



Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Testing stationarity of functional time series[☆]

Lajos Horváth^{a,*}, Piotr Kokoszka^b, Gregory Rice^a

^a Department of Mathematics, University of Utah, Salt Lake City, UT 84112-0090, USA

^b Department of Statistics, Colorado State University, Fort Collins, CO 80523-1877, USA

ARTICLE INFO

Article history:

Received 29 January 2013

Received in revised form

28 June 2013

Accepted 6 November 2013

Available online xxx

JEL classification:

C32

C50

C58

Keywords:

Change point

Functional data

Integrated time series

Intraday price curves

Test of stationarity

ABSTRACT

Economic and financial data often take the form of a collection of curves observed consecutively over time. Examples include, intraday price curves, yield and 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 such data, is whether these curves, perhaps suitably transformed, form a stationary functional time series. This paper formalizes the assumption of stationarity in the context of functional time series and proposes several procedures to test the null hypothesis of stationarity. The tests are nontrivial extensions of the broadly used tests in the KPSS family. The properties of the tests under several alternatives, including change-point and $I(1)$, are studied, and new insights, present only in the functional setting are uncovered. The theory is illustrated by a small simulation study and an application to intraday price curves.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Over the last two decades, functional data analysis has become an important and steadily growing area of statistics. Very early on, major applications and theoretical developments pertained to functions observed consecutively over time, for example one function per day, or one function per year, with many of these data sets arising in econometric research. The main model employed for such series has been the functional autoregressive model of order one, which has received a great deal of attention, see Bosq (2000), Antoniadis and Sapatinas (2003), Antoniadis et al. (2006) and Kargin and Onatski (2008), among many others. More recent research has considered functional time series which have non-linear dependence structure, see Hörmann and Kokoszka (2010), Gabrys et al. (2010), Horváth et al. (2013), Hörmann et al. (2013), as well as the review of Hörmann and Kokoszka (2012) and Chapter 16 of Horváth and Kokoszka (2012). As in traditional (scalar and vector) time series analysis, the underlying assumption for inference in such models is stationarity. Stationarity is also required for functional dynamic regression models like those studied by Hays

et al. (2012) and Kokoszka et al. (2013), for bootstrap and resampling methods for functional time series, see McMurry and Politis (2010) and for the functional analysis of volatility, see Müller et al. (2011).

Testing stationarity received due attention as soon as fundamental time series modeling principles have emerged. Early work includes Grenander and Rosenblatt (1957), Granger and Hatanaka (1964) and Priestley and Subba Rao (1969). The methods considered by these authors rest on the spectral analysis which dominated the field of time series analysis at that time. While such approaches remain useful, see Dwivedi and Subba Rao (2011), the spectral analysis of nonstationary functional time series has not been developed to a point where usable extensions could be readily derived. We note however the recent work of Panaretos and Tavakoli (2013a), Panaretos and Tavakoli (2013b) and Hörmann et al. (2013) who advance the spectral analysis of stationary functional time series.

We follow a time domain approach introduced in the seminal paper of Kwiatkowski et al. (1992) which is now firmly established in econometric theory and practice, and has been extended in many directions. 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

[☆] Research supported by NSF grants DMS 1305858 and DMS 0931948.

* Corresponding author. Tel.: +1 801 581 8159.

E-mail address: horvath@math.utah.edu (L. Horváth).

root. Since such tests have low power in samples of sizes occurring in many applications, Kwiatkowski et al. (1992) proposed that stationarity should be considered as the null hypothesis (they used a broader definition which allowed for deterministic trends), and the unit root should be the alternative. Rejection of the null of 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. At present, both the augmented Dickey–Fuller test and the KPSS test, as well as its robust version of de Jong et al. (1997), are typically applied to the same series to get a fuller picture. They are available in many packages, including R and Matlab implementations. The work of Lo (1991) is also very relevant to our approach. His contribution is crucial because he showed that 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, which leads to these distributions only if the observations are independent.

This paper seeks to develop a general methodology for testing the assumption that a functional time series to be modeled is indeed stationary and weakly dependent. Such a test should be applied before fitting one of the known stationary models (all of them are weakly dependent). In many cases, it will be applied to functions transformed to remove seasonality or obvious trends, or to model residuals. At present only CUSUM change point tests are available for functional time series, see Berkes et al. (2009), Horváth et al. (2010) and Zhang et al. (2011). These tests have high power to detect abrupt changes in the stochastic structure of a functional time series, either the mean or the covariance structure. Our objective is to develop more general tests of stationarity which also have high power against integrated and other alternatives.

It is difficult to explain the main contribution of this paper without introducing the required notation, but we wish to highlight in this paragraph the main difficulties which are encountered in the transition from the scalar or vector to the functional case. A stationary functional time series can be represented as

$$X_n(t) = \mu(t) + \sum_{j=1}^{\infty} \sqrt{\lambda_j} \xi_{jn} v_j(t),$$

where n is the time index that counts the functions (referring e.g. to a day), and t is the (theoretically continuous) argument of each function. The mean function μ and the functional principal components v_j are unknown deterministic functions which depend on the stochastic structure of the series $\{X_n\}$, and which are estimated by random functions $\hat{\mu}$ and \hat{v}_j . If $\{X_n\}$ is not stationary, one can still compute the estimators $\hat{\mu}$ and \hat{v}_j , but they will not converge to μ or v_j because these population quantities will not exist then. Thus the use of a data driven basis system v_j represents an aspect which is not encountered in the theory of scalar or vector valued tests. Therefore, after defining meaningful extensions to the functional setting, we must develop a careful analysis of the behavior of the tests under alternatives.

The paper is organized as follows. Section 2 formalizes the null hypothesis of stationarity and weak dependence of functional time series, introduces the tests, and explores their asymptotic properties under the null hypothesis. In Section 3, we turn to the behavior of the tests under several alternatives. Section 4 explains the details of the implementation, and contains the results of a simulation study, while Section 5 illustrates the properties of the tests by an application to intraday price curves. Appendices A and B contain, respectively, the proofs of the results stated in Sections 2 and 3.

2. Assumptions and test statistics

Linear functional time series, in particular functional AR(1) processes, have the form $X_n = \sum_j \Psi_j(\varepsilon_{n-j})$, where the ε_i are iid error functions, and the Ψ_j are bounded linear operators acting on the space of square integrable functions. In this paper, we assume merely that $X_n = \tilde{f}(\varepsilon_n, \varepsilon_{n-1}, \dots)$, for some, possibly nonlinear, function \tilde{f} . The operators Ψ_j or the function \tilde{f} arise as solutions to structural equations, very much like in the univariate econometric modeling, see e.g. Teräsvirta et al. (2010). For the functional autoregressive process, the norms of the operators Ψ_j decay exponentially fast. For the more general nonlinear moving averages, the rate at which the dependence of X_n on past errors ε_{n-j} decays with j can be quantified by a condition known as L^p - m -approximability stated in assumptions (2.1)–(2.4) below. In both cases, these functional models can be said to be in a class which is customarily referred to as weakly dependent or short memory time series. It is convenient to state the conditions for the error process, which we denote by $\eta = \{\eta_j\}_{-\infty}^{\infty}$, and which will be used to formulate the null and alternative hypotheses.

Throughout the paper, L^2 denotes the Hilbert space of square integrable functions on the unit interval $[0, 1]$ with the usual inner product $\langle \cdot, \cdot \rangle$ and the norm $\|\cdot\|$ it generates, \int means \int_0^1 .

η forms a sequence of Bernoulli shifts, i.e.

$$\eta_j = f(\varepsilon_j, \varepsilon_{j-1}, \dots) \tag{2.1}$$

for some measurable function $f : S^{\infty} \mapsto L^2$

and iid functions $\varepsilon_j, -\infty < j < \infty$,

with values in a measurable space S ,

$\varepsilon_j(t) = \varepsilon_j(t, \omega)$ is jointly measurable in (t, ω) ,

$$-\infty < j < \infty, \tag{2.2}$$

$E\eta_0(t) = 0$ for all t , and $E\|\eta_0\|^{2+\delta} < \infty$,

$$\text{for some } 0 < \delta < 1, \tag{2.3}$$

and

the sequence $\{\eta_n\}_{n=-\infty}^{\infty}$ can be approximated by

$$\ell\text{-dependent sequences } \{\eta_{n,\ell}\}_{n=-\infty}^{\infty} \text{ in the sense that} \tag{2.4}$$

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

where $\eta_{n,\ell}$ is defined by $\eta_{n,\ell} = g(\varepsilon_n, \varepsilon_{n-1}, \dots, \varepsilon_{n-\ell+1}, \varepsilon_{n,\ell}^*)$,

$\varepsilon_{n,\ell}^* = (\varepsilon_{n,\ell,n-\ell}^*, \varepsilon_{n,\ell,n-\ell-1}^*, \dots)$, where the $\varepsilon_{n,\ell,k}^*$'s

are independent copies of ε_0 ,

independent of $\{\varepsilon_i, -\infty < i < \infty\}$.

Assumptions similar to those stated above have been used extensively in recent theoretical work, as all stationary time series models in practical use can be represented as Bernoulli shifts, see Wu (2005), Shao and Wu (2007), Aue et al. (2009), Hörmann and Kokoszka (2010), among many other contributions. They have been used in econometric research even earlier, and the work of Pötscher and Prucha (1997) contributed to their popularity. Bernoulli shifts are stationary by construction; weak dependence is quantified by the summability condition in (2.4) which intuitively states that the function g decays so fast that the impact of shocks far back in the past is so small that they can be replaced by their independent copies, with only a small change in the distribution of the process.

We wish to test

$$H_0 : X_i(t) = \mu(t) + \eta_i(t), \quad 1 \leq i \leq N, \text{ where } \mu \in L^2.$$

The mean function μ is unknown. The null hypothesis is that the functional time series is stationary and weakly dependent, with the structure of dependence quantified by conditions (2.1)–(2.4).

The most general alternative is that H_0 does not hold, but some profound insights into the behavior of the tests can be obtained by considering some specific alternatives. We focus on the following.

Change point alternative:

$$H_{A,1} : X_i(t) = \mu(t) + \delta(t)I\{i > k^*\} + \eta_i(t), \quad 1 \leq i \leq N,$$

with some integer $1 \leq k^* < N$.

The mean function $\mu(t)$, the size of the change $\delta(t)$, and the time of the change, k^* , are all unknown parameters. We assume that the change occurs away from the end points, i.e.

$$k^* = \lfloor N\tau \rfloor \quad \text{with some } 0 < \tau < 1. \tag{2.5}$$

Integrated alternative:

$$H_{A,2} : X_i(t) = \mu(t) + \sum_{\ell=1}^i \eta_\ell(t), \quad 1 \leq i \leq N.$$

Deterministic trend alternative:

$$H_{A,3} : X_i(t) = \mu(t) + g(i/N)\delta(t) + \eta_i(t), \quad 1 \leq i \leq N \tag{2.6}$$

where

$$g(t) \text{ is a piecewise Lipschitz continuous function on } [0, 1]. \tag{2.7}$$

The trend alternative includes various change point alternatives, including $H_{A,1}$, but also those in which change can be gradual. It also includes the polynomial trend alternative, if $g(u) = u^\alpha$.

We emphasize that both under the null hypothesis and all alternatives, the mean function $\mu(t)$ is unknown.

The tests we propose can be shown to be consistent against any other sufficiently large departures from stationarity and weak dependence. In particular, functional long memory alternatives could be considered as well, as studied in the scalar case by Giraitis et al. (2003). Since long memory functional processes have not been considered in any applications yet, we do not pursue this direction at this point.

In the remainder of this section, we consider two classes of tests, those based on the curves themselves, and those based on the finite dimensional projections of the curves on the functional principal components. As will become clear, the tests of the two types are related.

2.1. Fully functional tests

Our approach is based on two tests statistics. The first is

$$T_N = \iint Z_N^2(x, t) dt dx,$$

where

$$Z_N(x, t) = S_N(x, t) - xS_N(1, t), \quad 0 \leq x, t \leq 1,$$

with

$$S_N(x, t) = N^{-1/2} \sum_{i=1}^{\lfloor Nx \rfloor} X_i(t), \quad 0 \leq x, t \leq 1.$$

The second test statistic is

$$M_N = T_N - \int \left(\int Z_N(x, t) dx \right)^2 dt$$

$$= \iint \left(Z_N(x, t) - \int Z_N(y, t) dy \right)^2 dx dt.$$

If $X_i(t) = X_i$, i.e. if the data are scalars (or constant functions on $[0, 1]$), the statistic T_N is the numerator of the KPSS statistic

of Kwiatkowski et al. (1992), and M_N is the numerator of the V/S statistic of Giraitis et al. (2003) who introduced centering to reduce the variability of the KPSS statistic and to increase power against “changes in variance” which are a characteristic of long memory in volatility. As pointed out by Lo (1991), to obtain parameter free limits under the null, statistics of this type must be divided by the long run variance. We now proceed with the suitable definitions in the functional case.

The null limit distributions of T_N and M_N depend on the eigenvalues of the long-run covariance function of the errors:

$$C(t, s) = E\eta_0(t)\eta_0(s) + \sum_{\ell=1}^{\infty} E\eta_0(t)\eta_\ell(s) + \sum_{\ell=1}^{\infty} E\eta_0(s)\eta_\ell(t). \tag{2.8}$$

It is proven in Horváth et al. (2013) that the series in (2.8) is convergent in L^2 . The function $C(t, s)$ is positive definite, and therefore there exist $\lambda_1 \geq \lambda_2 \geq \dots \geq 0$ and orthonormal functions $\varphi_i(t)$, $0 \leq t \leq 1$, satisfying

$$\lambda_i \varphi_i(t) = \int C(t, s) \varphi_i(s) ds, \quad 1 \leq i < \infty. \tag{2.9}$$

The following theorem specifies limit distributions of T_N and M_N under the stationarity null hypothesis. Throughout the paper, B_1, B_2, \dots are independent Brownian bridges.

Theorem 2.1. *If assumptions (2.1)–(2.4) and H_0 hold, then*

$$T_N \xrightarrow{d} \sum_{i=1}^{\infty} \lambda_i \int B_i^2(x) dx \tag{2.10}$$

and

$$M_N \xrightarrow{d} \sum_{i=1}^{\infty} \lambda_i \int \left(B_i(x) - \int B_i(y) dy \right)^2 dx. \tag{2.11}$$

According to Theorem A.1, under assumptions (2.1)–(2.4) the sum $\sum_{i=1}^{\infty} \lambda_i$ is finite, and therefore the variables T_0 and M_0 are finite with probability one.

Theorem 2.1 shows, in particular, that for functional time series a simple normalization with a long-run variance is not possible, and approaches involving the estimation of all large eigenvalues must be employed. The eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots$ can be easily estimated under the null hypothesis because then

$$C(t, s) = \text{cov}(X_0(t), X_0(s)) + \sum_{i=1}^{\infty} [\text{cov}(X_0(t), X_i(s)) + \text{cov}(X_0(s), X_i(t))],$$

so we can use the kernel estimator \hat{C}_N of Horváth et al. (2013) defined as

$$\hat{C}_N(t, s) = \hat{\gamma}_0(t, s) + \sum_{i=1}^{N-1} K\left(\frac{i}{h}\right) (\hat{\gamma}_i(t, s) + \hat{\gamma}_i(s, t)), \tag{2.12}$$

where

$$\hat{\gamma}_i(t, s) = \frac{1}{N} \sum_{j=i+1}^N (X_j(t) - \bar{X}_N(t)) (X_{j-i}(s) - \bar{X}_N(s))$$

with

$$\bar{X}_N(t) = \frac{1}{N} \sum_{i=1}^N X_i(t).$$

The kernel K in the definition of \hat{C}_N satisfies the following conditions:

$$K(0) = 1, \tag{2.13}$$

$$K(u) = 0 \quad \text{if } u > c \text{ with some } c > 0, \tag{2.14}$$

and

K is continuous on $[0, c]$, where c is given in (2.14). (2.15)

The window (or smoothing bandwidth) h must satisfy only

$$h = h(N) \rightarrow \infty \quad \text{and} \quad \frac{h(N)}{N} \rightarrow 0, \quad \text{as } N \rightarrow \infty. \quad (2.16)$$

Now the estimators for the eigenvalues and eigenfunctions are defined by

$$\hat{\lambda}_i \hat{\varphi}_i(t) = \int \hat{C}_N(t, s) \hat{\varphi}_i(s) ds, \quad 1 \leq i \leq N,$$

where $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \dots$ are the empirical eigenvalues and $\hat{\varphi}_1, \hat{\varphi}_2, \dots$ are the corresponding orthonormal eigenfunctions. We can thus approximate the limits in Theorem 2.1 with

$$\sum_{i=1}^d \hat{\lambda}_i \int B_i^2(x) dx \quad \text{and} \quad \sum_{i=1}^d \hat{\lambda}_i \int \left(B_i(x) - \int B_i(y) dy \right)^2 dx,$$

where d is suitably large. The details are presented in Section 4. We note that the $\hat{\lambda}_i$ and the $\hat{\varphi}_i$ are consistent estimators only under H_0 . Their behavior under the alternatives is complex. It is studied in Section 3.

2.2. Tests based on projections

Theorem 2.1 leads to asymptotic distributions depending on the eigenvalues λ_i , which can collectively be viewed as an analog of the long-run variance. In this section, we will see that by projecting on the eigenfunctions $\hat{\varphi}_i$ it is possible to construct statistics whose limit null distributions are parameter free. This procedure is a functional analog of dividing by an estimator of the long-run variance.

To have uniquely defined (up to the sign) eigenfunctions we assume

$$\lambda_1 > \lambda_2 > \dots > \lambda_d > \lambda_{d+1} > 0. \quad (2.17)$$

Define

$$T_N^0(d) = \sum_{i=1}^d \frac{1}{\hat{\lambda}_i} \int \langle Z_N(x, \cdot), \hat{\varphi}_i \rangle^2 dx,$$

$$T_N^*(d) = \sum_{i=1}^d \int \langle Z_N(x, \cdot), \hat{\varphi}_i \rangle^2 dx,$$

$$M_N^0(d) = \sum_{i=1}^d \frac{1}{\hat{\lambda}_i} \int \left(\langle Z_N(x, \cdot), \hat{\varphi}_i \rangle - \int \langle Z_N(u, \cdot), \hat{\varphi}_i \rangle du \right)^2 dx$$

and

$$M_N^*(d) = \sum_{i=1}^d \int \left(\langle Z_N(x, \cdot), \hat{\varphi}_i \rangle - \int \langle Z_N(u, \cdot), \hat{\varphi}_i \rangle du \right)^2 dx.$$

Theorem 2.2. *If assumptions (2.1)–(2.4), (2.13)–(2.16), (2.17) and H_0 hold, then*

$$T_N^0(d) \xrightarrow{\mathcal{D}} \sum_{i=1}^d \int B_i^2(x) dx, \quad (2.18)$$

$$T_N^*(d) \xrightarrow{\mathcal{D}} \sum_{i=1}^d \lambda_i \int B_i^2(x) dx, \quad (2.19)$$

$$M_N^0(d) \xrightarrow{\mathcal{D}} \sum_{i=1}^d \int \left(B_i(x) - \int B_i(u) du \right)^2 dx \quad (2.20)$$

and

$$M_N^*(d) \xrightarrow{\mathcal{D}} \sum_{i=1}^d \lambda_i \int \left(B_i(x) - \int B_i(u) du \right)^2 dx. \quad (2.21)$$

It is clear that T_N^* and M_N^* are just d -dimensional projections of T_N and M_N . The distribution of the limit in (2.18) can be found in Kiefer (1959). Critical values based on Monte Carlo simulations are given in Table 6.1 of Horváth and Kokoszka (2012). The distributions of the limits both in (2.18) and (2.20) can also be expressed in terms of sums of squared normals, see Shorack and Wellner (1986) and Section 4. It is also easy to derive normal approximations. By the central limit theorem we have, as $d \rightarrow \infty$,

$$\left(\frac{45}{d} \right)^{-1/2} \left[\sum_{i=1}^d \int B_i^2(x) dx - \frac{d}{6} \right] \xrightarrow{\mathcal{D}} N(0, 1),$$

where $N(0, 1)$ stands for a standard normal random variable. Aue et al. (2009) demonstrated that the limit in (2.18) can be approximated well with normal random variables even for moderate d . The limit in (2.20) can be approximated in a similar manner, namely, as $d \rightarrow \infty$,

$$\left(\frac{360}{d} \right)^{-1/2} \left[\sum_{i=1}^d \left\{ \int B_i^2(x) dx - \left(\int B_i(x) dx \right)^2 \right\} - \frac{d}{12} \right] \xrightarrow{\mathcal{D}} N(0, 1).$$

3. Asymptotic behavior under alternatives

The asymptotic behavior of the KPSS and related tests under alternatives is not completely understood, even for scalar data. This may be due to the fact that an asymptotic analysis of power is generally much more difficult than the theory under a null hypothesis. Giraitis et al. (2003) studied the behavior of the KPSS test, the R/S test of Lo (1991) and their V/S test under the alternative of long memory. Pelagatti and Sen (2013) established the consistency of their nonparametric version of the KPSS test under the integrated alternative. In this section, we present an asymptotic analysis, under alternatives, of the tests introduced in Section 2. In the functional setting, there is a fundamentally new aspect: convergence of a scalar estimator of the long run variance must be replaced by the convergence of the eigenvalues and the eigenfunctions of the long run covariance function. We derive precise rates of convergence and limits for this function, and use them to study the asymptotic power of the tests introduced in Section 2. In Section 4, we will see how these asymptotic insights manifest themselves in finite samples.

We expect that the tests introduced in Section 2 are also consistent against suitably defined long memory alternatives. While scalar long memory models have received a lot of attention in recent decades, long memory functional models have not been considered in econometric literature yet. To keep this contribution within reasonable limits, we do not pursue this direction here.

3.1. Change in the mean alternative

To state consistency results, we assume that the jump function is in L^2 , i.e.

$$\int \delta^2(t) dt < \infty. \quad (3.1)$$

We introduce the function

$$\delta_\tau(x, t) = \delta(t)[(x - \tau)I\{x \geq \tau\} - x(1 - \tau)] \quad (3.2)$$

and the Gaussian process $\Gamma^0(x, t)$ with $E\Gamma^0(x, t) = 0$ and $E\Gamma^0(x, t)\Gamma^0(y, s) = (\min(x, y) - xy)C(t, s)$.

The existence of the process $\Gamma^0(x, t)$ will be established in Appendix A.

Theorem 3.1. *If assumptions (2.1)–(2.4), (2.5), (3.1) and $H_{A,1}$ hold, then*

$$N^{-1/2} \left\{ T_N - \frac{N}{3} \tau^2 (1 - \tau)^2 \|\delta\|^2 \right\} \xrightarrow{D} 2 \iint \Gamma^0(x, t) \delta_\tau(x, t) dt dx \tag{3.3}$$

and

$$N^{-1/2} \left\{ M_N - \frac{N}{12} \tau^2 (1 - \tau)^2 \|\delta\|^2 \right\} \xrightarrow{D} 2 \iint \left(\Gamma^0(x, t) - \int \Gamma^0(y, t) dy \right) \times \left(\delta_\tau(x, t) - \int \delta_\tau(y, t) dy \right) dt dx. \tag{3.4}$$

It is easy to see that the limits in Theorem 3.1 are zero mean normal random variables. Their variances, computed in Appendix B, are positive if $C(t, s)$ is strictly positive definite. In that case, T_N and M_N increase like N . However, as we prove in Lemma B.2, $\hat{C}_N(t, s)$ does not converge to $C(t, s)$ under $H_{A,1}$, so it is not clear what the asymptotic behavior of the critical values under $H_{A,1}$ is. To show that the asymptotic power is 1, a more delicate argument is needed, which we now outline.

Applying Lemma B.2 with the result of Dunford and Schwartz (1988), p. 1091, we conclude that

$$\frac{\hat{\lambda}_1}{h} \xrightarrow{P} \gamma_{A,1} = 2\tau(1 - \tau)\|\delta\|^2 \int_0^c K(u)du, \tag{3.5}$$

and

$$\left\| \hat{\varphi}_1(t) - \hat{c}_1 \frac{\delta(t)}{\|\delta\|} \right\| = o_P(1). \tag{3.6}$$

According to (3.5), when we compute $\bar{c} = \bar{c}(h, N)$, the critical value from simulated copies of $\sum_{i=1}^d \hat{\lambda}_i \int B_i^2(t)dt$, then \bar{c} increases at most linearly with h . Therefore, using (2.16) with Theorem 3.1, we conclude that

$$\lim_{N \rightarrow \infty} P\{T_N \geq \bar{c}\} = 1 \quad \text{under } H_{A,1}. \tag{3.7}$$

This shows that the test based on T_N is consistent. The same argument applies to M_N .

We now turn to the tests based on projections, with the test statistics defined in Section 2.2. As we have seen, under $H_{A,1}$, the largest empirical eigenvalue $\hat{\lambda}_1$ increases to ∞ , as $N \rightarrow \infty$, and the corresponding empirical eigenfunction $\hat{\varphi}_1$ is asymptotically in the direction of the change. This means that both T_N^* and M_N^* are dominated by the first term under $H_{A,1}$. The precise asymptotic behavior of all statistics introduced in Section 2.2 is described in the following theorem.

Theorem 3.2. *If assumptions (2.1)–(2.4), (2.13)–(2.16), (3.1) and $H_{A,1}$ hold, then*

$$N^{-1/2} \left\{ T_N^*(1) - \frac{N}{3} \tau^2 (1 - \tau)^2 \langle \delta, \hat{\varphi}_1 \rangle^2 \right\} \xrightarrow{D} 2 \iint \Gamma^0(x, t) \delta_\tau(x, t) dx dt, \tag{3.8}$$

$$N^{-1/2} \left\{ M_N^*(1) - \frac{N}{12} \tau^2 (1 - \tau)^2 \langle \delta, \hat{\varphi}_1 \rangle^2 \right\} \xrightarrow{D} 2 \iint \left(\Gamma^0(x, t) - \int \Gamma^0(y, t) dy \right) \times \left(\delta_\tau(x, t) - \int \delta_\tau(y, t) dy \right) dt dx, \tag{3.9}$$

$$\frac{h}{N^{1/2}} 2\tau(1 - \tau)\|\delta\|^2 \int_0^c K(u)du \left\{ T_N^0(1) - \frac{N}{3\hat{\lambda}_1} \tau^2 (1 - \tau)^2 \langle \delta, \hat{\varphi}_1 \rangle^2 \right\} \xrightarrow{D} 2 \iint \Gamma^0(x, t) \delta_\tau(x, t) dx dt, \tag{3.10}$$

and

$$\frac{h}{N^{1/2}} 2\tau(1 - \tau)\|\delta\|^2 \int_0^c K(u)du \left\{ M_N^0(1) - \frac{N}{12\hat{\lambda}_1} \tau^2 (1 - \tau)^2 \langle \delta, \hat{\varphi}_1 \rangle^2 \right\} \xrightarrow{D} 2 \iint \left(\Gamma^0(x, t) - \int \Gamma^0(y, t) dy \right) \times \left(\delta_\tau(x, t) - \int \delta_\tau(y, t) dy \right) dt dx. \tag{3.11}$$

If in addition we assume that $h/N^{1/2} \rightarrow 0$ as $N \rightarrow \infty$, then

$$T_N^*(d) = \frac{N}{3} \tau^2 (1 - \tau)^2 \|\delta\|^2 (1 + o_P(1)), \tag{3.12}$$

$$M_N^*(d) = \frac{N}{12} \tau^2 (1 - \tau)^2 \|\delta\|^2 (1 + o_P(1)), \tag{3.13}$$

$$T_N^0(d) = \frac{N}{h} \frac{\tau(1 - \tau)}{6 \int_0^c K(u)du} (1 + o_P(1)), \tag{3.14}$$

and

$$M_N^0(d) = \frac{N}{h} \frac{\tau(1 - \tau)}{24 \int_0^c K(u)du} (1 + o_P(1)). \tag{3.15}$$

Observe that according to Theorems 3.1 and 3.2, the statistics T_N and $T_N^*(1)$ (M_N and $M_N^*(1)$, respectively) exhibit the same asymptotic behavior under the change point alternative. This is due to the fact that the projection in the direction of $\hat{\varphi}_1$ picks up all information on the change available in the data, as, by (3.6), $\hat{\varphi}_1$ is asymptotically aligned with the direction of the change.

Remark 3.1. In the local alternative model

$$X_i(t) = \mu(t) + \delta_N^*(t)I\{i > k^*\} + \eta_i(t), \quad 1 \leq i \leq N,$$

with some integer $1 \leq k^* < N$,

where $\|\delta_N^*\| \rightarrow 0$ as $N \rightarrow \infty$. We discuss briefly how the statistic T_N behaves under this model. If $N^{1/2}\|\delta_N^*\| \rightarrow 0$, then T_N converges in distribution to $\iint (\Gamma^0(x, t))^2 dt dx$ as is the case under H_0 . On the other hand, if $N^{1/2}\|\delta_N^*\| \rightarrow \infty$, then $T_N \xrightarrow{P} \infty$ and therefore consistency is retained. Moreover, under the additional assumption $N \iint C(t, s) \delta_N^*(t) \delta_N^*(s) dt ds \rightarrow \infty$ we show that

$$\frac{1}{A_N} \left\{ T_N - \frac{1}{N} \|\beta_N\|^2 \right\} \xrightarrow{D} N(0, 1), \tag{3.16}$$

where

$$A_N^2 = 4N \iint C(t, s) \delta_N^*(t) \delta_N^*(s) dt ds \times \iint (\min(x, y) - xy) \bar{\delta}_\tau(x) \bar{\delta}_\tau(y) dx dy.$$

In the critical case when $N^{1/2} \|\delta_N^*\| \rightarrow \delta^*$ in L^2 , where δ^* is some non zero function, then we have

$$T_N \xrightarrow{\mathcal{D}} \zeta + \sum_{\ell=1}^{\infty} \left\{ \lambda_\ell \|B_\ell\|^2 + 2\lambda_\ell^{1/2} \langle B_\ell, \bar{\delta}_\tau \rangle \langle \varphi_\ell, \delta^* \rangle \right\}, \quad (3.17)$$

where $\zeta = \|\bar{\delta}_\tau\|^2 \|\delta^*\|^2$, B_1, B_2, \dots are independent Brownian bridges, the λ_i 's and φ_i 's are defined in (2.9), and

$$\bar{\delta}_\tau(x) = (x - \tau)I\{x \geq \tau\} - x(1 - \tau). \quad (3.18)$$

The asymptotic behavior of M_N can be studied analogously in the local alternative change point model. The derivation of the asymptotic properties of $T_N^0(d)$, $T_N^*(d)$, $M_N^0(d)$, and $M_N^*(d)$ is much more involved since it requires the study of $\hat{C}_N(t, s)$ under this model. We will not pursue this line of inquiry in the present paper.

3.2. The integrated alternative

Let

$$\Delta(x, t) = \int_0^x \Gamma(u, t) du - x \int \Gamma(u, t) du, \quad (3.19)$$

where $\Gamma(x, t)$ is a Gaussian process with $E\Gamma(x, t) = 0$ and $E\Gamma(x, t)\Gamma(y, s) = \min(x, y)C(t, s)$. The existence of $\Gamma(x, t)$ is established in Theorem A.1.

For the fully functional tests of Section 2.1, we have the following result.

Theorem 3.3. *If assumptions (2.1)–(2.4) and $H_{A,2}$ hold, then*

$$\frac{1}{N^2} T_N \xrightarrow{\mathcal{D}} \iint \Delta^2(x, t) dt dx \quad (3.20)$$

and

$$\frac{1}{N^2} M_N \xrightarrow{\mathcal{D}} \iint \left(\Delta(x, t) - \int \Delta(u, t) du \right)^2 dt dx. \quad (3.21)$$

To find the limit distributions of the statistics based on projections, we need the following theorem.

Theorem 3.4. *If assumptions (2.1)–(2.4), (2.13)–(2.16) and $H_{A,2}$ hold, then*

$$\left\{ \frac{1}{N} Z_N(x, t), \frac{1}{Nh} \hat{C}_N(t, s), 0 \leq x, t, s \leq 1 \right\} \rightarrow \{ \Delta(x, t), Q(t, s), 0 \leq x, t, s \leq 1 \}$$

in $\mathcal{D}([0, 1] \times L^2)$, where

$$Q(t, s) = 2 \left(\int_0^c K(w) dw \right) \int R(z, t) R(z, s) dz,$$

with

$$R(z, t) = \int_0^z \Gamma(u, t) du - \int \left\{ \int_0^v \Gamma(u, t) du \right\} dv.$$

We show in Lemma B.5 that $Q(t, s)$ is non-negative definite with probability one, so there are random variables $\lambda_1^* \geq \lambda_2^* \geq \dots$ and random functions $\varphi_1^*(t), \varphi_2^*(t), \dots$ satisfying

$$\lambda_i^* \varphi_i^*(t) = \int Q(t, s) \varphi_i^*(s) ds, \quad 1 \leq i < \infty. \quad (3.22)$$

Combining Theorem 3.4 with Dunford and Schwartz (1988), we get that

$$\left(\hat{\lambda}_1/(Nh), \hat{\lambda}_2/(Nh), \dots, \dots, \hat{\lambda}_d/(Nh), \hat{\varphi}_1(t), \hat{\varphi}_2(t), \dots, \hat{\varphi}_d(t) \right) \xrightarrow{\mathcal{D}} \left(\lambda_1^*, \lambda_2^*, \dots, \lambda_d^*, \varphi_1^*(t), \varphi_2^*(t), \dots, \varphi_d^*(t) \right).$$

Thus the behavior of $T_N^0(d)$, $T_N^*(d)$, $M_N^0(d)$ and $M_N^*(d)$ is an immediate consequence of Theorem 3.4. An argument similar to that developed in Section 3.1 shows that the tests are consistent.

Theorem 3.5. *If assumptions (2.1)–(2.4), (2.13)–(2.16) and $H_{A,2}$ hold, then*

$$\frac{h}{N} T_N^0(d) \xrightarrow{\mathcal{D}} \sum_{i=1}^d \frac{1}{\lambda_i^*} \int \langle \Delta(x, \cdot), \varphi_i^*(\cdot) \rangle^2 dx, \quad (3.23)$$

$$\frac{1}{N^2} T_N^*(d) \xrightarrow{\mathcal{D}} \sum_{i=1}^d \int \langle \Delta(x, \cdot), \varphi_i^*(\cdot) \rangle^2 dx, \quad (3.24)$$

$$\frac{h}{N} M_N^0(d) \xrightarrow{\mathcal{D}} \sum_{i=1}^d \frac{1}{\lambda_i^*} \int \left(\langle \Delta(x, \cdot), \varphi_i^*(\cdot) \rangle - \int \langle \Delta(u, \cdot), \varphi_i^*(\cdot) \rangle du \right)^2 dx, \quad (3.25)$$

and

$$\frac{1}{N^2} M_N^*(d) \xrightarrow{\mathcal{D}} \sum_{i=1}^d \int \left(\langle \Delta(x, \cdot), \varphi_i^*(\cdot) \rangle - \int \langle \Delta(u, \cdot), \varphi_i^*(\cdot) \rangle du \right)^2 dx. \quad (3.26)$$

3.3. Deterministic trend alternative

Let

$$\bar{g}(x) = \int_0^x g(u) du - x \int g(u) du, \quad 0 \leq x \leq 1.$$

Theorem 3.6. *If assumptions (2.1)–(2.4), (2.6), (3.1) and $H_{A,3}$ hold, then*

$$N^{-1/2} \left\{ T_N - N \|\delta\|^2 \int \bar{g}^2(x) dx \right\} \xrightarrow{\mathcal{D}} 2 \iint \Gamma^0(x, t) \delta(t) \bar{g}(x) dt dx \quad (3.27)$$

and

$$N^{-1/2} \left\{ M_N - N \|\delta\|^2 \int \left(\bar{g}(x) - \int \bar{g}(y) dy \right)^2 dx \right\} \xrightarrow{\mathcal{D}} 2 \iint \left(\Gamma^0(x, t) - \int \Gamma^0(y, t) dy \right) \times \left(\bar{g}(x) - \int \bar{g}(y) dy \right) \delta(t) dt dx. \quad (3.28)$$

The limits in (3.27) and (3.28) are normal random variables with zero mean and variances which can be expressed in terms of the long run covariance kernel $C(\cdot, \cdot)$ and the functions δ and \bar{g} . We do not display these complex formulas to conserve space. They extend the formulas for the variances of the limits in Theorem 3.1 which are given in Appendix B. The consistency of the procedures based on projections can be established by extending the arguments used to prove Theorem 3.2, however with more abstract notation. Again, to keep this work within reasonable limits of space, we do not present the details.

4. Implementation and finite sample performance of the tests

4.1. Details of the implementation

To implement the tests introduced in Section 2, several issues must be considered. The choice of the kernel $K(\cdot)$ and the smoothing bandwidth h are the most obvious. Beyond that, to implement Monte Carlo tests based on statistics whose limits depend on the estimated eigenvalues, a fast method of calculating replications of these limits must be employed. The issues of bandwidth and kernel selection have been extensively studied in the econometric literature for over three decades, we cannot cite dozens, if not hundreds, of papers devoted to them. Perhaps the best known contributions are those of Andrews (1991) and Andrews and Monahan (1992) who introduced data driven bandwidth selection and prewhitening. While these approaches possess optimality properties in general regression models with heteroskedastic and correlated errors, they are not optimal in all specific applications. In particular, Jönsson (2006) found that the finite-sample distribution of the (scalar) KPSS test statistic can be very unstable when the Quadratic Spectral kernel (recommended by Andrews (1991)) is used and/or a prewhitening filter is applied. He recommends the Bartlett kernel. An elaboration on the finite sample properties of the KPSS test with many relevant references can be found in Jönsson (2011). Our paper focuses on the derivation and large sample theory for the stationarity tests for functional time series; we cannot present here a comprehensive and conclusive study of the finite sample properties, which are still being investigated even for scalar time series. We however wish to offer some practical guidance and report approaches which worked well for the data generating processes we considered.

Politis (2003, 2011) argues that the flat top kernel

$$K(t) = \begin{cases} 1, & 0 \leq t < 0.1 \\ 1.1 - |t|, & 0.1 \leq t < 1.1 \\ 0, & |t| \geq 1.1 \end{cases} \quad (4.1)$$

has better properties than the Bartlett or the Parzen kernels. In our empirical work, we used kernel (4.1). Our simulations showed that $h = N^{1/2}$ is satisfactory for our hypothesis testing problem when the observations are independent or weakly dependent (functional autoregressive processes). The empirical sizes and power functions change little if h is taken ± 5 lags smaller or larger. We note that the optimal rates derived in Andrews (1991) do not apply to kernel (4.1) because this piecewise function does not satisfy the regularity conditions assumed by Andrews (1991). It can be shown that the optimal rates for Bartlett and Parzen kernels remain the same in the functional case, but the multiplicative constants depend in a very complex way on the high order moments of the functions, and the arguments Andrews (1991) used to approximate them cannot be readily extended.

Once the kernel and the bandwidth have been selected, the eigenvalues $\hat{\lambda}_i$ can be computed. This allows us to compute the normalized statistics $T_N^0(d)$ and $M_N^0(d)$ and use the tests based on the asymptotic distribution of their limits. The critical values can

be computed by using the expansions analogous to (4.2) or (4.3) (without the $\hat{\lambda}_i$). Alternatively, since these limits do not depend on the distribution of the data, the critical values can be obtained by calculating a large number of replications of $T_N^0(d)$ and $M_N^0(d)$ for any specific functional time series. We used iid Brownian motions, and we refer to the tests which use the critical values so obtained as $T_N^0(d)(AM)$ and $M_N^0(d)(AM)$ (Alternative method). This method is extremely computationally intensive, if its performance is to be assessed by simulations; we needed almost two months of run time on the University of Utah Supercomputer (as of June 2013) to obtain the empirical rejection rates for $T_N^0(d)(AM)$ and $M_N^0(d)(AM)$ for samples of size 100 and 250 and values of d between 1 and 10.

The limits of statistics T_N and T_N^* must be approximated by the MC distribution of $\sum_{i=1}^d \hat{\lambda}_i \int B_i^2(x) dx$, and one must proceed analogously for M_N and M_N^* . Using the expansions discussed in Shorack and Wellner (1986), pp 210–211, we use the approximations

$$\hat{T}_{d,J} = \sum_{i=1}^d \hat{\lambda}_i \sum_{j=1}^J \frac{Z_{ij}^2}{j^2 \pi^2}, \quad (4.2)$$

and

$$\hat{M}_{d,J} = \sum_{i=1}^d \hat{\lambda}_i \sum_{j=1}^J \frac{Z_{2j-1}^2 + Z_{2j}^2}{4j^2 \pi^2}, \quad (4.3)$$

where $\{Z_{ij}\}_{j=1}^\infty$ are iid standard normal random variables. For large J , the sums over j approximate the integrals of the functionals of the Brownian bridge and eliminate the need to generate its trajectories and to perform numerical integration. In our work we used $J = 100$, and one thousand replications to obtain MC distributions.

To select d , we use the usual “cumulative variance” approach recommended by Ramsay and Silverman (2005) and Horváth and Kokoszka (2012); d is chosen so that roughly $v\%$ of the sample variance is explained by the first d principal components. In our implementation, we estimated the total of 49 largest eigenvalues (the largest number under which the estimation is numerically stable), and used $d = d_v$ such that

$$\frac{\hat{\lambda}_1 + \dots + \hat{\lambda}_{d_v}}{\hat{\lambda}_1 + \dots + \hat{\lambda}_{49}} \approx v.$$

A general recommendation is to use v equal to about 90 percent, but we report results for $v = .85, .90, .95$, to see how the performance of the tests is affected by the choice of d . This is a new aspect of the stationarity tests, which reflects the infinite dimensional structure of the functional data, and which is absent in tests for scalar or vector time series.

4.2. Empirical size and power

We first compare the empirical size of the tests implemented as described above. We consider two data generating processes (DGP’s): (1) iid copies of the Brownian motion (BM), (2) the functional AR process of order 1 (FAR(1)). There are a large number of stationary functional time series that could be considered. In our small simulation study, the focus on the BM is motivated by the application to cumulative intraday returns considered in Section 5; they approximately look like realizations of the BM, see Fig. 3. The FAR(1), with Brownian motion innovations, is used to generate temporal dependence: the tests should have correct size for general stationary functional time series, not just for iid functions. The FAR(1) process is defined by the equation

$$X_i(t) = \int_0^1 \psi(t, u) X_{i-1}(u) du + W_i(t), \quad 0 \leq t \leq 1, \quad (4.4)$$

Table 1

Empirical sizes for the iid BM and FAR(1) DGPs. We used $h = N^{1/2}$ and the flat-top kernel (4.1). The standard error is approximately 0.9% for the 10% level and 0.4% for the 5% level.

DGP	BM				FAR(1)				
	100		250		100		250		
Nominal	10%	5%	10%	5%	10%	5%	10%	5%	
Statistics	T_N								
v	.85	13.9	5.6	13.3	5.3	12.3	4.2	11.8	4.1
	.90	12.6	5.4	12.5	5.2	11.8	3.7	11.0	4.1
	.95	12.7	4.6	12.0	4.9	11.3	3.6	10.4	3.7
Statistics	M_N								
v	.85	9.7	2.2	11.4	3.6	11.4	2.3	11.8	4.2
	.90	8.8	1.5	10.5	3.0	9.4	1.6	11.2	4.0
	.95	8.2	0.9	9.9	2.9	8.5	1.2	10.1	3.4
Statistics	T_N^*								
v	.85	11.3	4.9	10.8	4.6	10.2	3.4	10.3	3.7
	.90	11.2	4.4	10.7	4.8	10.0	3.4	10.3	3.4
	.95	11.8	4.4	11.1	4.5	10.6	3.2	10.0	3.6
Statistics	M_N^*								
v	.85	6.3	0.7	8.8	2.6	11.0	2.1	11.2	4.2
	.90	7.0	0.8	8.9	2.3	8.9	1.3	10.7	4.0
	.95	6.9	0.7	8.8	2.7	8.2	1.1	10.0	3.2
Statistics	T_N^0								
v	.85	10.4	3.8	10.3	3.9	10.1	2.8	9.2	3.3
	.90	9.2	2.3	9.0	2.8	7.7	1.5	8.6	2.9
	.95	4.6	0.8	7.6	1.4	5.0	0.1	7.2	1.3
Statistics	M_N^0								
v	.85	6.1	0.5	6.7	2.1	6.7	1.5	7.7	2.8
	.90	4.2	0.8	5.4	1.7	5.8	1.0	7.2	2.4
	.95	2.9	0.3	5.6	1.4	3.3	0.0	5.1	0.9
Statistics	$T_N^0(AM)$								
v	.85	11.9	5.4	10.2	5.1	12.1	7.1	11.7	6.1
	.90	10.3	5.7	9.2	4.8	11.7	7.2	9.8	4.9
	.95	9.9	4.3	9.0	4.7	11.2	6.9	9.7	5.3
Statistics	$M_N^0(AM)$								
v	.85	8.8	5.1	10.7	4.6	12.7	7.7	10.8	5.8
	.90	8.6	5.3	10.0	4.5	12.1	7.3	10.5	5.4
	.95	8.5	4.7	9.8	5.2	11.9	7.1	10.6	5.4

where the W_i are independent Brownian motions on $[0, 1]$, and ψ is a kernel whose operator norm is not too large, the precise condition is somewhat technical, see Bosq (2000) or Chapter 13 of Horváth and Kokoszka (2012). A sufficient condition for a stationary solution to Eq. (4.4) to exist is that the Hilbert-Schmidt norm of ψ be less than 1. We work with the kernel

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

with $c = .3416$ so that the Hilbert-Schmidt norm of ψ is approximately 0.5.

We consider functional time series of length $N = 100$ and $N = 250$. Each DGP is simulated one thousand times, and the percentage of rejections of the null hypothesis is reported at the significance levels of 10% and 5%. The empirical sizes are reported in Table 1, which leads to the following conclusions:

- (1) The tests $T_N^0(AM)$ and $T_N^0(AM)$ have reasonably good empirical size, which does not depend on v . Note that we used the BM processes to obtain the critical values, so it is not surprising good results for the BM as the DGP. However the observations of the FAR(1) series are no longer BM's.
- (2) If the limit distribution is used to calculate the critical values, the tests based on the MC distributions (statistics T_N, M_N, T_N^*, M_N^*) are less sensitive to the choice of the cumulative variance v .

Table 2

Empirical power for change point and $I(1)$ alternatives. We used $h = N^{1/2}$ and the flat-top kernel (4.1).

DGP	Change point				I(1)				
	100		250		100		250		
Nominal	10%	5%	10%	5%	10%	5%	10%	5%	
Statistic	T_N								
v	.85	80.7	56.4	99.6	98.1	99.3	96.5	99.2	96.3
	.90	80.1	56.6	99.5	97.6	99.4	95.8	99.2	96.1
	.95	79.2	54.4	99.4	97.6	99.1	96.2	99.2	96.3
Statistic	M_N								
v	.85	50.6	14.7	95.2	84.1	93.8	68.3	97.7	92.5
	.90	46.7	11.0	94.7	82.7	92.8	64.5	97.6	92.4
	.95	79.2	54.4	99.4	97.6	90.9	61.4	99.2	96.3
Statistic	T_N^*								
v	.85	77.2	52.1	99.3	97.6	98.9	95.7	98.0	94.2
	.90	77.8	54.3	99.5	97.5	99.2	95.8	98.4	95.7
	.95	77.5	53.7	99.4	97.6	99.1	96.0	99.1	96.1
Statistic	M_N^*								
v	.85	39.6	8.5	93.7	78.1	93.5	67.9	94.9	88.2
	.90	39.9	7.7	93.9	79.6	92.7	63.9	96.2	89.3
	.95	40.8	6.8	94.5	79.8	90.5	61.2	96.6	90.0
Statistic	T_N^0								
v	.85	85.8	55.1	99.8	98.9	99.5	98.1	98.6	96.2
	.90	86.6	52.0	100	99.6	99.7	98.8	99.3	98.7
	.95	74.7	31.3	99.9	98.4	100	96.0	99.9	99.8
Statistic	M_N^0								
v	.85	35.0	7.8	97.2	77.7	86.1	75.2	97.9	92.7
	.90	31.0	5.9	98.0	71.5	90.9	73.4	99.2	95.4
	.95	21.1	4.8	93.0	63.0	96.8	75.9	100	98.5
Statistic	$T_N^0(AM)$								
v	.85	96.6	91.6	100	100	99.6	99.3	99.8	99.7
	.90	94.9	85.5	100	100	100	99.9	100	100
	.95	82.5	70.0	100	100	100	100	100	100
Statistic	$M_N^0(AM)$								
v	.85	85.3	71.3	99.9	99.8	93.4	85.0	99.8	99.3
	.90	68.7	52.3	99.7	98.8	96.3	90.3	99.9	99.9
	.95	43.7	28.5	97.0	92.0	98.6	96.3	100	100

- (3) The tests based M_N and M_N^* are generally too conservative at the 5% level.
- (4) Even though statistic T_N^* is too conservative at the 5% level in case of the FAR(1) model, it achieves a reasonable balance of empirical size at the 10% and 5% levels.
- (5) If the temporal dependence is not too strong, we recommend statistics T_N^* with $v = 90\%$.

We now turn to the investigation of the empirical power. The number of DGP's that could be considered under the alternative of nonstationarity is enormous. In our simulation study, we consider merely two examples intended to illustrate the theory developed in Section 3. Under the change point alternative, $H_{A,1}$, the DGP is

$$X_i(t) = \begin{cases} B_i(t) & \text{if } i < [N/2] \\ B_i(t) + \delta(t) & \text{if } i \geq [N/2], \end{cases}$$

where the B_i are iid Brownian bridges, and $\delta(t) = 2t(1-t)$, so that the change in the mean function is comparable to the typical size of the Brownian bridge. Under the $I(1)$ alternative, $H_{A,2}$, we consider the integrated functional sequence defined by

$$X_i(t) = X_{i-1}(t) + B_i(t), \quad 1 \leq i \leq N,$$

where $X_0(t) = B_0(t)$, and $\{B_i(t)\}_{i=0}^\infty$ are iid Brownian Bridges. Again, each data generating process is simulated 1000 times and the rejection rate of H_0 is reported when the significance level is 10% and 5%. Table 2 shows the results of these simulations. The following conclusions can be reached:

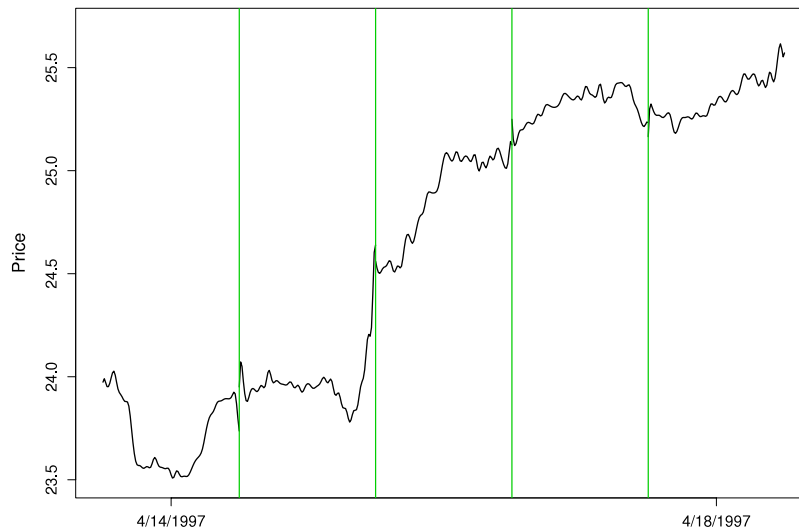


Fig. 1. Five functional data objects constructed from the 1-min average price of Disney stock. The vertical lines separate the days.

- (1) Under the change point alternative, the T statistics have higher power than the M statistics. This is in perfect agreement with Theorems 3.1 and 3.2, which show that the leading terms of the T statistics are four times larger than those of the corresponding M statistics.
- (2) The same observation remains true under the integrated alternative, and again it agrees with the theoretical rates obtained in Theorems 3.3 and 3.4. The multiplicative constants of leading terms of the T statistics are equal to second moments and those of the M statistics to corresponding variances.
- (3) As for empirical size, the T statistics are not sensitive to the choice of v .
- (4) The test based on T_N^* has slightly lower power than those based on T_N^0 and T_N , but this is because the latter two tests have slightly inflated sizes. Our overall recommendation remains to use T_N^* with $v = 0.90$. However, if very high power is of central importance, and computational time not a big concern, the method $T_N^0(AM)$ might be superior.

5. Application to intraday price curves

One of the most natural and obvious functional data are intraday price curves; five such functions are shown in Fig. 1. Not much quantitative research has however focused on the analysis of the information contained in the *shapes* of such curves, even though they very closely reflect the reactions and expectations of intraday investors. Extensive research has focused on scalar or vector summary statistics derived from intraday data, including realized volatility and noise variance estimation, see Barndorff-Nielsen and Shephard (2004) and Wang and Zou (2010), among many others. Several papers have however considered the shapes of suitably defined price or volatility curves, see Gabryś et al. (2010), Müller et al. (2011), Gabryś et al. (forthcoming), Kokoszka and Reimherr (2013) and Kokoszka et al. (2013). This paper focuses on statistical methodology and the underlying theory, and we cannot include a comprehensive empirical study of functional aspects of intraday price data. We merely show that the application of our tests leads to meaningful and useful insights.

Suppose $P_n(t_j)$, $n = 1, \dots, N$, $j = 1, \dots, m$, is the price of a financial asset at time t_j on day n . Fig. 1 shows five functional data objects constructed from the one minute average price of Disney stock interpolated by B-splines. In this case, the number of points t_j used to construct each object is $m = 390$. Each object is viewed

as a continuous curve making this data an excellent candidate for functional data analysis. As daily closing prices form a nonstationary scalar time series, we would expect the daily price curves to form a nonstationary functional time series. When our tests are applied to sufficiently long periods of time, they indeed always reject the null hypothesis of stationarity. For shorter periods of time, H_0 is sometimes rejected and sometimes is not, most likely due to reduced power. To illustrate, Fig. 2 displays the P-values for the test based on T_N applied to consecutive nonoverlapping segments of length N in the time period from 04/09/1997 to 04/02/2007, which comprises 2,510 trading days. This means that there are 50 segments of length $N = 50$, 25 segments of length $N = 100$ and 10 segments of length $N = 250$. If $N = 250$, H_0 is always rejected. We obtained very similar results for the other T statistics. When the M statistics are used, the rejection rates are marginally lower, but overall commensurate with those for the T statistics. We also applied the tests to several other stocks over the same period, including Chevron, Bank of America, Microsoft, IBM, McDonalds, and Walmart, and obtained nearly identical results. The results are also very similar for gold futures. The price of gold increased five fold between 2001 and 2011, with an almost linear trend. For segments of length $N = 100$, the null is sometimes not rejected if the curves do not show a clear increasing tendency over that period, but otherwise we obtained strong rejections.

In order to fit stationary functional time series models to intraday price curves, a suitable transformation should be applied. Gabryś et al. (2010) put forward the following definition.

Definition 1. Suppose $P_n(t_j)$, $n = 1, \dots, N$, $j = 1, \dots, m$, is the price of a financial asset at time t_j on day n . The functions

$$R_n(t_j) = 100[\ln P_n(t_j) - \ln P_n(t_1)], \quad j = 1, 2, \dots, m, \\ n = 1, \dots, N,$$

are called the *cumulative intraday returns* (CIDR's).

The idea behind Definition 1 is very simple. If the return from the start of a trading day until its close remains within the 5% range, $R_n(t_j)$ is practically equal to the simple return $[P_n(t_j) - P_n(t_1)]/P_n(t_1)$. Since $P_n(t_1)$ is fixed for every trading day, the $R_n(t_j)$ have practically the same *shape* as the price curves, see Fig. 3. However, since they always start from zero, level stationarity is enforced. The division by $P_n(t_1)$ helps reduce the scale inflation. It can thus be hoped that the CIDR's will form a stationary functional time series, which will be amenable to the statistical analysis of the

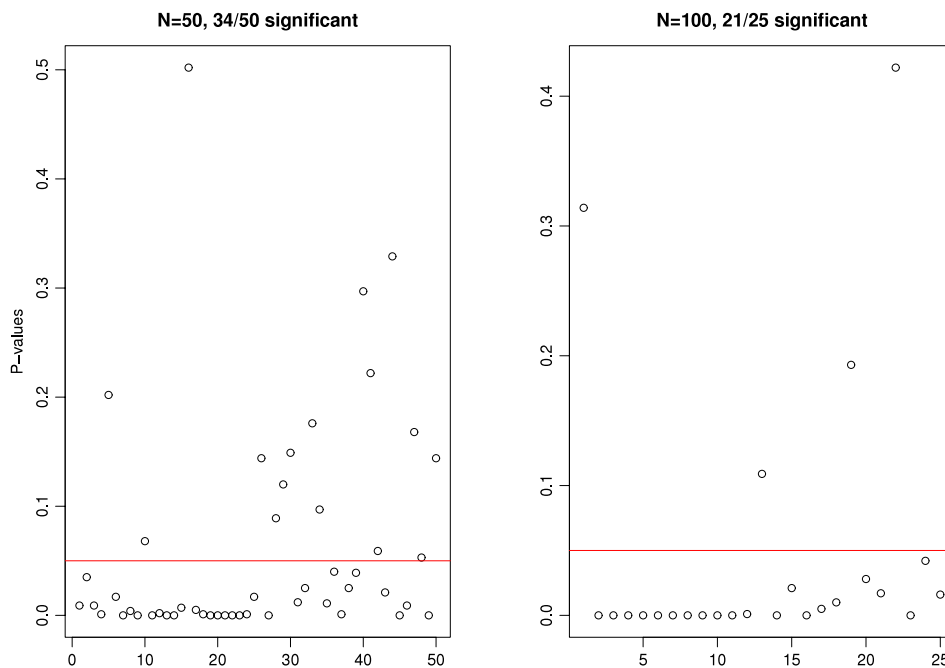


Fig. 2. P-values for consecutive segments of length N of the price curves $P_n(t)$ of the Disney stock computed using T_N with $\nu = .9$. The horizontal line shows the 5% threshold.

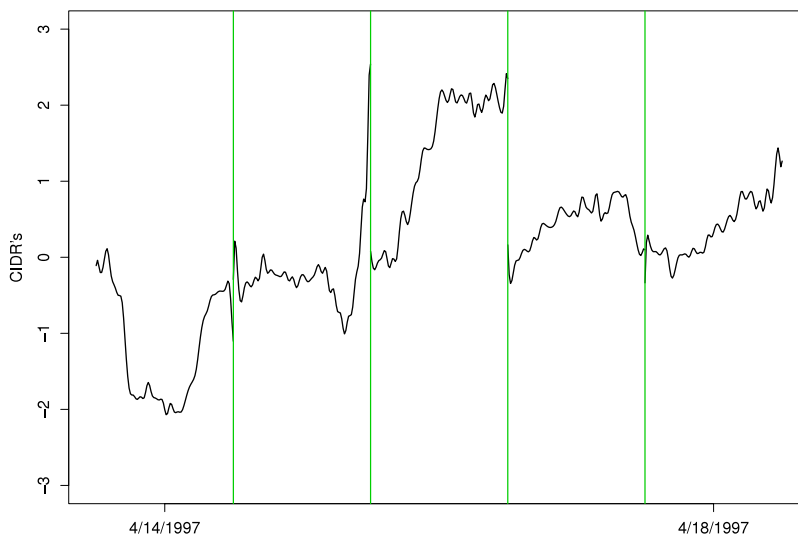


Fig. 3. Five cumulative intraday returns constructed from the intraday prices displayed in Fig. 1.

shapes of the intraday price curves. We note that the CIDR's are not readily comparable to daily returns because they do not include the overnight price change. They are designed to statistically analyze the evolution of the intraday shapes of an asset.

We wish to verify our conjecture of the stationarity of the CIDR's by application of our tests of stationarity. If the conjecture is true, the expectation is that the P-values will be roughly uniformly distributed on $(0, 1)$. Fig. 4 shows results of the test using T_N when applied to sequential segments of the CIDR curves of the Disney stock. We see that the P-values appear to be uniformly distributed which is consistent with the stationarity of the CIDR's. Again, the results for the other eight stocks are very similar.

Acknowledgments

We offer sincere thanks to the editor in charge, the associate editor and three referees. Their profound and constructive comments helped us improve this paper in very substantive ways.

Appendix A. Proofs of the results of Section 2

The proof of Theorem 2.1 is based on an approximation developed in Berkes et al. (2013) (Theorem A.1 below). Define

$$\Gamma(x, t) = \sum_{i=1}^{\infty} \lambda_i^{1/2} W_i(x) \varphi_i(t), \tag{A.1}$$

where W_i are independent and identically distributed Wiener processes (standard Brownian motions). Clearly, $\Gamma(x, t)$ is Gaussian with zero mean and $E \Gamma(x, t) \Gamma(y, s) = \min(x, y) C(t, s)$.

Theorem A.1. *If assumptions (2.1)–(2.4) hold, then*

$$\sum_{\ell=1}^{\infty} \lambda_{\ell} < \infty \tag{A.2}$$

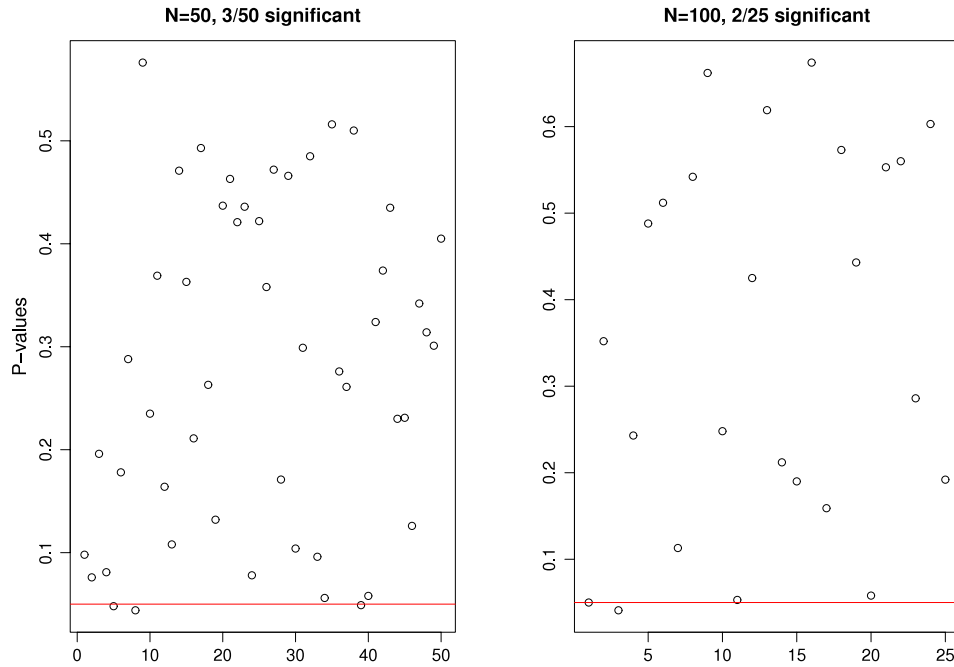


Fig. 4. P-values for consecutive segments of length N of the CIDR curves $R_n(t)$ for the Disney stock. The red line shows the 5% threshold.

and for every N we can define a sequence of Gaussian processes $\Gamma_N(x, t)$ such that

$$\{\Gamma_N(x, t), 0 \leq x, t \leq 1\} \stackrel{D}{=} \{\Gamma(x, t), 0 \leq x, t \leq 1\}$$

and

$$\sup_{0 \leq x \leq 1} \int (V_N(x, t) - \Gamma_N(x, t))^2 dt = o_p(1),$$

where

$$V_N(x, t) = \frac{1}{N^{1/2}} \sum_{i=1}^{\lfloor Nx \rfloor} \eta_i(t).$$

(It follows immediately from (A.2) that $\sup_{0 \leq x \leq 1} \int \Gamma^2(x, t) dt < \infty$ a.s.)

Proof of Theorem 2.1. Let

$$V_N^0(x, t) = V_N(x, t) - xV_N(1, t).$$

Under H_0

$$Z_N(x, t) = V_N^0(x, t) + \mu(t) \left[\frac{\lfloor Nx \rfloor - Nx}{N^{1/2}} \right]$$

and since $\mu \in L^2$ we get

$$\sup_{0 \leq x \leq 1} \|Z_N(x, t) - V_N^0(x, t)\| \leq \frac{1}{N^{1/2}} \|\mu\|.$$

Hence

$$T_N = \iint (V_N^0(x, t))^2 dt dx + o_p(1)$$

and

$$M_N = \iint \left(V_N^0(x, t) - \int V_N^0(y, t) dy \right)^2 dx dt + o_p(1).$$

Applying Theorem A.1 we get immediately that

$$T_N \xrightarrow{D} \iint (\Gamma^0(x, t))^2 dx dt$$

and

$$M_N \xrightarrow{D} \iint \left(\Gamma^0(x, t) - \int \Gamma^0(y, t) dy \right)^2 dx dt,$$

where

$$\Gamma^0(x, t) = \Gamma(x, t) - x\Gamma(1, t).$$

We also note that by the definition of $\Gamma(x, t)$ in (A.1) we have

$$\Gamma^0(x, t) = \sum_{i=1}^{\infty} \lambda_i^{1/2} B_i(x) \varphi_i(t), \tag{A.3}$$

where B_i are independent and identically distributed Brownian bridges. Using the fact that $\{\varphi_i(t), 0 \leq t \leq 1\}_{i=1}^{\infty}$ is an orthonormal system one can easily verify that

$$\iint (\Gamma^0(x, t))^2 dx dt = \sum_{i=1}^{\infty} \lambda_i \int B_i^2(x) dx$$

and

$$\begin{aligned} \iint \left(\Gamma^0(x, t) - \int \Gamma^0(y, t) dy \right)^2 dt dx \\ = \sum_{i=1}^{\infty} \lambda_i \int \left(B_i(x) - \int B_i(y) dy \right)^2 dx. \end{aligned}$$

The following lemma is an immediate consequence of the results in Section 2.7 of Horváth et al. (2013), or of Dunford and Schwartz (1988).

Lemma A.1. If assumptions (2.1)–(2.4), (2.13)–(2.16), (2.17) and H_0 hold, then

$$\max_{1 \leq i \leq d} |\hat{\lambda}_i - \lambda_i| = o_p(1) \quad \text{and} \quad \max_{1 \leq i \leq d} \|\hat{\varphi}_i - \hat{c}_i \varphi_i\| = o_p(1),$$

where $\hat{c}_1, \hat{c}_2, \dots, \hat{c}_d$ are unobservable random signs defined as $\hat{c}_i = \text{sign}(\langle \hat{\varphi}_i, \varphi_i \rangle)$.

Proof of Theorem 2.2. It follows from Theorem A.1 that

$$\sup_{0 \leq x \leq 1} |\langle S_N(x, \cdot) - \Gamma_N^0(x, \cdot), \varphi_i \rangle| \leq \sup_{0 \leq x \leq 1} \|S_N(x, \cdot) - \Gamma_N^0(x, \cdot)\| = o_p(1)$$

and by Lemma A.1 we get

$$\sup_{0 \leq x \leq 1} |\langle \Gamma_N^0(x, \cdot), \hat{\varphi}_i - \hat{c}_i \varphi_i \rangle| \leq \sup_{0 \leq x \leq 1} \|\Gamma_N^0(x, \cdot)\| \|\hat{\varphi}_i - \hat{c}_i \varphi_i\| = o_p(1).$$

It is immediate from (A.3) that for all N

$$\begin{aligned} & \{ \langle \Gamma_N^0(x, \cdot), \varphi_i \rangle, 0 \leq x \leq 1, 1 \leq i \leq d \} \\ & \stackrel{D}{=} \{ \lambda_i^{1/2} B_i(x), 0 \leq x \leq 1, 1 \leq i \leq d \}, \end{aligned}$$

where B_1, B_2, \dots, B_d are independent Brownian bridges. Thus we obtain that

$$\sum_{i=1}^d \frac{1}{\lambda_i} \langle \Gamma_N^0(x, \cdot), \hat{c}_i \varphi_i \rangle^2 \xrightarrow{D[0,1]} \sum_{i=1}^d B_i^2(x). \tag{A.4}$$

The weak convergence in (A.4) now implies (2.18). The same arguments can be used to prove (2.19)–(2.21). \square

Appendix B. Proofs of the results of Section 3

Proof of Theorem 3.1. First we introduce the function

$$\begin{aligned} \delta_N(x, t) &= \mu(t)\{[Nx] - Nx\} + \delta(t)\{([Nx] - k^*) \\ & \quad \times I\{k^* \leq [Nx]\} - x(N - k^*)\}. \end{aligned}$$

Under $H_{A,1}$ we can write

$$Z_N(x, t) = V_N^0(x, t) + N^{-1/2} \delta_N(x, t) \tag{B.1}$$

and therefore

$$\begin{aligned} T_N &= \iint Z_N^2(x, t) dt dx \\ &= \iint (V_N^0(x, t))^2 dt dx + \frac{2}{N^{1/2}} \iint V_N^0(x, t) \delta_N(x, t) dx dt \\ & \quad + \frac{1}{N} \iint \delta_N^2(x, t) dt dx. \end{aligned} \tag{B.2}$$

It follows from Theorem A.1 that

$$\iint (V_N^0(x, t))^2 dt dx = O_p(1). \tag{B.3}$$

It is easy to check that

$$\sup_{0 \leq x \leq 1} \left\| \frac{1}{N} \delta_N(x, t) - \delta_\tau(x, t) \right\| = O\left(\frac{1}{N}\right), \tag{B.4}$$

where $\delta_\tau(x, t)$ is defined in (3.2). Thus applying Theorem A.1 we conclude that

$$\frac{1}{N} \iint V_N^0(x, t) \delta_N(x, t) dx dt \xrightarrow{D} \iint \Gamma^0(x, t) \delta_\tau(x, t) dt dx. \tag{B.5}$$

Also,

$$\begin{aligned} \frac{1}{N} \iint \delta_N^2(x, t) dt dx &= N \int \delta^2(t) dt \left\{ \int_0^\tau x^2 (1 - \tau)^2 dx \right. \\ & \quad \left. + \int_\tau^1 (1 - x)^2 \tau^2 dx \right\} + O(1). \end{aligned} \tag{B.6}$$

Now (3.3) is an immediate consequence of (B.2)–(B.6).

The second part of Theorem 3.1 is proven analogously.

Variations of the limits in Theorem 3.1

The next lemma is used to show that the variances of the limits in Theorem 3.1 are strictly positive.

Lemma B.1. Let Θ be a L^2 valued Gaussian process such that $E\Theta(t) = 0$ and $E\Theta(t)\Theta(s)$ is a strictly positive definite function on $[0, 1]^2$. Let $g \in L^2$. Then $\text{var}(\int \Theta(t)g(t)dt) = 0$ if and only if $g = 0$ a.e.

Proof. By the Karhunen–Loève expansion and the assumption that $E\Theta(t)\Theta(s)$ is strictly positive definite we may write

$$\Theta(t) = \sum_{\ell=1}^{\infty} \rho_\ell N_\ell \phi_\ell(t), \quad 0 \leq t \leq 1,$$

where $\{N_i\}_{i=1}^{\infty}$ are iid standard normal random variables, $\{\phi_i(t)\}_{i=1}^{\infty}$ form an orthonormal basis, and $\rho_i > 0$ for all $i \geq 1$. It follows by a simple calculation that

$$\int \Theta(t)g(t)dt = \sum_{\ell=1}^{\infty} \rho_\ell N_\ell \langle \phi_\ell, g \rangle,$$

and hence

$$\text{var}\left(\int \Theta(t)g(t)dt\right) = \sum_{\ell=1}^{\infty} \rho_\ell^2 \langle \phi_\ell, g \rangle^2.$$

Since $\sum_{\ell=1}^{\infty} \rho_\ell^2 \langle \phi_\ell, g \rangle^2 = 0$ if and only if $g = 0$ a.e., the result follows. \square

It is easy to see that $\iint \Gamma^0(x, t) \delta_\tau(x, t) dt dx$ is a normal random variable with zero mean. Its variance is thus equal to

$$\begin{aligned} & E\left(\iint \Gamma^0(x, t) \delta_\tau(x, t) dt dx\right)^2 \\ &= \iiint \iiint C(t, s) \delta_\tau(x, t) \delta_\tau(y, s) (\min(x, y) - xy) dt ds dx dy \\ &= \left(\iint C(t, s) \delta(t) \delta(s) dt ds\right) \\ & \quad \times \left(\iint \bar{\delta}_\tau(x) \bar{\delta}_\tau(y) (\min(x, y) - xy) dx dy\right), \end{aligned} \tag{B.7}$$

where $\bar{\delta}_\tau(x)$ is defined in (3.18). Similarly to (3.3), the limit in (3.4) is normally distributed with zero mean and variance equal to

$$\begin{aligned} & E\left[\iint \left(\Gamma^0(x, t) - \int \Gamma^0(y, t) dy\right) \right. \\ & \quad \times \left.\left(\delta_\tau(x, t) - \int \delta_\tau(y, t) dy\right) dt dx\right]^2 \\ &= \left(\iint C(t, s) \delta(t) \delta(s) dt ds\right) \left(\iint \bar{\delta}_\tau(x) \bar{\delta}_\tau(y) \right. \\ & \quad \times \left[\min(x, y) - xy - \int (\min(y, z) - yz) dz \right. \\ & \quad \left. - \int (\min(x, z) - xz) dz + \iint (\min(z, z') - zz') dz dz'\right] dx dy) \\ &= \left(\iint C(t, s) \delta(t) \delta(s) dt ds\right) \\ & \quad \times \left(\iint \bar{\delta}_\tau(x) \bar{\delta}_\tau(y) \left[\min(x, y) - xy - \frac{y(1-y)}{2} \right. \right. \\ & \quad \left. \left. - \frac{x(1-x)}{2} + \frac{1}{12}\right] dx dy\right). \end{aligned}$$

If the bivariate function $C(t, s)$ is strictly positive definite, then $\iint C(t, s)\delta(t)\delta(s)dtds > 0$ if $\delta(t)$ is not the 0 function in L^2 . Observing that $\iint \bar{\delta}_\tau(x)\bar{\delta}_\tau(y)(\min(x, y) - xy)dxdy = \text{var}(\int B(x) \times \bar{\delta}_\tau(x)dx)$, where B is a Brownian bridge, the positivity of (B.7) follows by Lemma B.1 since $\bar{\delta}_\tau(x)$ is not the zero function and the covariance function of the Brownian bridge is strictly positive definite. A similar application of Lemma B.1 yields that

$$\iint \bar{\delta}_\tau(x)\bar{\delta}_\tau(y) \times \left[\min(x, y) - xy - \frac{y(1-y)}{2} - \frac{x(1-x)}{2} + \frac{1}{12} \right] dxdy > 0.$$

Lemma B.2. *If assumptions (2.1)–(2.4), (2.13)–(2.16), (3.1) and $H_{A,1}$ hold, then*

$$\left\| \hat{C}_N(t, s) - \left(2\tau(1-\tau)\delta(t)\delta(s) \sum_{i=1}^N K(i/h) + \bar{C}_N(t, s) \right) \right\| = O_p(h/N^{1/2}),$$

where

$$\bar{C}_N(t, s) = \bar{\gamma}_0(t, s) + \sum_{i=1}^{N-1} K\left(\frac{i}{h}\right) \{\bar{\gamma}_i(t, s) + \bar{\gamma}_i(s, t)\} \tag{B.8}$$

with

$$\bar{\gamma}_i(t, s) = \frac{1}{N} \sum_{j=i+1}^N (\eta_j(t) - \bar{\eta}_N(t)) \times (\eta_{j-i}(s) - \bar{\eta}_N(s)), \quad 0 \leq i \leq N-1.$$

Proof. First we write with $\mu_i(t) = EX_i(t)$ and observe that

$$\begin{aligned} \hat{\gamma}_i(t, s) &= \frac{1}{N} \sum_{j=i+1}^N (\eta_j(t) - \bar{\eta}_N(t) - [\bar{\mu}_N(t) - \mu_i(t)]) \\ &\quad \times (\eta_{j-i}(s) - \bar{\eta}_N(s) - [\bar{\mu}_N(s) - \mu_{j-i}(s)]) \\ &= \frac{1}{N} \sum_{j=i+1}^N (\eta_j(t) - \bar{\eta}_N(t)) (\eta_{j-i}(s) - \bar{\eta}_N(s)) \\ &\quad + \frac{1}{N} \sum_{j=i+1}^N (\eta_j(t) - \bar{\eta}_N(t)) (\mu_{j-i}(s) - \bar{\mu}_N(s)) \\ &\quad + \frac{1}{N} \sum_{j=i+1}^N (\mu_j(t) - \bar{\mu}_N(t)) (\eta_{j-i}(s) - \bar{\eta}_N(s)) \\ &\quad + \frac{1}{N} \sum_{j=i+1}^N (\mu_j(t) - \bar{\mu}_N(t)) (\mu_{j-i}(s) - \bar{\mu}_N(s)) \\ &= \bar{\gamma}_i(t, s) + \hat{\gamma}_i^{(1)}(t, s) + \hat{\gamma}_i^{(2)}(t, s) + \hat{\gamma}_i^{(3)}(t, s) \end{aligned}$$

with

$$\bar{\eta}_N(t) = \frac{1}{N} \sum_{\ell=1}^N \eta_\ell \quad \text{and} \quad \bar{\mu}_N(t) = \mu(t) + \frac{N - \lfloor N\tau \rfloor}{N} \delta(t).$$

By the triangle inequality we have

$$\left\| \hat{\gamma}_0^{(1)}(t, s) + \sum_{i=1}^{N-1} K\left(\frac{i}{h}\right) (\hat{\gamma}_i^{(1)}(t, s) + \hat{\gamma}_i^{(1)}(s, t)) \right\|$$

$$\leq \left\| \hat{\gamma}_0^{(1)}(t, s) \right\| + \left\| \sum_{i=1}^{N-1} K\left(\frac{i}{h}\right) \hat{\gamma}_i^{(1)}(t, s) \right\| + \left\| \sum_{i=1}^{N-1} K\left(\frac{i}{h}\right) \hat{\gamma}_i^{(1)}(s, t) \right\|.$$

Using Theorem A.1 we get

$$\left\| \hat{\gamma}_0^{(1)}(t, s) \right\| = O_p(N^{-1/2}).$$

Using again the triangle inequality we obtain that

$$E \left\| \sum_{i=1}^{N-1} K(i/h) \hat{\gamma}_i^{(1)}(t, s) \right\| \leq \sum_{i=1}^{N-1} K(i/h) E \|\hat{\gamma}_i^{(1)}(t, s)\|. \tag{B.9}$$

Furthermore by an application of the Cauchy–Schwarz inequality

$$\begin{aligned} E \|\hat{\gamma}_i^{(1)}(t, s)\| &\leq \left\| \frac{1}{N} \sum_{j=i+1}^N (\mu_{j-i}(s) - \bar{\mu}_N(s)) \right\| \\ &\quad \times E \left\| \frac{1}{N} \sum_{j=i+1}^N (\eta_j(t) - \bar{\eta}_N(t)) \right\|. \end{aligned}$$

It is clear that

$$\max_{1 \leq i \leq N} \left\| \frac{1}{N} \sum_{j=i+1}^N (\mu_{j-i}(s) - \bar{\mu}_N(s)) \right\| = O(1),$$

and by Berkes et al. (2013)

$$\max_{1 \leq i \leq N} E \left\| \frac{1}{N} \sum_{j=i+1}^N (\eta_j(t) - \bar{\eta}_N(t)) \right\| = O(N^{-1/2}).$$

Combining these bounds with (B.9) and assumptions (2.13)–(2.15) gives

$$E \left\| \sum_{i=1}^{N-1} K(i/h) \hat{\gamma}_i^{(1)}(t, s) \right\| = O(h/N^{1/2}),$$

and hence by Markov’s inequality

$$\left\| \sum_{i=1}^{N-1} K(i/h) \hat{\gamma}_i^{(1)}(t, s) \right\| = O_p(h/N^{1/2}).$$

Thus we conclude

$$\begin{aligned} \left\| \hat{\gamma}_0^{(1)}(t, s) + \sum_{i=1}^{N-1} K\left(\frac{i}{h}\right) (\hat{\gamma}_i^{(1)}(t, s) + \hat{\gamma}_i^{(1)}(s, t)) \right\| \\ = O_p(h/N^{1/2}). \end{aligned} \tag{B.10}$$

Similarly to (B.10) we have

$$\begin{aligned} \left\| \hat{\gamma}_0^{(2)}(t, s) + \sum_{i=1}^{N-1} K\left(\frac{i}{h}\right) (\hat{\gamma}_i^{(2)}(t, s) + \hat{\gamma}_i^{(2)}(s, t)) \right\| \\ = O_p(h/N^{1/2}). \end{aligned} \tag{B.11}$$

Using the definition of $\bar{\mu}_N(t)$ and $H_{A,1}$ we obtain that

$$\max_{0 \leq i \leq h} \|\hat{\gamma}_0^{(3)}(t, s) - \tau(1-\tau)\delta(t)\delta(s)\| = O(h/N). \tag{B.12}$$

The lemma now follows from (B.10)–(B.12). □

Proof of Theorem 3.2. The proof of Theorem 3.2 is based on the asymptotic properties of \hat{C}_N under $H_{A,1}$. It follows from Lemma B.2 that (3.5) and (3.6) hold assuming only (2.16). We write by (B.1)

$$\langle Z_N(x, \cdot), \hat{\varphi}_1 \rangle^2 = \langle V_N^0(x, \cdot), \hat{\varphi}_1 \rangle^2 + N^{-1} \langle \delta_N(x, \cdot), \hat{\varphi}_1 \rangle^2 + 2 \langle V_N^0(x, \cdot), \hat{\varphi}_1 \rangle N^{-1/2} \langle \delta_N(x, \cdot), \hat{\varphi}_1 \rangle.$$

Combining Theorem A.1 with the Cauchy–Schwarz inequality we get

$$\sup_{0 \leq x \leq 1} |\langle V_N^0(x, \cdot), \hat{\varphi}_1 \rangle| \leq \sup_{0 \leq x \leq 1} \|V_N^0(x, \cdot)\| = O_P(1).$$

Using (B.4) we conclude

$$\int N^{-1} \langle \delta_N(x, \cdot), \hat{\varphi}_1 \rangle^2 dx = \frac{N}{3} \tau^2 (1 - \tau)^2 \langle \delta, \hat{\varphi}_1 \rangle^2 (1 + O_P(1/N)).$$

Theorem A.1 and (3.6) yield

$$\begin{aligned} & N^{1/2} \int \langle V_N^0(x, \cdot), \hat{\varphi}_1 \rangle \langle \delta_N(x, \cdot), \hat{\varphi}_1 \rangle dx \\ & \xrightarrow{D} \frac{1}{\|\delta\|} \int \langle \Gamma_0(x, \cdot), \delta \rangle \langle \delta_\tau(x, \cdot), \delta \rangle dx \\ & = \int \left\{ \int \Gamma^0(x, t) \delta(t) dt \right\} [(x - \tau)I\{x \geq \tau\} - x(1 - \tau)] dx \\ & = \iint \Gamma^0(x, t) \delta_\tau(x, t) dx dt. \end{aligned}$$

This completes the proof of (3.8). It follows from (3.8) that

$$\begin{aligned} & \frac{\hat{\lambda}_1}{N^{1/2}} \left\{ T_N^0(1) - \frac{N}{3\hat{\lambda}_1} \tau^2 (1 - \tau)^2 \langle \delta, \hat{\varphi}_1 \rangle^2 \right\} \\ & \xrightarrow{D} 2 \iint \Gamma^0(x, t) \delta_\tau(x, t) dx dt, \end{aligned}$$

and therefore (3.8) implies (3.10). (3.9) and (3.11) are proven similarly.

If in addition we assume that $h/N^{1/2} \rightarrow 0$ as $N \rightarrow \infty$ then by Lemma B.2 and Dunford and Schwartz (1988) we have (3.5), (3.6), and for every fixed $i \geq 2$,

$$\hat{\lambda}_i \xrightarrow{P} \bar{\lambda}_i, \tag{B.13}$$

where $\bar{\lambda}_2 \geq \bar{\lambda}_3 \geq \dots \geq 0$ (might be different from $\lambda_i, i \geq 2$),

$$\|\hat{\varphi}_i(t) - \hat{c}_i \bar{\varphi}_i\| = o_P(1), \quad i \geq 2, \tag{B.14}$$

with some functions $\bar{\varphi}_2, \bar{\varphi}_3, \dots$, where $\hat{c}_i = \text{sign}(\langle \hat{\varphi}_i, \bar{\varphi}_i \rangle)$. (Of course, $\bar{\varphi}_i$ is only defined if $\hat{\lambda}_i > 0$.) Using again (B.1) with Theorem A.1 and (B.14), we obtain that

$$\int \langle Z_N(x, \cdot), \hat{\varphi}_i \rangle^2 dx = \frac{N}{3} \tau^2 (1 - \tau)^2 \langle \delta, \hat{\varphi}_i \rangle^2 + O_P(N^{1/2}).$$

Since δ and $\bar{\varphi}_i$ are orthogonal for all $i \geq 2$, (B.14) implies $\langle \delta, \hat{\varphi}_i \rangle = o_P(1)$. Hence (3.12) follows from (3.8). The results in (3.13)–(3.15) can be established similarly so the proofs are omitted.

Proof of Remark 3.1. Let

$$\begin{aligned} \beta_N(x, t) &= \mu(t) \{ [Nx] - Nx \} + \delta_N^*(t) \{ ([Nx] - k^*) \\ & \quad \times I\{k^* \leq [Nx]\} - x(N - k^*) \}. \end{aligned}$$

Using (B.2) with $\delta_N(x, t)$ replaced with $\beta_N(x, t)$ and Theorem A.1 we get

$$\begin{aligned} T_N - \frac{1}{N} \|\beta_N\|^2 &= \iint (\Gamma_N^0(x, t))^2 dt dx (1 + o_P(1)) \\ & \quad + 2N^{1/2} \iint \Gamma_N^0(x, t) \delta_N^*(t) \bar{\delta}_\tau(x) dt dx \\ & \quad \times (1 + o_P(1)). \end{aligned} \tag{B.15}$$

By the Cauchy–Schwarz inequality

$$\iint \Gamma_N^0(x, t) \delta_N^*(t) \bar{\delta}_\tau(x) dt dx = O_P(\|\delta_N^*\|). \tag{B.16}$$

Elementary arguments show that

$$\frac{1}{N} \|\beta_N\|^2 = \|\bar{\delta}_\tau\|^2 N \|\delta_N^*\|^2 (1 + o(1)), \tag{B.17}$$

as $N \rightarrow \infty$. If $N^{1/2} \|\delta_N^*\| \rightarrow 0$ as $N \rightarrow \infty$ then by (B.15)–(B.17), we obtain immediately that $T_N \xrightarrow{D} \iint (\Gamma^0(x, t))^2 dt dx$. If $N^{1/2} \|\delta_N^*\| \rightarrow \infty$, then again by (B.15)–(B.17) we see that $T_N \xrightarrow{P} \infty$. Since for every fixed N , $\iint \Gamma_N^0(x, t) \delta_N^*(t) \bar{\delta}_\tau(x) dt dx$ is normal with zero mean and variance $\iint (\min(x, y) - xy) \bar{\delta}_\tau(x) \bar{\delta}_\tau(y) dx dy \times \iint \delta_N^*(t) \delta_N^*(s) C(t, s) dt ds$, hence (3.16) follows. In the case when $N^{1/2} \delta_N^* \xrightarrow{L^2} \delta^*$, it follows from (B.17) that $(1/N) \|\beta_N\|^2 \rightarrow \zeta = \|\bar{\delta}_\tau\|^2 \|\delta^*\|^2 > 0$. Now by (B.15) and the representation of Γ_N^0 in (A.3) we conclude

$$\begin{aligned} T_N & \xrightarrow{D} \zeta (1 + o(1)) \\ & \quad + \sum_{\ell=1}^{\infty} \lambda_\ell^{1/2} \left[\lambda_\ell^{1/2} \int B_\ell^2(x) dx + \int B_\ell(x) \bar{\delta}_\tau(x) dx \right. \\ & \quad \left. \times \int \varphi_\ell(t) N^{1/2} \delta_N^*(t) dt \right] (1 + o_P(1)) \\ & \rightarrow \zeta + \sum_{\ell=1}^{\infty} \left\{ \lambda_\ell \|B_\ell\|^2 + 2\lambda_\ell^{1/2} \langle B_\ell, \bar{\delta}_\tau \rangle \langle \varphi_\ell, \delta^* \rangle \right\}, \end{aligned}$$

which completes the proof of (3.17). \square

Lemma B.3. If assumptions (2.1)–(2.4) hold, then

$$\sup_{0 \leq x \leq 1} \int \left(U_N(x, t) - \int_0^x \Gamma_N(u, t) du \right)^2 dt = o_P(1), \tag{B.18}$$

where

$$U_N(x, t) = \frac{1}{N^{3/2}} \sum_{k=1}^{[Nx]} \sum_{i=1}^k \eta_i(t),$$

and the Gaussian processes $\Gamma_N(x, t)$ are defined in Theorem A.1.

Proof. It is enough to verify that

$$\begin{aligned} & \sup_{0 \leq x \leq 1} \int \left(U_N(x, t) - \int_0^x V_N(u, t) du \right)^2 dt \\ & = \sup_{0 \leq x \leq 1} \left\| U_N(x, \cdot) - \int_0^x V_N(u, \cdot) \right\|^2 = o_P(1) \end{aligned}$$

and

$$\sup_{0 \leq x \leq 1} \int \left(\int_0^x \{V_N(u, t) - \Gamma_N(u, t)\} du \right)^2 dt = o_P(1).$$

Elementary arguments yield

$$\left| U_N(x, t) - \int_0^x V_N(u, t) du \right| \leq \frac{1}{N^{3/2}} \left| \sum_{i=1}^{[Nx]} \eta_i(t) \right|.$$

It follows from Theorem A.1 that

$$\sup_{0 \leq x \leq 1} \left\| N^{-1/2} \sum_{i=1}^{[Nx]} \eta_i(\cdot) \right\| = O_P(1),$$

and therefore

$$\sup_{0 \leq x \leq 1} \left\| U_N(x, \cdot) - \int_0^x V_N(u, \cdot) du \right\| = O_p \left(\frac{1}{N} \right).$$

Using the Cauchy–Schwarz inequality with **Theorem A.1** we conclude

$$\begin{aligned} & \int \left(\int_0^x (V_N(u, t) - \Gamma_N(u, t)) du \right)^2 dt \\ & \leq \int \int_0^x (V_N(u, t) - \Gamma_N(u, t))^2 dudt \\ & \leq \int \int (V_N(u, t) - \Gamma_N(u, t))^2 dudt = o_p(1). \end{aligned}$$

Now the proof of **Lemma B.3** is complete. \square

Proof of Theorem 3.3. First we note that under $H_{A,2}$ we have

$$\frac{1}{N^{3/2}} \sum_{k=1}^{\lfloor Nx \rfloor} X_k(t) = U_N(x, t) + \frac{\lfloor Nx \rfloor}{N^{3/2}} \mu(t). \tag{B.19}$$

Therefore

$$\frac{1}{N} Z_N(x, t) = U_N(x, t) - xU_N(1, t) + \frac{\lfloor Nx \rfloor - xN}{N^{3/2}} \mu(t).$$

Using **(B.19)** we get via the Cauchy–Schwarz inequality

$$\begin{aligned} & \left| \iint \left(\frac{1}{N} Z_N(x, t) \right)^2 dt dx - \iint (U_N(x, t) - xU_N(1, t))^2 dt dx \right| \\ & \leq \iint \left\{ \frac{1}{N} Z_N(x, t) - [U_N(x, t) - xU_N(1, t)] \right\}^2 dt dx \\ & \quad + 2 \iint \left| \frac{1}{N} Z_N(x, t) - [U_N(x, t) - xU_N(1, t)] \right| \\ & \quad \times |U_N(x, t) - xU_N(1, t)| dt dx \\ & \leq o_p(1) + o_p(1) \left\{ \iint (U_N(x, t) - xU_N(1, t))^2 dt dx \right\}^{1/2} = o_p(1), \end{aligned}$$

since by **Lemma B.3**

$$\iint (U_N(x, t) - xU_N(1, t))^2 dt dx = O_p(1).$$

It also follows from **Lemma B.3** that

$$\iint (U_N(x, t) - xU_N(1, t))^2 dt dx \xrightarrow{D} \iint \Delta^2(x, t) dt dx,$$

which completes the proof of **(3.20)**. The proof of **(3.21)** is similar to that of **(3.20)** and therefore the details are omitted.

Lemma B.4. Define

$$I_N(z, t) = \int_0^z \Gamma_N(u, t) du - \int \left\{ \int_0^v \Gamma_N(u, t) du \right\} dv,$$

where the Gaussian processes $\Gamma_N(x, t)$ are defined in **Theorem A.1**. Let

$$Q_N(t, s) = 2 \left(\int_0^c K(w) dw \right) \int_0^1 I_N(z, t) I_N(z, s) dz.$$

If assumptions **(2.1)–(2.4)**, **(2.13)–(2.16)** and $H_{A,2}$ hold, then

$$\left\| \frac{1}{Nh} \hat{C}_N(t, s) - Q_N(t, s) \right\| = o_p(1).$$

Proof. Since

$$\bar{X}_N(t) = \mu(t) + \frac{1}{N} \sum_{j=1}^N \sum_{\ell=1}^j \eta_\ell(t),$$

Theorem A.1 yields

$$\left\| N^{-1/2} (\bar{X}_N(t) - \mu(t)) - \int \left\{ \int_0^v \Gamma_N(u, t) du \right\} dv \right\| = o_p(1),$$

resulting in

$$\begin{aligned} & \max_{1 \leq i \leq N-1} \left\| \frac{1}{N} \hat{\gamma}_i(t, s) \right. \\ & \quad - \frac{1}{N} \sum_{j=i+1}^N \left(\int_0^{j/N} \Gamma_N(u, t) du - \int \left\{ \int_0^v \Gamma_N(u, t) du \right\} dv \right) \\ & \quad \times \left. \left(\int_0^{(j-i)/N} \Gamma_N(u, s) du - \int \left\{ \int_0^v \Gamma_N(u, s) du \right\} dv \right) \right\| \\ & = o_p(1). \end{aligned} \tag{B.20}$$

Next we use the almost sure continuity with $\Gamma_N(0, t) = 0$ to conclude

$$\begin{aligned} & \max_{1 \leq i \leq ch} \left\| \frac{1}{N} \sum_{j=i+1}^N \left(\int_0^{j/N} \Gamma_N(u, t) du - \int \left\{ \int_0^v \Gamma_N(u, t) du \right\} dv \right) \right. \\ & \quad \times \left. \left(\int_0^{(j-i)/N} \Gamma_N(u, s) du - \int \left\{ \int_0^v \Gamma_N(u, s) du \right\} dv \right) \right. \\ & \quad - \frac{1}{N} \sum_{j=i+1}^N \left(\int_0^{j/N} \Gamma_N(u, t) du - \int \left\{ \int_0^v \Gamma_N(u, t) du \right\} dv \right) \\ & \quad \times \left. \left(\int_0^{j/N} \Gamma_N(u, s) du - \int \left\{ \int_0^v \Gamma_N(u, s) du \right\} dv \right) \right\| \\ & = o_p(1). \end{aligned} \tag{B.21}$$

Putting together **(B.20)** and **(B.21)** we get

$$\begin{aligned} & \max_{1 \leq i \leq ch} \left\| \frac{1}{N} \hat{\gamma}_i(t, s) - \int_{i/N}^1 \left[\left(\int_0^z \Gamma_N(u, t) du \right) \right. \right. \\ & \quad - \left. \left. \int \left\{ \int_0^v \Gamma_N(u, t) du \right\} dv \right) \right. \right. \\ & \quad \times \left. \left. \left(\int_0^z \Gamma_N(u, s) du - \int \left\{ \int_0^v \Gamma_N(u, s) du \right\} dv \right) \right] dz \right\| \\ & = o_p(1) \end{aligned}$$

and

$$\begin{aligned} & \max_{1 \leq i \leq ch} \left\| \int_{i/N}^1 \left[\left(\int_0^z \Gamma_N(u, t) du - \int \left\{ \int_0^v \Gamma_N(u, t) du \right\} dv \right) \right. \right. \\ & \quad \times \left. \left. \left(\int_0^z \Gamma_N(u, s) du - \int \left\{ \int_0^v \Gamma_N(u, s) du \right\} dv \right) \right] dz \right. \\ & \quad - \int_0^1 \left[\left(\int_0^z \Gamma_N(u, t) du - \int \left\{ \int_0^v \Gamma_N(u, t) du \right\} dv \right) \right. \\ & \quad \times \left. \left. \left(\int_0^z \Gamma_N(u, s) du - \int \left\{ \int_0^v \Gamma_N(u, s) du \right\} dv \right) \right] dz \right\| = o_p(1). \end{aligned}$$

Since K satisfies conditions **(2.14)** and **(2.15)**, the proof of **Lemma B.4** is complete. \square

Lemma B.5. For every $N \geq 1$ we have

$$\{Q_N(t, s), 0 \leq t, s \leq 1\} \stackrel{D}{=} 2 \int_0^c K(w)dw \times \left\{ \sum_{i,j=1}^{\infty} \lambda_i^{1/2} \lambda_j^{1/2} \varphi_i(t) \varphi_j(s) v_{i,j} \right\}, \quad (B.22)$$

where $\lambda_1, \lambda_2, \dots, \varphi_1, \varphi_2, \dots$ are defined in (2.9) and for every $1 \leq i, j < \infty$

$$v_{i,j} \stackrel{D}{=} \int \left[\left\{ \int_0^z W_i(u)du - \int \left(\int_0^v W_i(u)du \right) dv \right\} \times \left\{ \int_0^z W_j(u)du - \int \left(\int_0^v W_j(u)du \right) dv \right\} \right] dz,$$

where W_1, W_2, \dots are independent Wiener processes. Also, $Q_N(t, s)$ is a positive non-negative function with for all N with probability one.

Proof. The representation in (B.22) is an immediate consequence of (A.1). It follows from (B.22) that $Q_N(t, s)$ is symmetric and $Q_N \in L^2$ with probability one. Also for any $g \in L^2$ we have

$$\begin{aligned} & \int \int Q_N(t, s) g(t) g(s) dt ds \\ &= \int \left(\int \sum_{i=1}^{\infty} \lambda_i^{1/2} \varphi_i(t) \left\{ \int_0^z W_i(u)du - \int \left(\int_0^v W_i(u)du \right) dv \right\} g(t) dt \right)^2 dz \\ & \geq 0, \end{aligned}$$

completing the proof. \square

Proof of Theorem 3.4. The result follows immediately from the proofs of Lemmas B.3 and B.4.

Proof of Theorem 3.5. The result in Theorem 3.4 and (3.22) yield that there are processes $\Gamma_N(x, t), \Delta_N(x, t), Q_N(t, s)$ such that

$$\begin{aligned} & \{\Gamma_N(x, t), \Delta_N(x, t), Q_N(t, s), 0 \leq x, t, s \leq 1\} \\ & \stackrel{D}{=} \{\Gamma(x, t), \Delta(x, t), Q_N(t, s), 0 \leq x, t, s \leq 1\} \end{aligned}$$

and

$$\begin{aligned} & \max_{0 \leq x \leq 1} \left\| \frac{1}{N} Z_N(x, t) - \Delta_N(x, t) \right\| = o_p(1), \\ & \left\| \frac{1}{Nh} \hat{C}_N(t, s) - Q_N(t, s) \right\| = o_p(1). \end{aligned}$$

Similarly to (3.22) we define $\lambda_{1,N}^* \geq \lambda_{2,N}^* \geq \dots$ and random functions $\varphi_{1,N}^*(t), \varphi_{2,N}^*(t), \dots$ satisfying

$$\lambda_{i,N}^* \varphi_{i,N}^*(t) = \int Q_N(t, s) \varphi_{i,N}^*(s) ds, \quad 1 \leq i < \infty. \quad (B.23)$$

Hence

$$\max_{1 \leq i \leq d} |\hat{\lambda}_i - \lambda_{i,N}^*| = o_p(1)$$

and

$$\max_{1 \leq i \leq d} \|\hat{\varphi}_i - \hat{c}_i \varphi_{i,N}^*\| = o_p(1),$$

where $\hat{c}_1, \hat{c}_2, \dots$ are random signs. By construction,

$$\begin{aligned} & \{\Delta_N(x, t), Q_N(t, s), \lambda_{1,N}^*, \dots, \lambda_{d,N}^*, (\varphi_{1,N}^*(t))^2, \\ & \dots, (\varphi_{d,N}^*(t))^2, 0 \leq x, t, s \leq 1\} \\ & \stackrel{D}{=} \{\Delta(x, t), Q_N(t, s), \lambda_1^*, \dots, \lambda_d^*, (\varphi_1^*(t))^2, \\ & \dots, (\varphi_d^*(t))^2, 0 \leq x, t, s \leq 1\}, \end{aligned}$$

which completes the proof.

References

Andrews, D.W.K., 1991. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* 59, 817–858.

Andrews, D.W.K., Monahan, J.C., 1992. An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. *Econometrica* 60, 953–966.

Antoniadis, A., Sapatinas, T., 2003. Wavelet methods for continuous time prediction using Hilbert-valued autoregressive processes. *J. Multivariate Anal.* 87, 133–158.

Antoniadis, A., Paparoditis, E., Sapatinas, T., 2006. A functional wavelet–kernel approach for time series prediction. *J. Roy. Statist. Soc. Ser. B* 68, 837–857.

Aue, A., Hörmann, S., Horváth, L., Reimherr, M., 2009. Break detection in the covariance structure of multivariate time series models. *Ann. Statist.* 37, 4046–4087.

Barndorff-Nielsen, O.E., Shephard, N., 2004. Econometric analysis of realized covariance: High frequency based covariance, regression and correlation in financial economics. *Econometrica* 72, 885–925.

Berkes, I., Gabrys, R., Horváth, L., Kokoszka, P., 2009. Detecting changes in the mean of functional observations. *Journal of the Royal Statistical Society (B)* 71, 927–946.

Berkes, I., Horváth, L., Rice, G., 2013. Weak invariance principles for sums of dependent random functions. *Stochastic Process. Appl.* 123 (2), 385–403.

Bosq, D., 2000. *Linear Processes in Function Spaces*. Springer, New York.

de Jong, R.M., Amsler, C., Schmidt, P., 1997. A robust version of the KPSS test based on indicators. *J. Econometrics* 137, 311–333.

Dickey, D.A., Fuller, W.A., 1979. Distributions of the estimators for autoregressive time series with a unit root. *J. Amer. Statist. Assoc.* 74, 427–431.

Dickey, D.A., Fuller, W.A., 1981. Likelihood ratio statistics for autoregressive time series with unit root. *Econometrica* 49, 1057–1074.

Dunford, N., Schwartz, J.T., 1988. *Linear Operators, Parts I and II*. Wiley.

Dwivedi, Y., Subba Rao, S., 2011. A test for second order stationarity based on the discrete Fourier transform. *J. Time Ser. Anal.* 32, 68–91.

Gabrys, R., Horváth, L., Kokoszka, P., 2010. Tests for error correlation in the functional linear model. *J. Amer. Statist. Assoc.* 105, 1113–1125.

Gabrys, R., Hörmann, S., Kokoszka, P., 2013. Monitoring the intraday volatility pattern. *J. Time Ser. Econom.* (forthcoming).

Giraitis, L., Kokoszka, P.S., Leipus, R., Teyssière, G., 2003. Rescaled variance and related tests for long memory in volatility and levels. *Journal of Econometrics* 112, 265–294.

Granger, C.W.J., Hatanaka, M., 1964. *Spectral Analysis of Economic Time Series*. Princeton University Press.

Grenander, U., Rosenblatt, M., 1957. *Statistical Analysis of Stationary Time Series*. Wiley, New York.

Hays, S., Shen, H., Huang, J.Z., 2012. Functional dynamic factor models with application to yield curve forecasting. *Ann. Appl. Stat.* 6, 870–894.

Hörmann, S., Kokoszka, P., 2010. Weakly dependent functional data. *Ann. Statist.* 38, 1845–1884.

Hörmann, S., Kokoszka, P., 2012. Functional time series. In: Rao, C.R., Subba Rao, T. (Eds.), *Handbook of Statistics*. In: Time Series, vol. 30. Elsevier.

Hörmann, S., Horváth, L., Reeder, R., 2013. A functional version of the ARCH model. *Econometric Theory* 29, 267–288.

Hörmann, S., Kidziński, L., Hallin, M., 2013. Dynamic functional principal components. Technical report. Université libre de Bruxelles.

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

Horváth, L., Hušková, M., Kokoszka, P., 2010. Testing the stability of the functional autoregressive process. *J. Multivariate Anal.* 101, 352–367.

Horváth, L., Kokoszka, P., Reeder, R., 2013. Estimation of the mean of functional time series and a two sample problem. *Journal of the Royal Statistical Society (B)* 75, 103–122.

Jönsson, K., (2006). Finite-sample stability of the KPSS test. Working Paper 2006:23, Department of Economics, Lund University.

Jönsson, K., 2011. Testing stationarity in small- and medium-sized samples when disturbances are serially correlated. *Oxf. Bull. Econom. Stat.* 73, 669–690.

Kargin, V., Onatski, A., 2008. Curve forecasting by functional autoregression. *J. Multivariate Anal.* 99, 2508–2526.

Kiefer, J., 1959. *K-sample analogues of the Kolmogorov-Smirnov and Cramér-Mises tests*. *Ann. Math. Statist.* 30, 420–447.

Kokoszka, P., Reimherr, M., 2013. Predictability of shapes of intraday price curves. *Econom. J.* 16, 285–308.

- Kokoszka, P., Miao, H., Zhang, X., 2013. Functional multifactor regression for intraday price curves Technical report. Colorado State University.
- Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., Shin, Y., 1992. Testing the null hypothesis of stationarity against the alternative of a unit root: how sure are we that economic time series have a unit root?. *Journal of Econometrics* 54, 159–178.
- Lee, D., Schmidt, P., 1996. On the power of the KPSS test of stationarity against fractionally integrated alternatives. *Journal of Econometrics* 73, 285–302.
- Lo, A.W., 1991. Long-term memory in stock market prices. *Econometrica* 59, 1279–1313.
- McMurry, T., Politis, D.N., 2010. Resampling methods for functional data. In: Ferraty, F., Romain, Y. (Eds.), *Oxford Handbook on Statistics and FDA*. Oxford University Press.
- Müller, H-G., Sen, R., Stadtmüller, U., 2011. Functional data analysis for volatility. *Journal of Econometrics* 165, 233–245.
- Panaretos, V.M., Tavakoli, S., 2013a. Fourier analysis of stationary time series in function space. *Ann. Statist.* 41, 568–603.
- Panaretos, V.M., Tavakoli, S., 2013b. Cramér–Karhunen–Loève representation and harmonic principal component analysis of functional time series. *Stochastic Process. Appl.* 123, 2779–2807.
- Pelagatti, M.M., Sen, P.K., 2013. Rank tests for short memory stationarity. *Journal of Econometrics* 172, 90–105.
- Politis, D.N., 2003. Adaptive bandwidth choice. *J. Nonparametr. Stat.* 25, 517–533.
- Politis, D.N., 2011. Higher-order accurate, positive semidefinite estimation of large sample covariance and spectral density matrices. *Econometric Theory* 27, 1469–4360.
- Pötscher, B., Prucha, I., 1997. *Dynamic Non-linear Econometric Models. Asymptotic Theory*. Springer.
- Priestley, M.B., Subba Rao, T., 1969. A test for non-stationarity of time-series. *Journal of the Royal Statistical Society (B)* 31, 140–149.
- Ramsay, J.O., Silverman, B.W., 2005. *Functional Data Analysis*. Springer.
- Said, S.E., Dickey, D.A., 1984. Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika* 71, 599–608.
- Shao, X., Wu, W.B., 2007. Asymptotic spectral theory for nonlinear time series. *Ann. Statist.* 35, 1773–1801.
- Shorack, G.R., Wellner, J.A., 1986. *Empirical Processes with Applications to Statistics*. Wiley.
- Teräsvirta, T., Tjøstheim, D., Granger, C.W.J., 2010. *Modeling Nonlinear Economic Time Series. Advanced Texts in Econometrics*. Oxford University Press.
- Wang, Y., Zou, J., 2010. Vast volatility matrix estimation for high-frequency financial data. *Ann. Statist.* 38, 953–978.
- Wu, W., 2005. Proceedings of The National Academy of Sciences of the United States. In: *Nonlinear System Theory: Another Look at Dependence*, vol. 102. National Academy of Sciences.
- Zhang, X., Shao, X., Hayhoe, K., Wuebbles, D., 2011. Testing the structural stability of temporally dependent functional observations and application to climate projections. *Electron. J. Stat.* 5, 1765–1796.