left\lfloor N x \right\rfloor}{N} \sum_{n=1}^{N} \eta_n(t)$$

$$+ \mu(t) \frac{1}{\sqrt{N}} \left\lfloor N x \right\rfloor - \mu(t) \frac{1}{\sqrt{N}} \left\lfloor N x \right\rfloor$$

$$+ \xi(t) \frac{1}{\sqrt{N}} \sum_{n=1}^{\left\lfloor N x \right\rfloor} n - \frac{\left\lfloor N x \right\rfloor}{N} \xi(t) \frac{1}{\sqrt{N}} \sum_{n=1}^{N} n$$

$$- \frac{1}{2} N^{3/2} \hat{\xi}(t) \left( \frac{\left\lfloor N x \right\rfloor}{N} \left( \frac{N x}{N} - 1 \right) \right).$$
By Lemma 5.2,

\[ Z_N(x,t) = \frac{1}{\sqrt{N}} \sum_{n=1}^{\lfloor Nx \rfloor} \eta_n(t) - \frac{\lfloor Nx \rfloor}{N} \sum_{n=1}^{N} \eta_n(t) \]
\[ + \xi(t) \frac{1}{\sqrt{N}} \frac{\lfloor Nx \rfloor (\lfloor Nx \rfloor + 1)}{2} - \frac{\lfloor Nx \rfloor}{N} \xi(t) \frac{1}{\sqrt{N}} \frac{N(N + 1)}{2} \]
\[ - \frac{1}{2} N^{3/2} \xi(t) \left( \frac{\lfloor Nx \rfloor}{N} \left( \frac{\lfloor Nx \rfloor}{N} - 1 \right) \right) \]
\[ = \frac{1}{\sqrt{N}} \sum_{n=1}^{\lfloor Nx \rfloor} \eta_n(t) - \frac{\lfloor Nx \rfloor}{N} \sum_{n=1}^{N} \eta_n(t) \]
\[ - \frac{1}{2} N^{3/2} \left( \xi(t) - \xi(t) \right) \left( \frac{\lfloor Nx \rfloor}{N} \left( \frac{\lfloor Nx \rfloor}{N} - 1 \right) \right) \]
\[ = \frac{1}{\sqrt{N}} \sum_{n=1}^{\lfloor Nx \rfloor} \eta_n(t) - \frac{\lfloor Nx \rfloor}{N} \sum_{n=1}^{N} \eta_n(t) \]
\[ - \frac{1}{2} \left( \frac{N^{3/2}}{s_N} \sum_{n=1}^{N} \left( n - \frac{N + 1}{2} \right) \eta_n(t) \right) \left( \frac{\lfloor Nx \rfloor}{N} \left( \frac{\lfloor Nx \rfloor}{N} - 1 \right) \right). \]

By Lemma 5.3,

\[ Z_N(x,t) = V_N(x,t) - \frac{\lfloor Nx \rfloor}{N} V_N(1,t) \]
\[ - \frac{1}{2} \frac{1}{N^3 s_N} \left( \frac{N - 1}{2N} \right) V_N(1,t) - \frac{1}{N} \sum_{k=1}^{N-1} V_N \left( \frac{k}{N}, t \right) \frac{\lfloor Nx \rfloor}{N} \left( \frac{\lfloor Nx \rfloor}{N} - 1 \right). \]

\[ \square \]

**Lemma 5.5** The following convergence holds

\[ \int \left\{ \frac{1}{N} \sum_{k=1}^{N} V_N \left( \frac{k}{N}, t \right) - \int_{0}^{1} \Gamma_N(y,t) \, dy \right\}^2 \, dt \xrightarrow{P} 0, \]

where the \( \Gamma_N \) are the Gaussian processes in Theorem 5.2.

**Proof:** Since \( V_N \) is a step function with jumps at \( y = \frac{k}{N} \),

\[ \frac{1}{N} \sum_{k=1}^{N} V_N \left( \frac{k}{N}, t \right) = \int_{0}^{1} V_N(y,t) \, dy. \]
We must thus show that
\[
\left\| \int_0^1 V_N(y, \cdot) dy - \int_0^1 \Gamma_N(y, \cdot) dy \right\| \xrightarrow{P} 0.
\]
Using the contractive property of integrals and relation (5.2), we have
\[
\left\| \int_0^1 V_N(y, \cdot) dy - \int_0^1 \Gamma_N(y, \cdot) dy \right\| \leq \int_0^1 \| V_N(y, \cdot) - \Gamma_N(y, \cdot) \| dy \\
\leq \int_0^1 \sup_{0 \leq x \leq 1} \| V_N(y, \cdot) - \Gamma_N(y, \cdot) \| dy \\
\leq \kappa_N = o_p(1)
\]
which proves Lemma 5.5.

**Lemma 5.6** Consider the process \( \Gamma(\cdot, \cdot) \) defined by (5.1) and set
\[
(5.3) \quad \Gamma^0(x, t) = \Gamma(x, t) + \left(2x - 3x^2\right) \Gamma(1, t) + \left(-6x + 6x^2\right) \int_0^1 \Gamma(y, t) dy.
\]
Then
\[
\int_0^1 \| \Gamma^0(x, \cdot) \|^2 dx = \sum_{i=1}^{\infty} \lambda_i \int_0^1 V_i^2(x) dx.
\]

**Proof:** Expansion (5.1) implies that
\[
\Gamma^0(x, t) = \Gamma(x, t) + \left(2x - 3x^2\right) \Gamma(x, t) + \left(-6x + 6x^2\right) \int_0^1 \Gamma(y, t) dy \\
= \sum_{i=1}^{\infty} \sqrt{\lambda_i} W_i(x) \phi_i(t) + \left(2x - 3x^2\right) \sum_{i=1}^{\infty} \sqrt{\lambda_i} W_i(1) \phi_i(t) \\
+ \left(-6x + 6x^2\right) \int_0^1 \sum_{i=1}^{\infty} \sqrt{\lambda_i} W_i(y) \phi_i(t) dy \\
= \sum_{i=1}^{\infty} \sqrt{\lambda_i} \left\{ W_i(x) + \left(2x - 3x^2\right) W_i(1) \\
+ \left(-6x + 6x^2\right) \int_0^1 W_i(y) dy \right\} \phi_i(t) \\
= \sum_{i=1}^{\infty} \sqrt{\lambda_i} V_i(x) \phi_i(t),
\]
where $V_1, V_2, \ldots$ are iid second–level Brownian bridges defined in (2.16). By the orthonormality of the eigenfunctions $\phi_i,$

$$
\int_0^1 \|\Gamma^0(x, \cdot)\|^2 dx = \iint (\Gamma^0(x, t))^2 dt dx \\
= \iint \left( \sum_{i=1}^{\infty} \sqrt{\lambda_i} V_i(x) \phi_i(t) \right)^2 dx dt \\
= \sum_{i=1}^{\infty} \lambda_i \int_0^1 V_i^2(x) dx.
$$

\[ \square \]

**Lemma 5.7** For the processes $Z_N(\cdot, \cdot)$ and $\Gamma^0_N(\cdot, \cdot)$ defined, respectively, in (3.2) and (5.3),

\[ (5.4) \sup_{0 \leq x \leq 1} \|Z_N(x, \cdot) - \Gamma^0_N(x, \cdot)\| \xrightarrow{P} 0. \]

**Proof:** Using the decomposition given in Lemma 5.4, $Z_N(x, t)$ can be algebraically manipulated to be expressed as

$$
Z_N(x, t) = V_N(x, t) + a_N(x) V_N(1, t) + b_N(x) \frac{1}{N} \sum_{k=1}^{N-1} V_N \left( \frac{k}{N}, t \right)
$$

where

\[ (5.5) a_N(x) = \left( \frac{1}{2N^{-3}s_N} \left( \frac{N-1}{2N} \right) - 1 \right) \left( \frac{N_x}{N} \right) - \frac{1}{2N^{-3}s_N} \left( \frac{N-1}{2N} \right) \left( \frac{N_x}{N} \right)^2 \]

and

$$
b_N(x) = \frac{1}{2N^{-3}s_N} \left( \frac{N_x}{N} \right) \left( \frac{N_x}{N} - 1 \right).
$$

Notice that

$$
\|Z_N(x, \cdot) - \Gamma^0_N(x, \cdot)\| \leq \left\| V_N(x, t) - \Gamma_N(x, t) \right\| \\
+ \left\| a_N(x) V_N(1, t) - \left( 2x - 3x^2 \right) \Gamma_N(1, t) \right\| \\
+ \left\| b_N(x) \frac{1}{N} \sum_{k=1}^{N-1} V_N \left( \frac{k}{N}, t \right) - \left( -6x + 6x^2 \right) \int_0^1 \Gamma_N(y, t) dy \right\|.
$$

Thus, Lemma 5.7 will be proven once we have established the following relations:

\[ (5.6) \sup_{0 \leq x \leq 1} \|V_N(x, \cdot) - \Gamma_N(x, \cdot)\| \xrightarrow{P} 0; \]
(5.7) \[ \sup_{0 \leq x \leq 1} \| a_n(x) V_N(1, \cdot) - (2x - 3x^2) \Gamma_N(1, \cdot) \| \xrightarrow{P} 0; \]

(5.8) \[ \sup_{0 \leq x \leq 1} \left\| \frac{1}{N} \sum_{k=1}^{N-1} V_N\left( \frac{k}{N}, \cdot \right) - (-6x + 3x^2) \int_0^1 \Gamma_N(y, \cdot) dy \right\| \xrightarrow{P} 0. \]

Relation (5.6) is the conclusion of Theorem 5.2. The verification of relation (5.7) follows next. Since

\[ \| a_n(x) V_N(1, \cdot) - (2x - 3x^2) \Gamma_N(1, \cdot) \| \leq |a_n(x)| \| V_N(1, \cdot) - \Gamma_N(1, \cdot) \| + \| a_N(x) - (2x - 3x^2) \| \| \Gamma_N(1, \cdot) \|, \]

relation (5.7) will hold once we have verified that

(5.9) \[ \sup_{N \geq 1} \left\{ \sup_{0 \leq x \leq 1} |a_N(x)| \right\} < \infty, \]

(5.10) \[ \| V_N(1, \cdot) - \Gamma_N(1, \cdot) \| \xrightarrow{P} 0, \]

(5.11) \[ \sup_{0 \leq x \leq 1} |a_N(x) - (2x - 3x^2)| \to 0, \]

(5.12) \[ \| \Gamma_N(1, \cdot) \| = O_p(1). \]

Relation (5.10) follows from Theorem 5.2. By (5.1),

\[ \| \Gamma_N(1, \cdot) \|^2 \xrightarrow{P} \sum_{i=1}^{\infty} \lambda_i Z_i^2, \]

where the \( Z_i \) are independent standard normal and \( \sum_{i=1}^{\infty} \lambda_i < \infty. \) Thus relation (5.12) holds trivially because the distribution of the left-hand side does not depend on \( N. \) To show (5.9), set

\[ d_N = \frac{1}{2N - 3s_N} \left( \frac{N - 1}{2N} \right). \]

By Lemma 5.1, \( d_N \to 3, \) and

\[ a_n(x) = (d_N - 1) \frac{\lfloor Nx \rfloor}{N} - d_N \left( \frac{\lfloor Nx \rfloor}{N} \right)^2. \]

22
Since $\lfloor Nx \rfloor \leq N$, $|a_N(x)| \leq |d_N - 1| + d_N$, and (5.9) follows. To show relation (5.11), first notice that

$$|a_N(x) - (2x - 3x^2)| \leq (d_N - 1) \frac{\lfloor Nx \rfloor}{N} - 2x + d_N \left(\frac{\lfloor Nx \rfloor}{N}\right)^2 - 3x^2.$$  

We must thus show that

$$\sup_{0 \leq x \leq 1} (d_N - 1) \frac{\lfloor Nx \rfloor}{N} - 2x \to 0,$$

and

$$\sup_{0 \leq x \leq 1} d_N \left(\frac{\lfloor Nx \rfloor}{N}\right)^2 - 3x^2 \to 0.$$  

To show (5.13), first notice that $Nx - 1 \leq \lfloor Nx \rfloor \leq Nx$, which implies

$$(d_N - 1)x - 2x - \frac{1}{N}(d_N - 1) \frac{\lfloor Nx \rfloor}{N} - 2x \leq (d_N - 1)x - 2x.$$  

Both sides of the inequality are linear in $x$ which implies the extrema occurs at the boundaries. Thus for any $x \in [0, 1]$,

$$\left|(d_N - 1) \frac{\lfloor Nx \rfloor}{N} - 2x\right| \leq |d_N - 3| \to 0,$$

which proves relation (5.13). To show (5.14), notice that $(Nx - 1)^2 \leq \lfloor Nx \rfloor^2 \leq (Nx)^2$, implying

$$d_N x^2 - 3x^2 - \left(\frac{2x}{N} - \frac{1}{N^2}\right) d_N \leq d_N \left(\frac{\lfloor Nx \rfloor}{N}\right)^2 - 3x^2 \leq d_N x^2 - 3x^2.$$  

For any $x \in [-1, 0]$, the quadratic functions on each side of the inequality are strictly decreasing. Hence, the extrema occurs at the boundaries implying for any $x \in [0, 1]$,

$$\left|d_N \left(\frac{\lfloor Nx \rfloor}{N}\right)^2 - 3x^2\right| \leq |d_N - 3| \to 0.$$  

This proves (5.14) and hence relation (5.11) holds true. Thus relations (5.9), (5.10), (5.11) and (5.11) hold true which proves relation (5.7). That is

$$\sup_{0 \leq x \leq 1} \|a_n(x)V_N(1, \cdot) - (2x - 3x^2)\Gamma_N(1, \cdot)\| \to 0.$$  

23
Next is the verification (5.8). As in the case of (5.7), it suffices to show that

\[(5.15) \sup_{N \geq 1} \left\{ \sup_{0 \leq x \leq 1} |b_N(x)| \right\} < \infty, \]

\[(5.16) \left\| \frac{1}{N} \sum_{k=1}^{N-1} V_N \left( \frac{k}{N}, \cdot \right) - \int_0^1 \Gamma_N(y, \cdot)dy \right\| \overset{P}{\to} 0, \]

\[(5.17) \sup_{0 \leq x \leq 1} |b_N(x) - (6x^2 - 6x)| \to 0, \]

\[(5.18) \left\| \int_0^1 \Gamma_N(y, \cdot)dy \right\| = O_P(1). \]

The verification of (5.15) and (5.17) uses the same arguments as the verification of (5.9) and (5.11) so we do not need to present the details. Relation (5.16) coincides with Lemma 5.5, while relation (5.18) follows from Lemma 5.6. This completes the proof of Lemma 5.7. □

Using the above lemmas, we can now present a compact proof of Theorem 3.1.

**Proof of Theorem 3.1:** Recall that the test statistic $R_N$ is defined by $R_N = \int \int Z_N^2(x,t)dxdt$, where

\[Z_N(x,t) = S_N(x,t) - \frac{|Nx|}{N} S_N(1,t) - \frac{1}{2} N^{3/2} \hat{\xi}(t) \left( \frac{|Nx|}{N} \left( \frac{|Nx|}{N} - 1 \right) \right)\]

with $S_N(x,t)$ and $\hat{\xi}(t)$ are respectively defined in equations (2.1) and (2.9). Recall that

\[\Gamma^0_N(x,t) = \Gamma_N(x,t) + \left( 2x - 3x^2 \right) \Gamma_N(1,t) + \left( -6x + 6x^2 \right) \int_0^1 \Gamma_N(y,t)dy, \]

and

\[\Gamma^0(x,t) = \Gamma(x,t) + \left( 2x - 3x^2 \right) \Gamma(1,t) + \left( -6x + 6x^2 \right) \int_0^1 \Gamma(y,t)dy. \]

From Lemma 5.7, we know that

\[\rho(Z_N(x,\cdot), \Gamma^0_N(x,\cdot)) = \sup_{0 \leq x \leq 1} \| Z_N(x,\cdot) - \Gamma^0_N(x,\cdot) \| \overset{P}{\to} 0. \]

By Theorem 5.2, $\Gamma^0_N(x,t) \overset{D}{=} \Gamma^0(x,t)$. Thus, Theorem 5.1 implies that

\[Z_N(x,t) \overset{D}{=} \Gamma^0(x,t). \]
By Lemma 5.6,
\[ \int \int (\Gamma^0(x,t))^2 \, dx \, dt \overset{d}{=} \sum_{i=1}^{d} \lambda_i \int_0^1 V_i^2(x) \, dx. \]
Thus, by the continuous mapping theorem,
\[ R_N = \int \int (Z_N(x,t))^2 \, dx \, dt \overset{D}{\to} \sum_{i=1}^{\infty} \lambda_i \int V_i^2(x) \, dx, \]
which proves the desired result. \( \square \)

### 5.3 Proof of Theorem 3.2

The key fact needed in the proof is the consistency of the sample eigenvalues \( \hat{\lambda}_i \) and eigenfunctions \( \hat{\phi}_i \). The required result, stated in (5.19), follows fairly directly from (2.4). However, the verification that (2.4) holds for the kernel estimator (2.14) is not trivial. The required result can be stated as follows.

**Theorem 5.3** Suppose Assumption 2.1 holds with \( \delta = 0 \) and \( \kappa = 2 \). If \( H_0 \) and Assumption 2.2 hold, then relation (2.4) holds.

Observe that assuming that relation (2.3) in Assumption 2.1 holds with \( \delta = 0 \) and \( \kappa = 2 \) weakens the universal assumption that it holds with some \( \delta > 0 \) and \( \kappa > 2 + \delta \).

We first present the proof of Theorem 3.2, which uses Theorem 5.3, and then turn to a rather technical proof of Theorem 5.3.

**Proof of Theorem 3.2:** If Assumptions 2.1, 2.2, condition (2.7) and \( H_0 \) hold, then

\[ \max_{1 \leq i \leq d} |\hat{\lambda}_i - \lambda_i| = o_p(1) \quad \text{and} \quad \max_{1 \leq i \leq d} \|\hat{\phi}_i - \hat{c}_i \phi_i\| = o_p(1), \]

where \( \hat{c}_1, \hat{c}_2, \ldots, \hat{c}_d \) are unobservable random signs defined as \( \hat{c}_i = \text{sign}(\langle \hat{\phi}_i, \phi_i \rangle) \). Indeed, Theorem 5.3 states that relation (2.4) holds under \( H_0 \) and Assumptions 2.1 and 2.2. Relations (5.19) follow from (2.4) and Lemmas 2.2. and 2.3 of Horváth and Kokoszka (2012) which state that the differences of the eigenvalues and eigenfunctions are bounded by the Hilbert–Schmidt norm of the difference of the corresponding operators.

Using (5.1), it is easy to see that for all \( N \)

\[ \{\langle \Gamma_N^0(x,\cdot), \phi_i \rangle, 0 \leq x \leq 1, 1 \leq i \leq d\} \overset{D}{=} \{\sqrt{\lambda_i} V_i(x), 0 \leq x \leq 1, 1 \leq i \leq d\}. \]

We first show that

\[ \sup_{0 \leq x \leq 1} |\langle Z_N(x,\cdot), \hat{\phi}_i \rangle - \langle \Gamma_N^0(x,\cdot), \hat{c}_i \phi_i \rangle| \overset{P}{\to} 0. \]
By the Cauchy-Schwarz inequality and Lemma 5.7, we know
\[
\sup_{0 \leq x \leq 1} |\langle Z_N(x, \cdot) - \Gamma^0_N(x, \cdot), \hat{\phi}_i \rangle| \leq \sup_{0 \leq x \leq 1} \| Z_N(x, \cdot) - \Gamma^0_N(x, \cdot) \| = o_p(1).
\]

Again by the Cauchy-Schwarz inequality and (5.19), we have
\[
\sup_{0 \leq x \leq 1} |\langle \Gamma^0_N(x, \cdot), \hat{\phi}_i - \hat{c}_i \phi_i \rangle| \leq \sup_{0 \leq x \leq 1} \| \Gamma^0_N(x, \cdot) \| \| \hat{\phi}_i - \hat{c}_i \phi_i \| = o_p(1).
\]

Then using the triangle inequality and inner product properties,
\[
\sup_{0 \leq x \leq 1} |\langle Z_N(x, \cdot), \hat{\phi}_i \rangle - \langle \Gamma_N(x, \cdot), \hat{\phi}_i \rangle| = o_p(1),
\]
which proves relation (5.21). Thus by Theorem 5.1, (5.19), (5.21), (5.20) and the continuous mapping theorem,
\[
R^0_N = \sum_{i=1}^{d} \frac{1}{\lambda_i} \int \langle Z_N(x, \cdot), \hat{\phi}_i \rangle^2 dx \overset{d}{\to} \sum_{i=1}^{d} \int V^2_i(x) dx.
\]

**Proof of Theorem 5.3:** Recall definitions of the kernels $c$ and $\hat{c}$ given, respectively, in (2.5) and (2.14). The claim will follow if we can show that
\[
\int \int \{ \hat{\gamma}_0(t, s) - E[\eta_0(t)\eta_0(s)] \}^2 dt ds = o_P(1)
\]
and
\[
\int \int \left\{ \sum_{i=1}^{N-1} K\left( \frac{i}{h} \right) \hat{\gamma}_i(t, s) - \sum_{i \geq 1} E[\eta_0(s)\eta_i(t)] \right\}^2 dt ds = o_P(1).
\]

These relations are established in a sequence of Lemmas which split the argument by isolating the terms related to the estimation of trend from those related to the autocovariances of the $\eta_i$. The latter terms were treated in Horváth et al. (2013), so the present proof focuses on the extra terms appearing in our context.
Lemma 5.8 Under model (1.1), the following relation holds

\[ \hat{\gamma}_0(t, s) = \frac{1}{N} \sum_{i=1}^{N} (\eta_i(t) - \bar{\eta}(t))(\eta_i(s) - \bar{\eta}(s)) \]

\[ - \frac{1}{NsN} \left\{ \sum_{i=1}^{N} \left( i - \frac{N+1}{2} \right) \eta_i(t) \right\} \left\{ \sum_{i=1}^{N} \left( i - \frac{N+1}{2} \right) \eta_i(s) \right\}. \]

Proof: Observe that \( \sum_{i=1}^{N} \left( i - \frac{N+1}{2} \right) = 0 \), and so

\[ \sum_{i=1}^{N} \left( i - \frac{N+1}{2} \right) i = \sum_{i=1}^{N} \left( i - \frac{N+1}{2} \right)^2 = s_N. \]

Also recall that

\[ \hat{\xi}(t) - \xi(t) = \frac{1}{s_N} \sum_{i=1}^{N} \left( i - \frac{N+1}{2} \right) \eta_i(t) = \frac{1}{s_N} \sum_{i=1}^{N} (\eta_i(t) - \bar{\eta}(t)) \left( i - \frac{N+1}{2} \right). \]

We can express the residuals \( e_n(t) \) as

\[ e_i(t) = (X_i(t) - \bar{X}(t)) - \hat{\xi}(t) \left( i - \frac{N+1}{2} \right) \]

\[ = \mu(t) + i\xi(t) + \eta_i(t) - \frac{1}{N} \sum_{i=1}^{N} (\mu(t) + i\xi(t) + \eta_i(t)) - \hat{\xi}(t) \left( i - \frac{N+1}{2} \right) \]

\[ = (\eta_i(t) - \bar{\eta}(t)) - (\hat{\xi}(t) - \xi(t)) \left( i - \frac{N+1}{2} \right). \]
Then, by the above relations, 

\[ \hat{\gamma}_0(t, s) = \frac{1}{N} \sum_{i=1}^{N} e_i(t) e_i(s) \]

\[ = \frac{1}{N} \sum_{i=1}^{N} \left[ (\eta_i(t) - \bar{\eta}(t)) - (\xi_i(t) - \xi(t)) \left( i - \frac{N+1}{2} \right) \right] \left[ (\eta_i(s) - \bar{\eta}(s)) - (\xi_i(s) - \xi(s)) \left( i - \frac{N+1}{2} \right) \right] \]

\[ = \frac{1}{N} \sum_{i=1}^{N} (\eta_i(t) - \bar{\eta}(t)) (\eta_i(s) - \bar{\eta}(s)) \]

\[ - (\hat{\xi}(s) - \xi(s)) \frac{1}{N} \sum_{i=1}^{N} (\eta_i(t) - \bar{\eta}(t)) \left( i - \frac{N+1}{2} \right) \]

\[ - (\hat{\xi}(t) - \xi(t)) \frac{1}{N} \sum_{i=1}^{N} (\eta_i(s) - \bar{\eta}(s)) \left( i - \frac{N+1}{2} \right) \]

\[ + (\hat{\xi}(t) - \xi(t)) (\hat{\xi}(s) - \xi(s)) \frac{1}{N} \sum_{i=1}^{N} \left( i - \frac{N+1}{2} \right)^2 \]

\[ = \frac{1}{N} \sum_{i=1}^{N} (\eta_i(t) - \bar{\eta}(t)) (\eta_i(s) - \bar{\eta}(s)) \]

\[ - \frac{1}{N s_N} \sum_{i=1}^{N} \left( i - \frac{N+1}{2} \right) \eta_i(s) \sum_{i=1}^{N} \left( i - \frac{N+1}{2} \right) \eta_i(t) \]

\[ - \frac{1}{N s_N} \sum_{i=1}^{N} \left( i - \frac{N+1}{2} \right) \eta_i(t) \sum_{i=1}^{N} \left( i - \frac{N+1}{2} \right) \eta_i(s) \]

\[ + \frac{1}{N s_N} \sum_{i=1}^{N} \left( i - \frac{N+1}{2} \right) \eta_i(t) \frac{1}{s_N} \sum_{i=1}^{N} \left( i - \frac{N+1}{2} \right) \eta_i(s) s_N. \]

The claim thus follows because the last two terms cancel. \[\square\]

To lighten the notation, in the remainder of this section we set 

\[ k_i = i - \frac{N+1}{2}. \]

**Lemma 5.9** Relation (5.22) holds under the assumptions of Theorem 5.3.

**Proof:** We must show that \( \| \hat{\gamma}_0 - \gamma_0 \| \overset{P}{\to} 0 \), where \( \gamma_0(t, s) = E[\eta_0(t)\eta_0(s)] \) and the norm is in \( L^2([0,1] \times [0,1]) \). We will use the decomposition of Lemma 5.8, i.e.

\[ \hat{\gamma}_0(t, s) = \tilde{\gamma}_0(t, s) - v_N(t, s) \]

28
where
\[ \tilde{\gamma}_0(t, s) = \frac{1}{N} \sum_{i=1}^{N} (\eta_i(t) - \bar{\eta}(t))(\eta_i(s) - \bar{\eta}(s)) \]
and
\[ v_N(t, s) = \frac{1}{N s_N} \left\{ \sum_{i=1}^{N} k_i \eta_i(t) \right\} \left\{ \sum_{i=1}^{N} k_i \eta_i(s) \right\}. \]

It will be enough to show that
\[ \| \tilde{\gamma}_0 - \gamma_0 \| \overset{P}{\to} 0 \quad \text{and} \quad \| v_N \| \overset{P}{\to} 0. \]

The first convergence is the consistency of the sample covariance function which was proven by Horváth et al. (2013). The remainder of the proof is devoted to the verification that \( \| v_N \| \overset{P}{\to} 0. \)

Since
\[ \| v_N \|^2 = \frac{1}{N^2 s_N^2} \int \left\{ \sum_{i=1}^{N} k_i \eta_i(t) \right\}^2 dt \int \left\{ \sum_{i=1}^{N} k_i \eta_i(s) \right\}^2 ds, \]
we must show that
\[
(5.24) \quad \| v_N \| = \frac{1}{N s_N} \left\| \sum_{i=1}^{N} k_i \eta_i \right\|^2 \overset{P}{\to} 0,
\]
where the norm in (5.24) is in \( L^2([0, 1]) \). Using diagonal summation, we get
\[
\left\| \sum_{i=1}^{N} k_i \eta_i \right\|^2 = \sum_{i=1}^{N} \sum_{j=1}^{N} k_i k_j \langle \eta_i, \eta_j \rangle \\
= \sum_{i=1}^{N} k_i^2 \| \eta_i \|^2 + 2 \sum_{l=1}^{N-1} \sum_{i=1}^{N-l} k_i k_{i+l} \langle \eta_i, \eta_{i+l} \rangle.
\]

Since \( \sum_{i=1}^{N} k_i^2 = s_N \), we obtain
\[
(5.25) \quad E \left[ \sum_{i=1}^{N} k_i^2 \| \eta_i \|^2 \right] = s_N E \| \eta_0 \|^2.
\]

We now turn to the expectation of the second term
\[
E \left[ \sum_{l=1}^{N-1} \sum_{i=1}^{N-l} k_i k_{i+l} \langle \eta_i, \eta_{i+l} \rangle \right] = \sum_{l=1}^{N-1} \sum_{i=1}^{N-l} k_i k_{i+l} E \langle \eta_i, \eta_{i+l} \rangle.
\]

29
By stationarity, \( E\langle \eta_0, \eta_\ell \rangle = E\langle \eta_{n-\ell}, \eta_n \rangle \). By Assumption 2.1, \( \eta_n \) can be approximated by \( \eta_{n,\ell} \) which is independent of \( \eta_{n-\ell} \). Therefore

\[
E\langle \eta_{n-\ell}, \eta_n \rangle = E\langle \eta_{n-\ell}, \eta_{n,\ell} \rangle + E\langle \eta_{n-\ell}, \eta_n - \eta_{n,\ell} \rangle = E\langle \eta_{n-\ell}, \eta_n - \eta_{n,\ell} \rangle,
\]

because \( E\langle \eta_{n-\ell}, \eta_{n,\ell} \rangle = 0 \). Observe that, using the Cauchy-Schwarz inequality twice,

\[
|E\langle \eta_{n-\ell}, \eta_n - \eta_{n,\ell} \rangle| \leq E\|\eta_{n-\ell}\| E\|\eta_n - \eta_{n,\ell}\| \leq \{E\|\eta_{n-\ell}\|^2\}^{1/2}\{E\|\eta_n - \eta_{n,\ell}\|^2\}^{1/2}.
\]

It follows that

\[
\left| E \left[ \sum_{l=1}^{N-1} \sum_{i=1}^{N-\ell} k_l k_{i+\ell} \langle \eta_i, \eta_{i+\ell} \rangle \right] \right| \leq \sum_{l=1}^{N-1} \sum_{i=1}^{N-\ell} k_l k_{i+\ell} \left| E \langle \eta_i, \eta_{i+\ell} \rangle \right| \leq \{E\|\eta_{n-\ell}\|^2\}^{1/2} \sum_{l=1}^{N-1} \sum_{i=1}^{N-\ell} k_l k_{i+\ell} \{E\|\eta_n - \eta_{n,\ell}\|^2\}^{1/2} \leq CN^3 \sum_{\ell=1}^{\infty} \{E\|\eta_n - \eta_{n,\ell}\|^2\}^{1/2}.
\]

Thus assuming

\[
\sum_{\ell=1}^{\infty} \{E\|\eta_n - \eta_{n,\ell}\|^2\}^{1/2} < \infty,
\]

we obtain

(5.26) \[
\left| E \left[ \sum_{l=1}^{N-1} \sum_{i=1}^{N-\ell} k_l k_{i+\ell} \langle \eta_i, \eta_{i+\ell} \rangle \right] \right| \leq O(N^3) = O(s_N).
\]

Combining (5.25) and (5.26), we see that

(5.27) \[
E\|v_N\| = \frac{1}{N s_N} E\left\| \sum_{i=1}^{N} k_i \eta_i \right\|^2 = \frac{1}{N s_N} O(s_N) = O(N^{-1}).
\]

This proves relation (5.24). \( \square \)

**Lemma 5.10** Relation (5.23) holds under the assumptions of Theorem 5.3.
Proof: Under $H_0$,

$$\hat{\gamma}_i(t, s) = \frac{1}{N} \sum_{j=i+1}^{N} e_j(t)e_{j-i}(s)$$

$$= \frac{1}{N} \sum_{j=i+1}^{N} \left[ (\eta_j(t) - \bar{\eta}(t)) - (\hat{\xi}(t) - \xi(t))k_j \right] \cdot \left[ (\bar{\eta}_{j-i}(s) - \bar{\eta}(s)) - (\hat{\xi}(s) - \xi(s))k_{j-i} \right]$$

$$= \frac{1}{N} \sum_{j=i+1}^{N} (\eta_j(t) - \bar{\eta}(t))(\eta_{j-i}(s) - \bar{\eta}(s))$$

$$- \frac{1}{Ns_N} \sum_{l=1}^{N} k_l \eta_l(s) \sum_{j=i+1}^{N} k_{j-i}(\eta_j(t) - \bar{\eta}(t))$$

$$- \frac{1}{Ns_N} \sum_{l=1}^{N} k_l \eta_l(t) \sum_{j=i+1}^{N} k_j(\eta_{j-i}(s) - \bar{\eta}(s))$$

$$+ \frac{1}{Ns_N^2} \sum_{l=1}^{N} k_l \eta_l(s) \sum_{m=1}^{N} k_m \eta_m(t) \sum_{j=i+1}^{N} k_j k_{j-i}.$$

Set

$$\bar{\gamma}_i(s, t) = \frac{1}{N} \sum_{j=i+1}^{N} (\eta_j(t) - \bar{\eta}(t))(\eta_{j-i}(s) - \bar{\eta}(s)).$$

Then

$$\sum_{i=1}^{N-1} K\left(\frac{i}{h}\right) \hat{\gamma}_i(t, s) = \sum_{i=1}^{N-1} K\left(\frac{i}{h}\right) \bar{\gamma}_i(t, s)$$

$$- \frac{1}{Ns_N} \sum_{l=1}^{N} k_l \eta_l(s) \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} K\left(\frac{i}{h}\right) k_{j-i}(\eta_j(t) - \bar{\eta}(t))$$

$$- \frac{1}{Ns_N} \sum_{l=1}^{N} k_l \eta_l(t) \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} K\left(\frac{i}{h}\right) k_j(\eta_{j-i}(s) - \bar{\eta}(s))$$

$$+ \frac{1}{Ns_N^2} \sum_{l=1}^{N} k_l \eta_l(s) \sum_{m=1}^{N} k_m \eta_m(t) \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} K\left(\frac{i}{h}\right) k_j k_{j-i}.$$

Thus, in order to prove Lemma 5.10, we must establish all of the following relations:

$$\int \int \left( \sum_{i=1}^{N-1} K\left(\frac{i}{h}\right) \bar{\gamma}_i(t, s) - \sum_{i \geq 1} E[\eta_0(s)\eta_i(t)] \right)^2 dt ds = o_P(1);$$

(5.28)
\[(5.29) \quad \iint \left( \frac{1}{N s_N} \sum_{l=1}^{N} k_l \eta_l(s) \sum_{i=1}^{N-1} K \left( \frac{i}{h} \right) \sum_{j=i+1}^{N} k_{j-i}(\eta_j(t) - \bar{\eta}(t)) \right)^2 ds dt = o_p(1); \]

\[(5.30) \quad \iint \left( \frac{1}{N s_N} \sum_{l=1}^{N} k_l \eta_l(t) \sum_{i=1}^{N-1} K \left( \frac{i}{h} \right) \sum_{j=i+1}^{N} k_{j-i}(\eta_j(s) - \bar{\eta}(s)) \right)^2 ds dt = o_p(1); \]

\[(5.31) \quad \iint \left( \frac{1}{N s_N^2} \sum_{l=1}^{N} k_l \eta_l(s) \sum_{m=1}^{N} k_m \eta_m(t) \sum_{i=1}^{N-1} K \left( \frac{i}{h} \right) \sum_{j=i+1}^{N} k_j k_{j-i} \right)^2 ds dt = o_p(1). \]

Relation (5.28) has been established by Horváth et al. (2013), so it remains to deal with the remaining three relations which are due to the estimation of the trend.

Relations (5.29) and (5.30) follow by application of similar arguments, so we display only the verification of (5.29). Observe that the left-hand side of (5.29) is equal to

\[(5.32) \quad \left( \frac{1}{N s_N} \left\| \sum_{l=1}^{N} k_l \eta_l \right\|^2 \right) \left( \frac{1}{N s_N} \left\| \sum_{i=1}^{N-1} K \left( \frac{i}{h} \right) \sum_{j=i+1}^{N} k_{j-i}(\eta_j - \bar{\eta}) \right\|^2 \right) \]

A bound for the expectation of the first factor is given in (5.27). In the second factor, the centering by $\bar{\eta}$ contributes asymptotically negligible terms, so this factor has the same order as

\[F_N = \frac{1}{N s_N} \left\| \sum_{i=1}^{N-1} K \left( \frac{i}{h} \right) \sum_{j=i+1}^{N} k_{j-i} \eta_j \right\|^2.\]

Observe that

\[E \left[ \sum_{i=1}^{N-1} K \left( \frac{i}{h} \right) \sum_{j=i+1}^{N} k_{j-i} \right]^2 = \sum_{i=1}^{N-1} \sum_{i'=1}^{N-1} K \left( \frac{i}{h} \right) K \left( \frac{i'}{h} \right) \sum_{l=1}^{N-1} \sum_{l'=1}^{N-1} k_l k_{l'} E \langle \eta_{i+l}, \eta_{i'+l'} \rangle \]

\[= E \| \eta_0 \|^2 \sum_{i=1}^{N-1} \sum_{i'=1}^{N-1} K \left( \frac{i}{h} \right) K \left( \frac{i'}{h} \right) \sum_{l=1}^{N-1} \sum_{l'=1}^{N-1} k_l k_{l'} \]

\[= O(N^3) \left( \sum_{i=1}^{N-1} K \left( \frac{i}{h} \right) \right)^2 \]

Notice that

\[\sum_{i=1}^{N-1} K \left( \frac{i}{h} \right) = h \sum_{i=1}^{N-1} \frac{1}{h} K \left( \frac{i}{h} \right) = O \left( h \int_0^1 |k(u)| du \right),\]
we see that $EF_N = O(N^{-4}N^3h^2) = O(N^{-1}h^2)$. Thus, (5.32) is of the order of (5.27),

$$O_P(N^{-1})O_P(N^{-1}h^2) = O_P\left(\frac{h^2}{N^2}\right) = o_p(1).$$

We now turn to the verification of (5.31) whose left-hand side can be written as

$$\left\{ \int \left( \frac{1}{s_N} \sum_{l=1}^N k_l \eta_l(t) \right)^2 dt \right\}^2 \left\{ \sum_{i=1}^{N-1} K\left( \frac{i}{h} \right) \frac{1}{N} \sum_{j=i+1}^N k_j k_{j-i} \right\}^2$$

$$= \frac{N^2}{S_N^2} \|v_N\|^2 \left\{ \sum_{i=1}^{N-1} K\left( \frac{i}{h} \right) \frac{1}{N} \sum_{j=i+1}^N k_j k_{j-i} \right\}^2.$$

Using (5.27), we see that the order of the above expression is

$$O\left(\frac{N^2}{N^6}\right)O_P(N^{-2})\{O_P(hN^2)\}^2 = O(N^{-4})O_P(N^{-2})O(h^2N^4) = O_P\left(\frac{h^2}{N^2}\right) = o_P(1).$$

This completes the proof of Lemma 5.10. □

### 5.4 Proof of Theorem 3.3

The proof of Theorem 3.3 is constructed from several lemmas.

**Lemma 5.11** Under the alternative (1.2), for the functional slope estimate $\hat{\xi}$ defined by (2.9),

$$N^{3/2}(\hat{\xi}(t) - \xi(t)) = \frac{1}{N^{-3}s_N} \left\{ \left( \frac{N-1}{2N} \right) V_N(1,t) - \frac{1}{N} \sum_{k=1}^{N-1} V_N\left( \frac{k}{N},t \right) \right\}$$

$$+ \frac{1}{N^{-3}s_N} \left\{ \sum_{n=1}^{N} \frac{n}{N} Y_N\left( \frac{n}{N},t \right) - \left( \frac{N+1}{2N} \right) \sum_{n=1}^{N} Y_N\left( \frac{n}{N},t \right) \right\},$$

where $V_N(x,t)$ is the partial sum process of the errors $\eta_n$ defined in (2.2) and $Y_N(x,t)$ is the partial sum process of the random walk errors $u_n$ defined by

$$Y_N(x,t) = \frac{1}{\sqrt{N}} \sum_{n=1}^{[Nx]} u_n(t).$$
Proof: Recall that \( k_n = n - (N + 1)/2 \), \( \sum_{n=1}^{N} k_n = 0 \), \( \sum_{n=1}^{N} nk_n = s_N \). Therefore, under alternative (1.2),

\[
\dot{\xi}(t) - \xi(t) = \frac{1}{s_N} \sum_{n=1}^{N} k_n X_n(t) - \xi(t)
\]

\[
= \frac{1}{s_N} \sum_{n=1}^{N} k_n \left( \mu(t) + n\xi(t) + \sum_{i=1}^{n} u_i(t) + \eta_n(t) \right) - \xi(t)
\]

\[
= \frac{1}{s_N} \left\{ \mu(t) \sum_{n=1}^{N} k_n + \xi(t) \sum_{n=1}^{N} nk_n \right\}
\]

\[
+ \frac{1}{s_N} \sum_{n=1}^{N} k_n \eta_n(t) - \xi(t) + \frac{1}{s_N} \sum_{n=1}^{N} k_n \sum_{i=1}^{n} u_i(t)
\]

\[
= \frac{1}{s_N} \xi(t) s_N + \frac{1}{s_N} \sum_{n=1}^{N} k_n \eta_n(t) - \xi(t) + \frac{1}{s_N} \sum_{n=1}^{N} k_n \sum_{i=1}^{n} u_i(t)
\]

\[
= \frac{1}{s_N} \sum_{n=1}^{N} k_n \eta_n(t) + \frac{1}{s_N} \sum_{n=1}^{N} k_n \sum_{i=1}^{n} u_i(t).
\]

Using definition (5.33),

\[
\frac{N^{-3/2}}{N^{-3/2}} \frac{1}{s_N} \sum_{n=1}^{N} k_n \sum_{i=1}^{n} u_i(t) = \frac{N^{-3/2}}{N^{-3/2}} \left\{ \frac{N}{N} \sum_{n=1}^{N} \sum_{i=1}^{n} u_i(t) - \frac{N+1}{2} \sum_{n=1}^{N} \sum_{i=1}^{n} u_i(t) \right\}
\]

\[
= \frac{1}{N^{-3/2}} \left\{ \sum_{n=1}^{N} \frac{1}{N} \sum_{i=1}^{n} u_i(t) - \frac{N+1}{2N} \sum_{n=1}^{N} \frac{1}{N} \sum_{i=1}^{n} u_i(t) \right\}
\]

\[
= \frac{1}{N^{-3/2}} \left\{ \sum_{n=1}^{N} \frac{1}{N} Y_N \left( \frac{n}{N}, t \right) - \frac{N+1}{2N} \frac{1}{N} \sum_{n=1}^{N} Y_N \left( \frac{n}{N}, t \right) \right\}.
\]

The claim follows from the above relations and Lemma 5.3.

□
Lemma 5.12 Under the alternative, $Z_N(x,t)$ defined in (3.2) can be expressed as

$$Z_N(x,t) = V_N(x,t) - \frac{[Nx]}{N} V_N(1,t)$$

$$- \frac{1}{2} N^{-3} s_N \left\{ \frac{N-1}{2N} V_N(1,t) - \frac{1}{N} \sum_{k=1}^{N-1} V_N\left(\frac{k}{N}, t\right) \right\} \left( \frac{[Nx]}{N} \left( \frac{[Nx]}{N} - 1 \right) \right)$$

$$+ \sum_{n=1}^{[Nx]} Y_N\left(\frac{n}{N}, t\right) - \frac{[Nx]}{N} \sum_{n=1}^{N} Y_N\left(\frac{n}{N}, t\right)$$

$$- \frac{1}{2} N^{-3} s_N \left\{ \sum_{n=1}^{N} n Y_N\left(\frac{n}{N}, t\right) - \frac{N+1}{2N} \sum_{n=1}^{N} Y_N\left(\frac{n}{N}, t\right) \right\} \left( \frac{[Nx]}{N} \left( \frac{[Nx]}{N} - 1 \right) \right).$$

Proof: Under $H_A$, the partial sum process $S_N(x,t)$ can be expressed as

$$S_N(x,t) = \frac{1}{\sqrt{N}} \sum_{n=1}^{[Nx]} X_n(t)$$

$$= \frac{1}{\sqrt{N}} \sum_{n=1}^{[Nx]} \left( \mu(t) + n \xi(t) + \sum_{i=1}^{n} u_i(t) + \eta_n(t) \right)$$

$$= \left( \frac{1}{\sqrt{N}} \mu(t) [Nx] + \frac{1}{\sqrt{N}} \xi(t) \frac{[Nx](|Nx|+1)}{2} + \sum_{n=1}^{[Nx]} \frac{1}{\sqrt{N}} \sum_{i=1}^{n} u_i(t) + \frac{1}{\sqrt{N}} \sum_{n=1}^{[Nx]} \eta_n(t) \right)$$

$$= \left( \frac{1}{\sqrt{N}} \mu(t) [Nx] + \frac{1}{\sqrt{N}} \xi(t) \frac{[Nx](|Nx|+1)}{2} + \sum_{n=1}^{[Nx]} Y_N\left(\frac{n}{N}, t\right) + V_N(x,t) \right) .$$

By Lemma 5.8,

$$Z_N(x,t) = S_N(x,t) - \frac{[Nx]}{N} S_N(1,t) - \frac{1}{2} N^{3/2} \hat{\xi}(t) \left( \frac{[Nx]}{N} \left( \frac{[Nx]}{N} - 1 \right) \right)$$

$$= \left( \frac{1}{\sqrt{N}} \mu(t) [Nx] + \frac{1}{\sqrt{N}} \xi(t) \frac{[Nx](|Nx|+1)}{2} + \sum_{n=1}^{[Nx]} Y_N\left(\frac{n}{N}, t\right) + V_N(x,t) \right)$$

$$- \frac{[Nx]}{N} \left( \frac{1}{\sqrt{N}} \mu(t) N + \frac{1}{\sqrt{N}} \xi(t) \frac{N(N+1)}{2} + \sum_{n=1}^{N} Y_N\left(\frac{n}{N}, t\right) + V_N(1,t) \right)$$

$$- \frac{1}{2} N^{3/2} \hat{\xi}(t) \left( \frac{[Nx]}{N} \left( \frac{[Nx]}{N} - 1 \right) \right).$$
Therefore,
\[
Z_N(x, t) = \frac{1}{\sqrt{N}} \mu(t) [N x] - \frac{1}{\sqrt{N}} \mu(t) [N x] + \frac{1}{2} N^{3/2} \xi(t) \left( \frac{[N x]}{N} \left( \frac{[N x]}{N} - 1 \right) \right) \\
+ \sum_{n=1}^{[N x]} Y_N \left( \frac{n}{N}, t \right) - \frac{[N x]}{N} \sum_{n=1}^{N} Y_N \left( \frac{n}{N}, t \right) + V_N(x, t) - \frac{[N x]}{N} V_N(1, t) \\
- \frac{1}{2} N^{3/2} \hat{\xi}(t) \left( \frac{[N x]}{N} \left( \frac{[N x]}{N} - 1 \right) \right) \\
= \sum_{n=1}^{[N x]} Y_N \left( \frac{n}{N}, t \right) - \frac{[N x]}{N} \sum_{n=1}^{N} Y_N \left( \frac{n}{N}, t \right) + V_N(x, t) - \frac{[N x]}{N} V_N(1, t) \\
- \frac{1}{2} N^{3/2} \hat{\xi}(t) - \hat{\xi}(t) \left( \frac{[N x]}{N} \left( \frac{[N x]}{N} - 1 \right) \right).
\]

Then reexpressing \( N^{3/2} \hat{\xi}(t) - \hat{\xi}(t) \), we get the desired expression. □

Since the \( u_i \) satisfy Assumption 2.1, an analog of Theorem 5.2 holds, i.e. there exist Gaussian processes \( \Lambda_N \) equal in distribution to
\[
\Lambda(x, t) = \sum_{i=1}^{\infty} \tau_i^{1/2} W_i(x) \psi_i(t),
\]
where \( \tau_i, \psi_i \) are, respectively, the eigenvalues and the eigenfunctions of the kernel (3.5).

Moreover, for the partial sum process \( Y_N \) defined by (5.33),
\[
l_n = \sup_{0 \leq x \leq 1} \| Y_N(x, \cdot) - \Lambda_N(x, \cdot) \| = o_p(1).
\]

**Lemma 5.13** Under the alternative, the following convergence holds
\[
\sup_{0 \leq x \leq 1} \| N^{-1} Z_N^A(x, \cdot) - \Delta_N^0(x, \cdot) \| \overset{P}{\to} 0,
\]
where the processes \( Z_N^A(\cdot, \cdot) \) and \( \Delta_N^0(\cdot, \cdot) \) are respectively defined by
\[
Z_N^A(x, t) = \sum_{n=1}^{[N x]} Y_N \left( \frac{n}{N}, t \right) - \frac{[N x]}{N} \sum_{n=1}^{N} Y_N \left( \frac{n}{N}, t \right) \\
- \frac{1}{2} N^{-3/2} \sum_{n=1}^{N} \frac{n}{N} Y_N \left( \frac{n}{N}, t \right) - \frac{N + 1}{2N} \sum_{n=1}^{N} Y_N \left( \frac{n}{N}, t \right) \left( \frac{[N x]}{N} \left( \frac{[N x]}{N} - 1 \right) \right) \\
\text{and}
\]
\[
\Delta_N^0(x, t) = \int_0^x \Lambda_N(y, t) dy + (3x^2 - 4x) \int_0^1 \Lambda_N(y, t) dy + (-6x^2 + 6x) \int_0^1 y \Lambda_N(y, t) dy.
\]

36
Proof: Notice that \( N^{-1}Z^A_N(x,t) \) can be expressed as
\[
\frac{1}{N}Z^A_N(x,t) = \sum_{n=1}^{\lfloor Nx \rfloor} Y_N\left(\frac{n}{N}, t\right) \frac{1}{N} + f_N(x) \sum_{n=1}^{N} Y_N\left(\frac{n}{N}, t\right) \frac{1}{N} + g_N(x) \sum_{n=1}^{N} \frac{n}{N} Y_N\left(\frac{n}{N}, t\right) \frac{1}{N}
\]
where
\[
f_N(x) = \frac{1}{2N^{-3}s_N} \left( \frac{N+1}{2N} \right) \left( \frac{|N_x|}{N} - 1 \right) - \frac{|N_x|}{N}
\]
and
\[
g_N(x) = \frac{1}{2N^{-3}s_N} \left( \frac{|N_x|}{N} \right) \left( 1 - \frac{|N_x|}{N} \right).
\]
By the triangle inequality,
\[
\|N^{-1}Z_N(x, \cdot) - \Delta^0_N(x, \cdot)\| \leq \left\| \sum_{n=1}^{\lfloor Nx \rfloor} Y_N\left(\frac{n}{N}, \cdot\right) \frac{1}{N} - \int_0^x \Lambda_N(y, \cdot)dy \right\|
\]
\[
+ \left\| f_N(x) \sum_{n=1}^{N} Y_N\left(\frac{n}{N}, \cdot\right) \frac{1}{N} - (3x^2 - 4x) \int_0^1 \Lambda_N(y, \cdot)dy \right\|
\]
\[
+ \left\| g_N(x) \sum_{n=1}^{N} \frac{n}{N} Y_N\left(\frac{n}{N}, \cdot\right) \frac{1}{N} - (-6x^2 + 6x) \int_0^1 y\Lambda_N(y, \cdot)dy \right\|.
\]
Thus, Lemma 5.1 will be proven once we have established the following relations:

(5.38) \[
\sup_{0 \leq x \leq 1} \left\| \sum_{n=1}^{\lfloor Nx \rfloor} Y_N\left(\frac{n}{N}, \cdot\right) \frac{1}{N} - \int_0^x \Lambda_N(y, \cdot)dy \right\| \overset{p}{\to} 0;
\]

(5.39) \[
\sup_{0 \leq x \leq 1} \left\| f_N(x) \sum_{n=1}^{N} Y_N\left(\frac{n}{N}, \cdot\right) \frac{1}{N} - (3x^2 - 4x) \int_0^1 \Lambda_N(y, \cdot)dy \right\| \overset{p}{\to} 0;
\]

(5.40) \[
\sup_{0 \leq x \leq 1} \left\| g_N(x) \sum_{n=1}^{N} \frac{n}{N} Y_N\left(\frac{n}{N}, \cdot\right) \frac{1}{N} - (-6x^2 + 6x) \int_0^1 y\Lambda_N(y, \cdot)dy \right\| \overset{p}{\to} 0.
\]
Relations (5.39) and (5.40) follow from arguments fully analogous to those used in the proof of Lemma 5.7. The verification of (5.38) is not difficult either. Observe that
\[
\sum_{n=1}^{\lfloor Nx \rfloor} Y_N\left(\frac{n}{N}, \cdot\right) \frac{1}{N} = \int_0^x Y_N(y, \cdot)dy - r_N(x),
\]
37
where
\[ r_N(x) = \int_{[N x]/N}^x Y_N(y, \cdot)dy. \]

Relation
\[ \sup_{0 \leq x \leq 1} \left\| \int_0^x Y_N(y, \cdot)dy - \int_0^x A_N(y, \cdot)dy \right\| = o_p(1) \]
follows from (5.35) and the contractive property of the integral, which also implies that
\[ \sup_{0 \leq x \leq 1} \left\| r_N(x) \right\| \leq \sup_{0 \leq x \leq 1} \int_{[N x]/N}^x \left\| Y_N(y, \cdot) \right\|dy \]
\[ \leq \frac{1}{N} \sup_{0 \leq x \leq 1} \left\| Y_N(y, \cdot) \right\| = O_p(N^{-1}). \]

This completes the proof of Lemma 5.13.

\[ \square \]

**LEMMA 5.14** Under the alternative, the following convergence holds
\[ \sup_{0 \leq x \leq 1} \left\| N^{-1}Z_N(x, \cdot) - \Delta_N^0(x, \cdot) \right\| \overset{P}{\to} 0, \]
where the process \( Z_N(\cdot, \cdot) \) is defined in (3.2) and the process \( \Delta_N^0(\cdot, \cdot) \) in (5.37).

**PROOF:** Under the alternative, Lemma 5.12 implies that, \( Z_N(x, t) \) can be expressed as
\[ Z_N(x, t) = Z_N^0(x, t) + Z_N^A(x, t), \]
where
\[ Z_N^0(x, t) = V_N(x, t) - \frac{|N x|}{N} V_N(1, t) \]
\[ - \frac{1}{2} \frac{1}{N^{-3} s_N} \left\{ \left( \frac{N - 1}{2N} \right) V_N(1, t) - \frac{1}{N} \sum_{k=1}^{N-1} V_N \left( \frac{k}{N}, t \right) \right\} \left( \frac{|N x|}{N} \left( \frac{|N x|}{N} - 1 \right) \right) \]
and \( Z_N^A(x, t) \) is defined by (5.36). Notice under the null hypothesis, Lemma 5.4 implies that \( Z_N^0(x, t) = Z_N(x, t) \). Hence from Lemma 5.7, \( \sup_{0 \leq x \leq 1} \left\| Z_N^0(x, \cdot) - \Gamma_N^0(x, \cdot) \right\| \overset{P}{\to} 0, \) implying
\[ \sup_{0 \leq x \leq 1} \left\| \frac{1}{N} Z_N^0(x, \cdot) \right\| = \frac{1}{N} O_p(1). \]

Thus the claim follows from Lemma 5.13 and the triangle inequality. \( \square \)
Lemma 5.15 Consider the process \( \Lambda(\cdot, \cdot) \) defined by (5.34) and set
\[
\Delta^0(x, t) = \int_0^x \Lambda(y, t)dy + (3x^2 - 4x) \int_0^1 \Lambda(y, t)dy + (-6x^2 + 6x) \int_0^1 y \Lambda(y, t)dy.
\]
Then
\[
\int_0^1 \|\Delta^0(x, \cdot)\|^2dx = \sum_{i=1}^{\infty} \tau_i \int_0^1 \Delta_i^2(x)dx,
\]
where \( \tau_1, \tau_2, \ldots \) are eigenvalues of the long-run covariance function of the \( u_i \), i.e. (3.5), and \( \Delta_1, \Delta_2, \ldots \) are independent copies of the process \( \Delta \) defined in (3.4).

Proof: Expansion (5.34) implies that
\[
\Delta^0(x, t) = \int_0^x \Lambda(y, t)dy + (3x^2 - 4x) \int_0^1 \Lambda(y, t)dy + (-6x^2 + 6x) \int_0^1 y \Lambda(y, t)dy
\]
\[
= \int_0^x \sum_{i=1}^{\infty} \sqrt{\tau_i} W_i(y) \psi_i(t)dy + (3x^2 - 4x) \int_0^1 \sum_{i=1}^{\infty} \sqrt{\tau_i} W_i(y) \psi_i(t)dy
\]
\[
+ (-6x^2 + 6x) \int_0^1 y \sum_{i=1}^{\infty} \sqrt{\tau_i} W_i(y) \psi_i(t)dy
\]
\[
= \sum_{i=1}^{\infty} \sqrt{\tau_i} \left\{ \int_0^x W_i(y)dy + (3x^2 - 4x) \int_0^1 W_i(y)dy
\right\} \psi_i(t)
\]
\[
= \sum_{i=1}^{\infty} \sqrt{\tau_i} \Delta_i(x) \psi_i(t).
\]
The claim then follows from the orthonormality of the eigenfunctions \( \phi_i \).

\[\square\]

Proof of Theorem 3.3: Recall that the test statistic \( R_N \) is defined by \( R_N = \int \int Z_N^2(x, t)dxdt \), with \( Z_N \) defined by (3.2). We want to show that under the alternative model (1.2),
\[
\frac{1}{N^2} R_N \xrightarrow{d} \sum_{i=1}^{\infty} \tau_i \int_0^1 \Delta_i^2(x)dx,
\]
where \( \Delta_1, \Delta_2, \ldots \) are independent copies of the process defined by (3.4) and \( \tau_1, \tau_2, \ldots \) are the eigenvalues of the long-run covariance kernel (3.5). By Lemma 5.14,
\[
\rho(N^{-1} Z_N(x, \cdot), \Delta_N^0(x, \cdot)) = \sup_{0 \leq x \leq 1} \|N^{-1} Z_N(x, \cdot) - \Delta_N^0(x, \cdot)\| \xrightarrow{P} 0.
\]

39
By construction, $\Delta^0_N \overset{d}{=} \Delta^0$, so Theorem 5.1 implies that $N^{-1}Z_N \overset{d}{\to} \Delta^0$. By Lemma 5.15,
\[
\int \int (\Delta^0(x,t))^2 dx dt \overset{d}{=} \sum_{i=1}^{\infty} \lambda_i \int_{0}^{1} \Delta^2_i(x) dx.
\]
Thus by the continuous mapping theorem,
\[
\frac{1}{N^2} R_N = \int \int (N^{-1}Z_N(x,t))^2 dx dt \overset{d}{\to} \sum_{i=1}^{\infty} \lambda_i \int \Delta^2_i(x) dx.
\]
\[\square\]

References


