Change point analysis of covariance functions: a weighted cumulative sum approach

Lajos Horváth[a], Gregory Rice[b,*], Yuqian Zhao[c]

[a] *Department of Mathematics, University of Utah, Salt Lake City, UT 84112–0090 USA*
[b] *Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Canada*
[c] *Essex Business School, University of Essex, Colchester, CO4 3SQ, United Kingdom*

Abstract

We develop and study change point detection and estimation procedures for the covariance kernel of functional data based on the norms of a generally weighted process of partial sample estimates. It is shown under mild weak dependence and moment conditions on the data that in the absence of a change point a detector based on integrating such a process over the partial sample parameter is asymptotically distributed as the norm of a Gaussian process, which furnishes a consistent change point detection procedure. We further derive consistency and local asymptotic results for this detector in the presence of a change in the covariance function. The corresponding change point estimator based on such a process is also shown to be rate optimal for estimating an existing change point, and further is asymptotically distributed as the argument maximum of a Gaussian process under a local asymptotic framework. We study the detector and change point estimator in a small simulation study to detect changes in the covariance of functional autoregressive and generalized conditionally heteroscedastic processes, which demonstrate that the use of the weighted CUSUM statistics in this context generally improves performance over existing methods. These new statistics are demonstrated in an application to detecting changes in the volatility of high resolution intraday asset price curves derived from oil futures prices.

Keywords: functional data change point detection Bernoulli shift approximation of partial sums of functions
2010 MSC: Primary 62H12, Secondary 62F12

1. Introduction

Functional data analysis has emerged as a vibrant area of research in statistics over the past several decades, owing to the multitude of data now collected, often at a high resolution, over a continuum. Such data can be viewed as discrete observations from functional data objects taking values in a function space. In a number of examples of interest, functional data objects are obtained sequentially as functional time series. We refer the reader to Ramsey and Silverman (2002), Ferraty and Vieu (2006), Horváth and Kokoszka (2012), and Kokoszka and Reimherr (2017) for textbook length treatments of functional data analysis, and Cuevas (2014) for a survey of modern research topics. Seminal work on functional time series analysis is summarized in Bosq (2000), and further reviewed in Kokoszka (2012) and Hörmann and Kokoszka (2012).

In the setting of functional time series, one often encounters series of curves exhibiting nonstationarity that appears as "shocks" or structural changes in the data generating mechanism. A simple model for such data is a change point model in which various features of the series are allowed to change at unknown points over the observation period; see Horváth and Rice (2014) for a survey of change point analysis. Inference for change points in the level or mean function with functional data have been avidly studied recently. Berkes et al. (2009), Aue et al. (2009), Aston and Kirch (2012a), Fraiman et al. (2014) consider mean change point analysis of serially independent functional data, and many of these methods were extended to cover potential serial dependence in Sharipov et al. (2016), Aston and Kirch (2012b), and Aue et al. (2018). Bardsley et al. (2017) develops a test for changes in the mean of a functional time series evolving according to heteroscedastic functional factor model.

*Corresponding author. Email address: grice@uwaterloo.ca

Preprint submitted to Journal of Multivariate Analysis *March 29, 2021*

Many change points appear though as changes to the underlying variance or covariance structure, rather than level shifts in the mean. Change point analysis for the variance and covariance matrices of scalar or vector valued observations enjoys an enormous literature going back to at least Inclan and Tiao (1994) in the scalar case, and more recent references for change point analysis of the covariance matrix include Aue et al. (2009), Wied et al. (2012), Galeano and Wied (2014), Berens et al. (2015), and Wied (2017). The problem of conducting change point analysis for the covariance function or operator describing the second order behaviour of functional data has been comparatively less explored. Jarušková (2013) is apparently the first to consider the change point detection problem for the covariance operator of independent functional data, and their approach is based on an initial dimension reduction step using functional principal component analysis. Stoehr et al. (2020) generalized these detection methods as well as several other dimension reduction based approaches to allow for general forms of serial dependence. Aue et al. (2020) and Dette and Kutta (2019) consider change point inference under similar weak dependence conditions for the spectrum and eigenfunctions, respectively, of covariance operators. Evaluating potential changes in the covariance operator is of particular interest in functional data analysis, since the covariance operator is central to a wide range of dimension reduction approaches for functional data, as described in the prior surveys Goia and Vieu (2016) and Aneiros et al. (2019).

Most closely related to the present paper are Sharipov and Wendler (2020), who develop change point detection methods for the covariance operator based on norms of a suitably constructed functional cumulative sum (CUSUM) process under general weak dependence conditions. The recent preprint Jiao et al. (2020) considers change point detection as well as estimation using similar norm-based techniques and conditions to Sharipov and Wendler (2020), with applications to functional data objects derived from rat brain studies. A notable feature of each of these procedures is that they are based on norms of the standard CUSUM process. As is well known in general with change point analysis, detection and estimation can be improved when change points are located away from the middle of the sample by applying weights to the standard CUSUM process in order to make the "variance" of the process more comparable at each potential change point; see e.g. Horváth et al. (2020) for a detailed discussion. Although it is conceptually simple to apply such weights, the application in this context encounters some technical challenges in that suitable weighted approximations for the CUSUM process of random elements in general infinite dimensional Hilbert space are not available.

In this paper, we develop and study change point detection and estimation procedures for the covariance operator based on the norms of weighted functional CUSUM processes. In the absence of a change point we establish the asymptotic distribution of a change point detector based on integrating such processes across the partial sample parameter under general weak dependence conditions similar to those considered in Stoehr et al. (2020) and Sharipov and Wendler (2020), and we further derive consistency and local asymptotic results for this detector in the presence of a change in the covariance function. Additionally, we show that the natural change point estimator based on such processes is rate optimal for estimating an existing change point, and further is asymptotically distributed as the argument maximum of a Gaussian process under a local asymptotic framework. In place of suitable weighted approximations, we establish Hájek–Rényi style inequalities for the norms of partial sample estimates of the covariance function, which underpin these results. We study the detector and change point estimator in a small simulation study to detect changes in the covariance of functional autoregressive and generalized conditionally heteroscedastic processes, which demonstrate that the use of the weighted CUSUM statistics in this context generally improves performance over existing methods. These new statistics are demonstrated in an application to detecting changes in the volatility of high resolution intraday asset price curves derived from oil futures prices.

The rest of the paper is organized as follows. In Section 2 we introduce the change point model and assumptions, define the weighted CUSUM change point detector, and detail its asymptotic behaviour. Section 3 details some specific examples, functional linear processes, in which changes in the covariance function arise and can be quantified based on changes in the model parameters. We define and present the asymptotic properties of the change point estimator in Section 4. The results of the Monte Carlo simulation study and analysis of intraday asset price curves are contained in Sections 5 and 6, respectively. All technical details and some concluding remarks follow these sections.

2. Detecting changes in the covariance function

We consider real valued functional observations $X_1(t), \dots, X_N(t)$, $t \in [0,1]$. Here each functional observation X_i is treated as a stochastic process with sample path in $L^2([0,1])$, where $L^2([0,1])$ denotes the space of (measurable) real valued square integrable functions defined on $[0,1]$. These results could be cast for functional observations in general, separable Hilbert space, but in order to maintain similarity to the literature introduced above, and in view of the data analysis that we consider below, we focus on this example. All random functions considered are assumed to be defined over a common probability space $(\Omega, \mathcal{F}, P)$. We assume that these observations are generated by the model, for $t \in [0,1]$,

$$X_i(t) = \begin{cases} \mu(t) + \epsilon_i(t), & 1 \le i \le k^* \\ \mu(t) + \epsilon_{i,A}(t), & k^* + 1 \le i \le N, \end{cases} \tag{1}$$

where $\mu(t)$ is the common mean, $E\epsilon_i(t) = E\epsilon_{i,A}(t) = 0$ for all $t \in [0,1]$, and $k^* \in \{1, ..., N\}$ denotes the unknown potential time of change. Implicitly under this model we make the simplifying assumption that the mean is constant with respect to i (although the mean function $\mu(t)$ may not be constant). This could be a reasonable assumption when considering for instance data constructed from financial asset returns, or in general suitably differenced or transformed time series data, as in our data example on intra-day asset price returns below. If this assumption is in question the mean might be estimated initially by applying a mean segmentation procedure as described for example in Aue et al. (2018). It is assumed that the covariance function

$$\mathfrak{C}(t,s) = E\epsilon_i(t)\epsilon_i(s), \quad 1 \le i \le k^*,$$

may change after k^*. Let

$$\mathfrak{C}_A(t,s) = E\epsilon_{k,A}(t)\epsilon_{k,A}(s), \quad k^* \le k \le N$$

denote the covariance function after the change. We use the notation $\mathfrak{C}_\Delta(t,s) = \mathfrak{C}(t,s) - \mathfrak{C}_A(t,s)$ to denote the difference between the covariance kernels before and after the change point. We are then interested in testing the null hypothesis that there is no change in the covariance function, versus the alternative of the presence of such a change:

$$H_0 : \|\mathfrak{C}_\Delta\| = 0, \text{ versus } \quad H_A : \|\mathfrak{C}_\Delta\| > 0.$$

We assume that the error functions might be serially dependent in that they evolve as decomposable Bernoulli shifts. Let $\|\cdot\|_2$ denote the norm induced by the inner product $\langle\cdot,\cdot\rangle_2$ in the Hilbert space $L^2([0,1]^p)$, the dimension p being clear based on the input function.

Assumption 1. (i) $\{\epsilon_i, -\infty < i < \infty\}$ and $\{\epsilon_{i,A}, -\infty < i < \infty\}$ form Bernoulli shifts, i.e. $\epsilon_i = g(\eta_i, \eta_{i-1}, \ldots)$ and $\epsilon_{i,A} = g_N(\eta_i, \eta_{i-1}, \ldots)$ for some non random measurable functions g,g_N: $\mathcal{S}^\infty \to L^2([0,1])$, where $\{\eta_j, -\infty < j < \infty\}$ are independent and identically distributed elements of a measurable space $\mathcal{S}$. (2) For each i, ϵ_i and $\epsilon_{i,A}$ are jointly measurable on $\Omega \times [0,1]$. (3) $E\epsilon_i(t) = 0$, $E\epsilon_{i,A}(t) = 0$, $E\|\epsilon_i\|_2^\nu < \infty$, $E\|\epsilon_{i,A}\|_2^\nu < \infty$ with some $\nu > 4$, and for some $\alpha > 2$ and a positive constant c,

$$\left(E\|\epsilon_0 - \epsilon_{0,\ell}\|_2^\nu\right)^{1/\nu} \le c k^{-\alpha}, \text{ and } \left(E\|\epsilon_{0,A} - \epsilon_{0,\ell,A}\|_2^\nu\right)^{1/\nu} \le c k^{-\alpha},$$

where $\epsilon_{i,\ell} = g(\eta_i, \ldots, \eta_{i-\ell+1}, \eta^*_{i-\ell}, \eta^*_{i-\ell-1}, \ldots)$, $\epsilon_{i,\ell,A} = g_N(\eta_i, \ldots, \eta_{i-\ell+1}, \eta^*_{i-\ell}, \eta^*_{i-\ell-1}, \ldots)$, and $\{\eta^*_\ell, -\infty < \ell < \infty\}$ are independent and identically distributed as η_0, and independent of $\{\eta_\ell, -\infty < \ell < \infty\}$.

Aue et al. (2014) and Kokoszka and Reimherr (2017) provide numerous examples where Assumption 1 holds, which include the stationary solutions to most functional time series models under standard regularity conditions.

In order to test H_0 it is natural to consider functionals of the CUSUM process of the sample covariance function:

$$Z_N(u,t,s) = N^{-1/2}\left(\sum_{i=1}^{\lfloor Nu \rfloor}[X_i(t) - \bar{X}_N(t)][X_i(s) - \bar{X}_N(s)] - \frac{\lfloor Nu \rfloor}{N}\sum_{i=1}^{N}[X_i(t) - \bar{X}_N(t)][X_i(s) - \bar{X}_N(s)]\right),$$

where $\bar{X}_N(t) = (1/N)\sum_{i=1}^N X_i(t)$. is the sample mean. The asymptotic behaviour of the process $Z_N(u,t,s)$ depends on the long run covariance function

$$\mathfrak{D}(t,t',s,s') = \sum_{\ell=-\infty}^{\infty} E[(\epsilon_0(t)\epsilon_0(s) - \mathfrak{C}(t,s))(\epsilon_\ell(t')\epsilon_\ell(s') - \mathfrak{C}(t',s'))]. \tag{2}$$

It follows from Assumption 1 that the infinite sum defining $\mathfrak{D}(t,t',s,s')$ is absolutely convergent in $L^2([0,1]^4)$. Throughout this paper we use $\int$ for $\int_0^1$.

Theorem 1. If H_0 and Assumption 1 hold, then we can define a sequence of Gaussian processes $\{\Gamma_N(u,t,s), 0 \le u,t,s \le 1\}$, $N \ge 1$, such that

$$\sup_{0<u<1} \iint (Z_N(u,t,s) - \Gamma_N(u,t,s))^2 \, dtds = o_P(1),$$

with $E\Gamma_N(u,t,s) = 0$ and $E\Gamma_N(u,t,s)\Gamma_N(u',t',s') = (\min(u,u') - uu')\mathfrak{D}(t,t',s,s')$.

The test statistics considered in Sharipov and Wendler (2020) and Jiao et al. (2020) coincide with

$$T_N = \iiint Z_N^2(u,t,s)dudtds, \text{ and } S_N = \sup_{0\le u\le 1} \int\int Z_N^2(u,t,s)dudtds.$$

It has been observed that reweighting the CUSUM process can increase the power of CUSUM based change point detectors, and improve the efficiency of corresponding change point estimators in the presence of multiple change points or change points occurring near the end points of the sample. We use the weight function $w_\kappa(u) = [u(1-u)]^\kappa$, where $\kappa \in [0, 1/2]$. We then have the following approximation similar to Theorem 1 with Z_N reweighted by w_κ.

Theorem 2. If H_0, Assumption 1 hold and $0 \le \kappa < 1/2$, then with the Gaussian processes $\{\Gamma_N(u,t,s), 0 \le u,t,s \le 1\}$ of Theorem 1 we have

$$\sup_{1/(N+1)<u<1-1/(N+1)} \frac{1}{w_\kappa^2(u)} \iint (Z_N(u,t,s) - \Gamma_N(u,t,s))^2 \, dtds = o_P(1).$$

This result may be used to evaluate the asymptotic properties of weighted functionals of Z_N. Before doing so, we first investigate the behaviour of functionals of the weighted CUSUM process Z_N under the alternative. We write the time of change in the form of $k^* = \lfloor N\theta_N \rfloor$, $\theta_N \in (0,1)$. We allow that the size of the change described by $\mathfrak{C}_\Delta(t,s)$ depends on N but it cannot converge to infinity.

Theorem 3. We assume that H_A, Assumption 1 are satisfied, $0 \le \kappa < 1/2$ and

$$\limsup_{N\to\infty} \|\mathfrak{C}_\Delta\|_2 < \infty.$$

(i) If $N(\theta_N(1-\theta_N))^{2-2\kappa}\|\mathfrak{C}_\Delta\|_2^2 \ \to \ \infty$, then

$$\frac{1}{N(\theta_N(1-\theta_N))^{2-2\kappa}\|\mathfrak{C}_\Delta\|_2^2} \sup_{0<u<1} \frac{1}{w_\kappa^2(u)} \iint Z_N^2(u,t,s)dtds \ \xrightarrow{P} \ 1. \tag{3}$$

(ii) If $N(\theta(1-\theta_N))^{3-2\kappa}\|\mathfrak{C}_\Delta\|_2^2 \ \to \ \infty$, then

$$\frac{1}{N(\theta_N(1-\theta_N))^{3-2\kappa}\|\mathfrak{C}_\Delta\|_2^2} \iiint \frac{1}{w_\kappa^2(u)} Z_N^2(u,t,s)dtdsdu \ \xrightarrow{P} \ c_0, \tag{4}$$

for some $c_0 > 0$.

According to Theorems 1 and 3, a consistent test of H_0 versus H_A is obtained by rejecting H_0 if

$$T_N(\kappa) = \iiint \frac{Z_N^2(u,t,s)}{w_\kappa^2(u)} dudtds$$

is large. Now we discuss how to obtain asymptotically valid critical values for $T_N(\kappa)$ under H_0 based on Theorem 2. Let $\{\Gamma(u,t,s), 0 \le u,t,s \le 1\}$ be a Gaussian process with zero mean and $E\Gamma(u,t,s)\Gamma(u',t',s') = (\min(u,u') - uu')\mathfrak{D}(t,t',s,s')$, where $\mathfrak{D}$ is defined in (2). We note that the distribution of the process Γ coincides for each N with Γ_N defined in Theorems 1 and 2. Noticing that the covariance function of Γ separates into a product of $\mathfrak{D}$ and the function $\min(u,u') - uu'$, which is the covariance function of a Brownian bridge, we obtain using the Karhunen-Loéve expansion of Γ that

$$\iiint \Gamma^2(u,t,s)dudtds \overset{\mathcal{D}}{=} \sum_{k,\ell=1}^{\infty} \frac{\lambda_\ell}{(\pi k)^2} \mathcal{N}_{k,\ell}^2,$$

where $\{\mathcal{N}_{k,\ell}, 1 \leq k, \ell < \infty\}$ are independent, identically distributed standard normal random variables and $\lambda_1 \geq \lambda_2 \geq \dots$ satisfy

$$\lambda_i \phi_i(t,s) = \int\int \mathfrak{D}(t,t',s,s')\phi_i(t',s')dt'ds', \quad 1 \leq i < \infty, \;\; \langle \phi_i, \phi_j \rangle_2 = \mathbb{1}\{i=j\}. \tag{5}$$

The (long run) covariance $\mathfrak{D}$ is of course generally unknown, but we can estimate it from the sample. We define the residuals

$$\hat{\epsilon}_i(t) = X_i(t) - \bar{X}_N(t), \quad 1 \leq i \leq N. \tag{6}$$

Let

$$z_i(t,s) = \hat{\epsilon}_i(t)\hat{\epsilon}_i(s), \quad 1 \leq i \leq N, \text{ and } \bar{z}_N(t,s) = \frac{1}{N}\sum_{i=1}^{N} z_i(t,s). \tag{7}$$

We suggest using the long run kernel covariance estimator

$$\hat{\mathfrak{D}}_N(t,s,t's') = \sum_{\ell=-(N-1)}^{N-1} K\left(\frac{\ell}{h}\right)\hat{\gamma}_\ell(t,s,t',s') \tag{8}$$

where K is symmetric about zero Lipschitz continuous window function with bounded support such that $K(0) = 1$, $h = h(N)$ is the window (smoothing parameter) satisfying $1/h + N/h \to 0$ as $N \to \infty$, and

$$\hat{\gamma}_\ell(t,s,t',s') = \begin{cases} \displaystyle\sum_{i=1}^{N-\ell}(z_i(t,s) - \bar{z}_N(t,s))(z_{i+\ell}(t',s') - \bar{z}_N(t',s')), & \ell \geq 0 \\ \displaystyle\frac{1}{N-|\ell|}\sum_{i=-(\ell-1)}^{N}(z_i(t,s) - \bar{z}_N(t,s))(z_{i+\ell}(t',s') - \bar{z}_N(t',s')), & \ell < 0. \end{cases}$$

If $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \dots$ denote the empirical eigenvalues of $\hat{\mathfrak{D}}_N$, i.e.

$$\hat{\lambda}_i\hat{\phi}_i(t,s) = \int\int \hat{\mathfrak{D}}_N(t,t',s,s')\hat{\phi}_i(t',s')dt'ds', \quad 1 \leq i \leq N, \;\; \langle \hat{\phi}_i, \hat{\phi}_j \rangle_2 = \mathbb{1}\{i=j\}. \tag{9}$$

We use $c_N(\alpha)$ the solution of the equation with a suitably large constant d,

$$P\left\{\sum_{\ell=1}^{d}\sum_{k=1}^{\infty}\frac{\hat{\lambda}_\ell}{(\pi k)^2}\mathcal{N}_{k,\ell}^2 > c_N(\alpha)\right\} = \alpha.$$

to approximate the $1-\alpha$ critical value for $T_N(0)$ of H_0. It follows similarly as in Berkes et al. (2016) that under these conditions and Assumption 1

$$\|\mathfrak{D} - \hat{\mathfrak{D}}_N\|_2 \xrightarrow{P} 0, \tag{10}$$

if H_0 and Assumption 1 hold. We obtain from (10) using Lemma 2.2 in Horváth and Kokoszka (2012) that for any $d \geq 1$,

$$\max_{1\leq i\leq d}|\lambda_i - \hat{\lambda}_i| \xrightarrow{P} 0.$$

Theorem 1 now implies that under H_0

$$\lim_{d\to\infty}\lim_{N\to\infty} P\{T_N(0) > c_N(\alpha)\} = \alpha.$$

Under the alternative H_A might not hold. Let

$$\Delta_N(t,s,t',s') = \mathfrak{C}_\Delta(t,s)\mathfrak{C}_\Delta(t',s').$$

Under the alternative, it may be established as in the proof of Theorem 3.2 of Horváth et al. (2014) that

$$\iiiint \left(\frac{1}{h}\hat{\mathfrak{D}}_N(t,s,t',s') - \theta_N(1-\theta_N)\Delta_N(t,s,t',s')\int_{-c}^{c}K(u)du\right)^2 dtdsdt'ds' = o_P(1). \tag{11}$$

Using again Lemma 2.2 in Horváth and Kokoszka (2012) we now conclude that under alternative

$$\hat{\lambda}_1 = O_P\left(h\theta_N(1-\theta_N)\|\mathfrak{C}_\Delta\|_2^2\right).$$

If

$$N(\theta(1-\theta_N))^2\|\mathfrak{C}_\Delta\|_2^2 \;\to\; \infty, \text{ and } \frac{N}{h}\theta_N(1-\theta_N) \to \infty,$$

hold, then under the alternative

$$\lim_{d\to\infty}\lim_{N\to\infty} P\{T_N(0) > c_N(\alpha)\} = 1$$

for any $d \geq 1$. Hence the test based on $T_N(0)$ has approximately correct, for large d, asymptotic size, and is consistent under H_A. If the statistic $T_N(\kappa)$, $\kappa \in (0, 1/2)$, is used, then the above formula may be modified by defining the eigenvalues $\zeta_{1,\kappa} \geq \zeta_{2,\kappa} \geq \cdots \geq 0$ satisfying

$$\zeta_{i,\kappa}\psi_{i,\kappa}(u) = \int \frac{\min(u,v) - uv}{w_\kappa(u)w_\kappa(v)}\psi_{i,\kappa}(v)dv, \quad 1 \leq i < \infty, \;\; \langle\psi_{i,\kappa}, \psi_{j,\kappa}\rangle_2 = \mathbb{1}\{i = j\}.$$

According to the Karhunen–Loéve representation,

$$\iiint \frac{\Gamma^2(u,t,s)}{w_\kappa^2(u)}dudtds = \sum_{k,\ell=1}^{\infty} \lambda_\ell \zeta_{j,\kappa} \mathcal{N}_{j,\ell}^2, \tag{12}$$

where $\{\mathcal{N}_{j,\ell}, 1 \leq j, \ell < \infty\}$ are independent, identically distributed standard normal random variables and $\lambda_1 \geq \lambda_2 \geq \ldots$ are defined in (5). We can use the previous arguments to estimate λ_ℓ and obtain asymptotic critical values for $T_N(\kappa)$ by estimating the distribution on the right hand side of (12) using simulation.

With regards to the choice of d, there are a number of sensible methods to do this, including using the total variation explained criterion common in functional principal component analysis. Noting that the quantiles of the random variable on the right hand side of (12) are increasing as a function of d, below we use the conservative approach of selecting d as the number of positive eigenvalues estimated according to (9).

3. Functional Time Series Models with Covariance Change Points

As is the case of scalar and vector valued observations, the estimator for the time of change has nice asymptotic properties if the size of the change converges to zero when the sample sizes goes to ∞. Hence, before we start discussing the estimation of the time of the change, we consider some examples in that describe such small changes in the covariance function. We show that changes in the parameters of popular functional time series models lead to changes in the covariance function of the observations. These examples were also considered in the simulation study presented in Section 5.

Example 1. We consider functional AR(1) models. We assume that ϵ_i, $i \in \mathbb{Z}$ is the stationary solution of

$$\epsilon_i(t) = \int \mathcal{K}(t,s)\epsilon_{i-1}(s)ds + \eta_i(t), \tag{13}$$

where $\mathcal{K}$ is an autoregressive kernel function. It is well known (see for example, Bosq, 2000) that the equation in (13) has a unique, non anticipative stationary solution if

$$\|\mathcal{K}\|_2 < 1, \tag{14}$$

and

$$\begin{aligned}&\{\eta_i(t), 0 \leq t \leq 1, -\infty < i < \infty\} \quad \text{are independent and identically distributed}\\ &\text{random functions with } \; E\eta_0(t) = 0 \; \text{ and } \; E\|\eta_0(t)\|_2^\nu < \infty \quad \text{with some} \quad \nu > 4.\end{aligned} \tag{15}$$

In this example, comparing to (1) we assume the random functions $\epsilon_{i,A}$, $i \in \mathbb{Z}$ are the unique, non anticipative stationary solution of

$$\epsilon_{i,A}(t) = \int \mathcal{K}_N(t,s)\epsilon_{i-1,A}(s)ds + \eta_i(t), \tag{16}$$

with

$$\mathcal{K}_N(t,s) = \mathcal{K}(t,s) + a_N \mathscr{k}(t,s),$$

where, in addition to (13) (14) and (15), $a_N > 0$ is small enough that

$$\|\mathcal{K}\|_2 + a_N \|\mathscr{k}\|_2 < 1.$$

Let

$$\mathcal{K}^{(\ell)}(x_1, x_{n+1}) = \int \cdots \int \mathcal{K}(x_1,x_2)\mathcal{K}(x_2,x_3)\ldots\mathcal{K}(x_n,x_{n+1}) \prod_{i=2}^{n} dx_i$$

and the associated operator be

$$\mathcal{K}^{(\ell)}[f](t) = \int \mathcal{K}^{(\ell)}(t,s)f(s)ds, \quad \ell \geq 1 \quad \text{and} \quad \mathcal{K}^{(0)}[f](t) = f(t).$$

It may be shown under these assumptions that the stationary solutions satisfy

$$\epsilon_i(t) = \sum_{\ell=0}^{\infty} \mathcal{K}^{(\ell)}[\eta_i](t), \text{ and similarly } \epsilon_{i,A}(t) = \sum_{\ell=0}^{\infty} \mathcal{K}_N^{(\ell)}[\eta_i](t).$$

It is clear that Assumption 1 holds for these variables due to the explicit forms of the solutions of (13) and (16). Let

$$\mathcal{L}^{(\ell)}(x_1, x_{\ell+1}) = \sum_{k=1}^{\ell} \int \cdots \int \prod_{i=1}^{k-1} \mathcal{K}(x_i, x_{i+1}) \mathscr{k}(x_k, x_{k+1}) \prod_{j=k+1}^{\ell} \mathcal{K}(x_j, x_{j+1}) \prod_{m=2}^{\ell} dx_m,$$

with $\prod_{i\in\emptyset} = 1$. The corresponding operator is

$$\mathcal{L}^{(\ell)}[f](t) = \int \mathcal{L}^{(\ell)}(t,s)f(s)ds, \quad \ell \geq 1 \quad \text{and} \quad \mathcal{L}^{(0)}[f](t) = f(t).$$

Elementary arguments give

$$\epsilon_{i,A}(t) - \epsilon_i(t) = a_N\delta_i(t) + a_N^2 R_{i,N}(t), \text{ with } \delta_i(t) = \sum_{\ell=0}^{\infty} \mathcal{L}^{(\ell)}[\eta_{i-\ell}](t), \text{ and } \limsup_{N\to\infty} E\|R_{0,N}\|_2^{\nu} < \infty.$$

From this it follows by simple calculation that

$$\|\mathfrak{C}_\Delta - a_N(E\epsilon_0 \otimes \delta_0 + E\delta_0 \otimes \epsilon_0)\|_2 = O(a_N^2),$$

where here for functions $f, g \in L^2([0,1])$, $f \otimes g$ is the function in $L^2([0,1]^2)$ defined by $f \otimes g(t,s) = f(t)g(s)$. In this example therefore a small change on the order of a_N to the functional autoregressive operator defined by $\mathcal{K}$ induces a change of the same magnitude in the covariance kernels. This example may also be extended to general functional linear processes.

The second example we consider is concerned with nonlinear functional time series processes.

Example 2. Following Aue et al. (2017), Cerovecki et al. (2019), and Künert (2020) we define the functional GARCH(1,1) (FGARCH(1,1)) process

$$\epsilon_i(t) = \sigma_i(t)\eta_i(t),$$

where $E\eta_i(t) = 0$, $E\eta_i^2(t) = 1$, and

$$\sigma_i^2(t) = \omega(t) + \int \alpha(t,s)\epsilon_{i-1}^2(s)ds + \int \beta(t,s)\sigma_{i-1}^2(s)ds, \tag{17}$$

and the non-negative parameter functions $\omega(t)$, α and β satisfy the regularity conditions of Theorem 1 of Aue et al. (2017), which imply that a stationary solution ϵ_i satisfying (17) exists in the function space $C[0,1]$ of continuous functions defined on the unit interval. One of these conditions in particular is that $\inf_{0\leq t\leq 1} \omega(t) > 0$. A change in the variance of the process may be modelled by changes in these parameter functions. For example, a "level shift" in the pointwise variance of the functional observations is induced in (1) by setting $\epsilon_{i,A}(t) = \sigma_{i,A}(t)\eta_i(t)$, with

$$\sigma_{i,A}^2(t) = \omega(t) + a_N c(t) + \int \alpha(t,s)\epsilon_{i-1}^2(s)ds + \int \beta(t,s)\sigma_{i-1}^2(t,s)ds,$$

and the function c is taken to satisfy $\inf_{0\le t\le 1} c(t) > 0$. Since the stationary solution σ_0 of (17) is independent of η_0, we get

$$\mathfrak{C}_\Delta(t,s) = \left(E\sigma_{i,A}(t)\sigma_{i,A}(s) - E\sigma_i(t)\sigma_i(s)\right) E\eta_0(t)\eta_0(s),$$

and by the mean-value theorem

$$E\left\|\sigma_{i,A} - \sigma_i - \frac{a_N c}{2\sigma_i}\right\|_2^2 = O(a_N^4).$$

From this it may be shown that

$$\int\int \left(\mathfrak{C}_\Delta(t,s) - \frac{1}{2}\left\{E\left(\frac{\sigma_0(s)}{\sigma_0(t)}\right)c(t) + E\left(\frac{\sigma_0(t)}{\sigma_0(s)}\right)c(s)\right\}a_N\right)^2 dtds = O(a_N^4).$$

Therefore a change of magnitude a_N to the level of the conditional variance process induces a change of the same magnitude in the covariance functions. A similar change arises when any of the other parameter functions are changed, and it may be shown as above that if

$$\sigma_{A,i}^2(t) = \omega + \int (\alpha(t,s) + a_N\delta_1(t,s))\epsilon_{i-1}^2(s)ds + \int (\beta(t,s) + a_N\delta_2(t,s))\sigma_{i-1}^2(s)ds,$$

then as a_N tends to zero there exists a nonzero kernel $\mathcal{A}$ so that $\|\mathfrak{C}_\Delta - a_N\mathcal{A}\|_2 = O(a_N^2)$.

4. Estimation of k^*

The weighted CUSUM process can also be used to define an estimator for the time of change. We let

$$\hat{k}_N(\kappa) = \min\left\{k :\ \left(\frac{N}{k(N-k)}\right)^\kappa \int\int Z_N^2\left(\frac{k}{N},t,s\right)dtds = \max_{1\le j<N}\left(\frac{N}{j(N-j)}\right)^\kappa \int\int Z_N^2\left(\frac{j}{N},t,s\right)dtds\right\}.$$

We investigate the asymptotic behaviour of this estimator when the time of change is proportional to the sample size:

Assumption 2. $k^* = \lfloor N\theta \rfloor$ with some $0 < \theta < 1$.

Again here we characterize the size of the change in the covariance by $\mathfrak{C}_\Delta(t,s) = \mathfrak{C}(t,s) - \mathfrak{C}_A(t,s)$. We only consider the case of a local alternative in which the size of the change is shrinking as a function of N. In this case the limit distribution of the time of change has a simple, easily computable form depending on only a small number of nuisance parameters.

Assumption 3. $\|\mathfrak{C}_\Delta\|_2 \to 0$ and $N\|\mathfrak{C}_\Delta\|_2^2 \to \infty$.

We also assume that the change between the distributions of the ϵ_i and $\epsilon_{i,A}$ vanishes as N increases.

Assumption 4. With some $\nu > 4$

$$E\left\|\epsilon_i - \epsilon_{i,A}\right\|_2^\nu \to 0, \quad \text{as} \quad N \to \infty.$$

The examples detailed in Section 3 satisfy Assumptions 3 and 4 when the parameter a_N decreases at a suitable rate. Under Assumption 4 the covariance function $\mathfrak{C}_A$ is close to $\mathfrak{C}$, if the sample size N is large. We also assume that the standardized difference has a quantifiable limit:

Assumption 5. There is $\mathfrak{C}^* \in L^2([0,1]\times[0,1])$ such that

$$\iint \left(\frac{\mathfrak{C}_\Delta(t,s)}{\|\mathfrak{C}_\Delta\|_2} - \mathfrak{C}^*(t,s) \right)^2 dtds = o(1).$$

Let

$$\tau^2 = \sum_{\ell=-\infty}^{\infty} \mathrm{cov}\left(\iint \epsilon_0(t,s)\mathfrak{C}^*(t,s)dtds, \iint \epsilon_\ell(t,s)\mathfrak{C}^*(t,s) \right) dtds, \tag{18}$$

which appears in the normalization of the difference between $\hat{k}_N$ and k^*. In order to define the limit distribution of $\hat{k}_N - k^*$, we also define

$$m_\kappa(t) = \begin{cases} (1-\kappa)(1-\theta) + \kappa\theta, & \text{if} \quad t < 0 \\ 0, \quad \text{if} \quad t = 0 \\ (1-\kappa)\theta + \kappa(1-\theta), & \text{if} \quad t > 0 \end{cases} \tag{19}$$

and

$$W(t) = \begin{cases} W_1(-t), & \text{if} \quad t < 0 \\ W_2(t), & \text{if} \quad t \geq 0 \end{cases} \tag{20}$$

where $\{W_1(t), t \geq 1\}$ and $\{W_2(t), t \geq 1\}$ are independent Wiener processes. The process $\{W(t), -\infty < t < \infty\}$ is sometimes called a two sided Wiener process. There is an almost surely unique random variable $\xi(\kappa)$ defined as

$$\xi(\kappa) = \mathrm{argmax}_{t\in\mathbb{R}} \{W(t) - |t| m_\kappa(t)\}. \tag{21}$$

We note that the random variable $\xi(0)$ appeared in Dümbgen (1991) and $\xi(\kappa),\ 0 \leq \kappa \leq 1/2$ in Antoch et al. (1995, 1997).

Theorem 4. We assume that H_A and Assumptions 1–5 are satisfied.
(i) If $0 \leq \kappa < 1/2$, then

$$\frac{\|\mathfrak{C}_\Delta\|_2^2}{\tau^2} \left(\hat{k}_N(\kappa) - k^* \right) \overset{\mathcal{D}}{\to} \xi(\kappa).$$

(ii) If in addition

$$N^{1/2} \|\mathfrak{C}_\Delta\|_2 (\log N)^{-2/\nu} \to \infty$$

holds, then

$$\frac{\|\mathfrak{C}_\Delta\|_2^2}{\tau^2} \left(\hat{k}_N(1/2) - k^* \right) \overset{\mathcal{D}}{\to} \xi(1/2),$$

where ν is from Assumption 1 and $\xi(\kappa)$ is defined in (21).

The asymptotic distribution of $\hat{k}_N(0) - k^*$ was obtained by Jiao et al. (2020) when $\|\mathfrak{C}_\Delta\|_2$ is a constant. Theorem 4 can be used to produce asymptotically conservative confidence intervals for k^*. Although we do not study the empirical properties of such intervals here, producing them requires the estimation of τ^2 in (18), which we describe. τ^2 may be viewed as the long run variance of the stationary sequence

$$e_j = \iint \epsilon_j(t)\epsilon_j(s)\mathfrak{C}^*(t,s)dtds.$$

The series e_j is not observed, but can be estimated from the sample by replacing $\epsilon_j(t)\epsilon_j(s)$ with the residuals defined in (6). We note that

$$\tau^2 = \int \cdots \int \mathfrak{C}^*(t,s)\mathfrak{D}(t,t',s,s')\mathfrak{C}^*(t',s')dtdt'dsds'.$$

The estimator for the covariance function before the change may be defined as,

$$\hat{\mathfrak{C}}_{N,1}(t,s) = \frac{1}{\hat{k}_N} \sum_{i=1}^{\hat{k}_N} (X_i(t) - \bar{X}_N(t))(X_i(s) - \bar{X}_N(s))$$

and similarly after the change as

$$\hat{\mathfrak{C}}_{N,2}(t,s) = \frac{1}{N-\hat{k}_N}\sum_{i=\hat{k}_N+1}^{N}(X_i(t)-\bar{X}_N(t))(X_i(s)-\bar{X}_N(s)).$$

$\mathfrak{C}^*(t,s)$ may be estimated with

$$\hat{\mathfrak{C}}_N^*(t,s) = \frac{\hat{\mathfrak{C}}_{N,1}(t,s)-\hat{\mathfrak{C}}_{N,2}(t,s)}{\|\hat{\mathfrak{C}}_{N,1}-\hat{\mathfrak{C}}_{N,2}\|_2}.$$

We already discussed the estimation of $\mathfrak{D}$ in (8), but this must be modified in the presence of a change point. Let

$$\bar{\mathfrak{D}}_N(t,s,t's') = \sum_{\ell=-(N-1)}^{N-1} K\left(\frac{\ell}{h}\right)\bar{\gamma}_\ell(t,s,t',s')$$

where again K is window function, and $h = h(N)$ is the window (smoothing parameter) and

$$\bar{\gamma}_\ell(t,s,t',s') = \begin{cases} \dfrac{1}{N-\ell}\displaystyle\sum_{i=1}^{N-\ell} v_i(t,s)v_i(t',s'), & \ell \ge 0 \\[2ex] \dfrac{1}{N-|\ell|}\displaystyle\sum_{i=-(\ell-1)}^{N} v_i(t,s)v_i(t',s'), & \ell < 0. \end{cases}$$

with

$$v_i(t,s) = \begin{cases} \hat{\epsilon}_i(t)\hat{\epsilon}_i(s) - \hat{\mathfrak{C}}_{N,1}(t,s), & 1 \le i \le \hat{k}_N, \\[1ex] \hat{\epsilon}_i(t)\hat{\epsilon}_i(s) - \hat{\mathfrak{C}}_{N,2}(t,s), & \hat{k}_N + 1 \le i \le N. \end{cases}$$

Similarly to (10) one can show that $\|\bar{\mathfrak{D}} - \mathfrak{D}\|_2 \xrightarrow{P} 0$. Combining Theorem 4 and the ergodic theorem we conclude $\|\mathfrak{C} - \hat{\mathfrak{C}}_{N,1}\|_2 \xrightarrow{P} 0$ and $\|\mathfrak{C}_A - \hat{\mathfrak{C}}_{N,2}\|_2 \xrightarrow{P} 0$. Hence $\hat{\tau}_N^2 \xrightarrow{P} \tau^2$, where

$$\hat{\tau}^2 = \int\cdots\int \hat{\mathfrak{C}}_N^*(t,s)\bar{\mathfrak{D}}_N(t,t',s,s')\hat{\mathfrak{C}}_N^*(t',s')dtdt'dsds'.$$

Remark 1. Under H_0 and the conditions of Theorem 2, it may be shown using the argmax continuous mapping theorem as in Kim and Pollard (1990), that for $0 \le \kappa < 1/2$ the break fraction $\hat{k}_N(\kappa)/N$ satisfies

$$\frac{\hat{k}_N(\kappa)}{N} \xrightarrow{\mathcal{D}} \arg\sup_{0<u<1} \frac{1}{w_\kappa^2(u)}\int\int \Gamma^2(u,t,s)dtds.$$

The distribution on the right hand side has mode at $1/2$ for each value of κ, but becomes more uniform over the unit interval as κ approaches $1/2$.

5. Simulation Results

In this section we present the results of a Monte Carlo simulation study that we used to assess the empirical size and power of the proposed change point tests, and evaluate the advantages and drawbacks of using weighted CUSUM test statistics compared to existing methods. We generated data from the functional AR(1) and GARCH(1,1) change point models defined in Examples 1 and 2. In both data generating processes (DGPs), we take the innovation terms to follow a time-homogeneous Ornstein–Uhlenbeck process,

$$\eta_i(t) = e^{-t/2}W_i(e^t), \quad t \in [0,1],$$

where $\{W_i(t), t \ge 0\}$ are independent and identically distributed standard Brownian motions. Each such curve is evaluated on a grid of equally space points on $[0,1]$ of size 50.

Table 1: Empirical size based on 1000 independent simulations using the FAR and FGARCH data generating processes, based on the statistics T_N, $T_N(1/4)$, and $T_N(2/5)$.

			T_N			$T_N(1/4)$			$T_N(2/5)$		
	Nominal Level	N=	100	250	500	100	250	500	100	250	500
	90%		0.10	0.12	0.11	0.13	0.12	0.11	0.11	0.12	0.11
FAR	95%		0.06	0.05	0.05	0.05	0.06	0.06	0.05	0.06	0.05
	99%		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	90%		0.14	0.14	0.13	0.13	0.14	0.13	0.14	0.15	0.14
FGARCH	95%		0.07	0.07	0.06	0.07	0.07	0.07	0.07	0.07	0.07
	99%		0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.02

We take the kernel $\mathcal{K}$ in Example 1 to be $\mathcal{K}(t,s) = 12t(1-t)s(1-s)$, which is changed after the change point k^* to $\mathcal{K}(t,s) + a_j t(1-t)s(1-s)$, where $a_0 = 0$, giving a setting under H_0, and $a_1 = 6$, $a_2 = 9$, $a_3 = 12$, giving settings under H_A.

We set the parameter functions α, β, and ω in Example 2 to $\alpha(t,s) = 2t(1-t)s(1-s)$, $\beta(t,s) = 10t(1-t)s(1-s)$ and $\omega(t) = 0.1t(1-t) + 10^{-9}$. We considered two types of change points in such FGARCH models: one of which we denote FGARCH(ω), where the function ω changes after k^* to $\omega(t) + a_j c(t) = [0.1 + a_j]t(1-t) + 10^{-9}$, with $a_0 = 0$, $a_1 = 0.1$, $a_2 = 0.2$, and $a_3 = 0.3$, and another that we denote FGARCH(α,β), in which the functions α and β change after k^* to $\alpha(t,s) + a_j\delta_1(t,s) = [2 + a_j]t(1-t)s(1-s)$ and $\beta(t,s) + a_j\delta_2(t,s) = [10 + a_j]t(1-t)s(1-s)$ where $a_0 = 0$, $a_1 = 2$, $a_2 = 4$, and $a_3 = 6$.

We considered three potential change point locations, which we call an early change, $k^* = \lfloor 0.2N \rfloor$, a central change, $k^* = \lfloor 0.5N \rfloor$, and a late change, $k^* = \lfloor 0.9N \rfloor$. We considered the sample sizes $N = 100, 250$, and 500. To each sample we applied the above described tests of H_0 using both the unweighted CUSUM test statistic $T_N = T_N(0)$, as well as the weighted statistics $T_N(1/4)$ and $T_N(2/5)$ to detect potential change points. In order to compute the critical and p-values of these tests, we followed the basic outline in Section 2, and estimated the long–run covariance kernel $\mathfrak{D}$ with the Bartlett window function and bandwidth $h = N^{1/5}$. Estimating the eigenvalues satisfying 9 was conducted by estimating the eigenvalues of a discretized 4-way tensor approximation of the kernel $\mathfrak{D}$, which can be computationally intensive if the resolution of this discretization is large. For example, letting J_u denote the points at which the functional data are observed, the intra–day return curves derived from 5–minute frequency equity price data that we consider below are observed on points in the set J_u, where J_u are equally spaced points on the unit interval with$|J_u| = 67$, where $|A|$ is the cardinality of the set A, and calculating the resulting 4-way tensor eigenvalues leads to a computation akin to estimating the leading eigenvalues of a square, dense matrix of approximate dimension 4500×4500. In order to reduce this computational burden, we employ standard Monte Carlo integration with a trapezoid rule to numerically approximate the integral in 9, where the grid $J_u \times J_u \times J_u \times J_u$ in the computation of the long–run covariance is exchanged with a grid $J_{mc} \times J_{mc} \times J_{mc} \times J_{mc}$, where J_{mc} are uniformly randomly sampled points from J_u, and $|J_{mc}| < |J_u|$. In the simulations and data analysis, we set $|J_{mc}| = 20$, and found that this leads to reliable results and relatively fast computation in the examples we considered.

Each simulation was repeated independently 1000 times, and the number of rejections of H_0 at several nominal levels are collected in Table 1, which shows results on the empirical size, and Figure 1, which display power curves when the change point is located at $k^* = \lfloor 0.2N \rfloor$, respectively.

In summary, we observed that the application of weights did not have much of an effect on the size of the tests for the examples considered, which were close to the nominal levels for all sample sizes and weighting functions considered. In terms of power, we saw in simulations that are reported in the supplement to this article that applying weights has little effect, good or bad, in terms of detecting changes in the covariance function that occur near the middle of the sample. However, in Figure 1 one may observe that the power of the test improves for detecting changes closer to the end point of the sample by incorporating larger weights. This improvement is observed most starkly for small sample sizes, in which in some cases as much as a 50% increase in power was observed when the weight function $w_{2/5}(u)$ was used compared to no weights, but this improvement was evident at each sample size considered.

Fig. 1: Power functions, which display the percentage of rejections of H_0 based on 1000 independent simulations as a function of the size of the change point, using the statistics T_N, $T_N(1/4)$, and $T_N(2/5)$. The nominal level set at 5% with $k^* = \lfloor 0.2N \rfloor$, and the data was generated according to the FAR and FGARCH data generating processes.

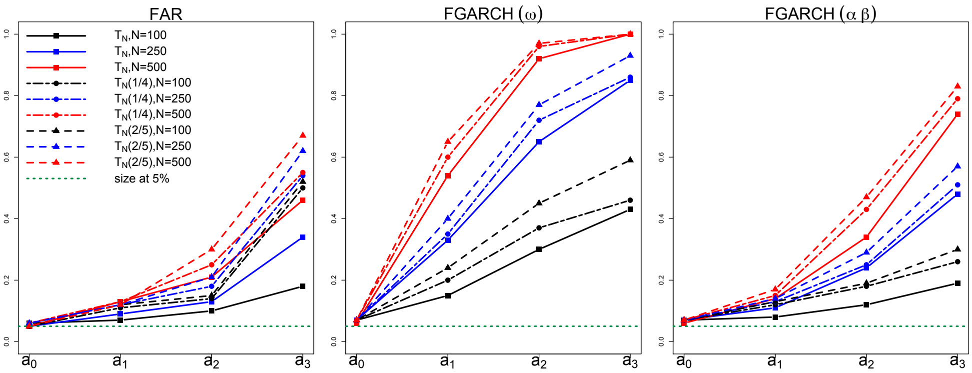
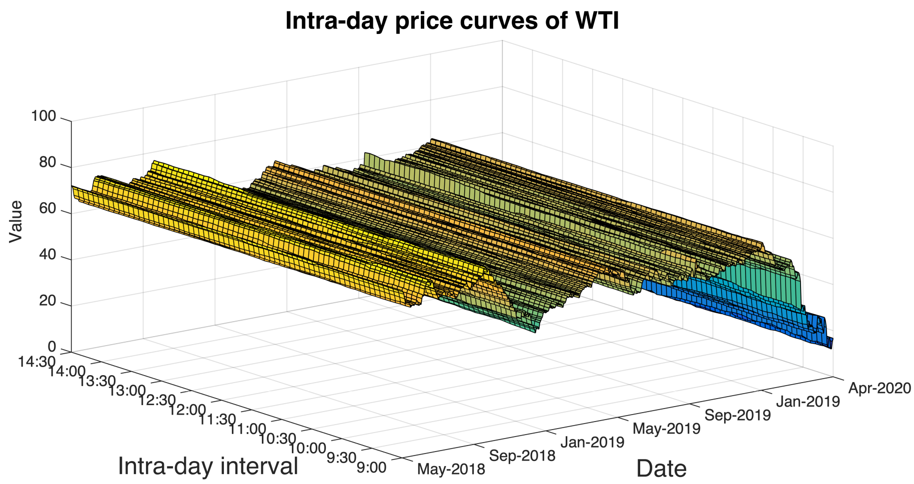


Fig. 2: Daily price curves from 12 May, 2018 to 30 Apr, 2020 of WTI commodity futures. Each curve is constructed from 5-minute frequency front-month prices of WTI futures, from 9:00 am to 2:30 pm, each trading day, which totals 502 days. There are 67 discrete observations of the price within each day, which are linearly interpolated to produce intra-day price curves.



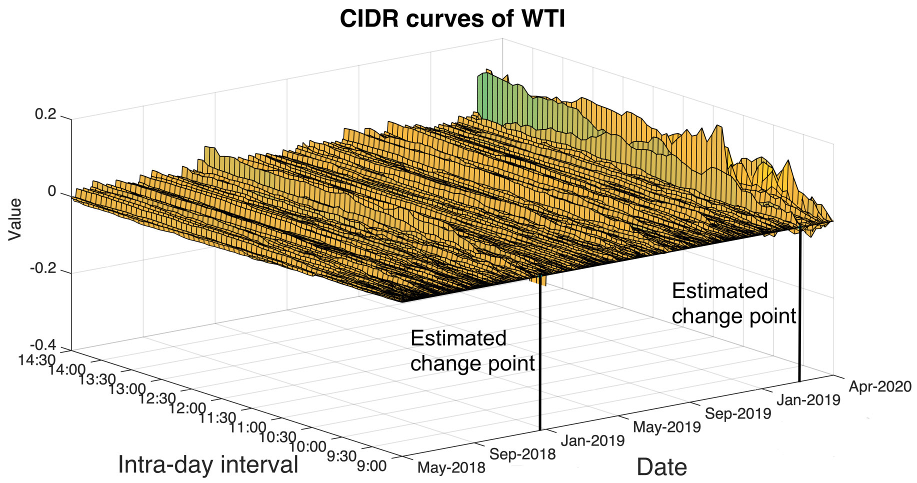
6. Detecting covariance changes in Crude oil intra-day return curves

In this section, we illustrate the use of the weighted CUSUM covariance change point detector and estimator in a data application to detect changes in the covariance of intra-day return curves derived from crude oil futures prices. Crude oil futures are believed to be sensitive to the overall sentiment of the market, and their variability is affected by major macroeconomic events (see Guo and Kliesen (2005), Liu and Gong (2020), Sharif et al. (2020)). As such, analyzing and forecasting oil futures volatility is an important and oft studied topic, and evaluating whether there are structural breaks in the covariance structure in futures prices series is an important step in doing so.

We consider two benchmark assets in the international crude oil pricing system: West Texas Intermediate (WTI) and Brent crude oil futures. The raw data that we consider were obtained from `www.backtestmarket.com`, and are comprised of 5-minute frequency front-month prices of WTI and Brent futures, from 9:00 am to 2:30 pm, each trading day from 12 May, 2018 to 30 Apr, 2020, which totals 502 days. As such there are 67 discrete observations of the price within each day, which we linearly interpolated to produce intra-day price curves of the form $p_i(t)$ for each asset. Visualizations of these price curves for WTI are shown in Figure 2. Visualizations of the Brent prices curves are available in the supplement to this article.

We take as a goal of this analysis to evaluate whether the variability of the curves modelled by their covariance kernel undergoes structural breaks during the observation period. In order to study this series as a mean stationary functional time series, we transform them to cumulative intra-day return curves (CIDRs) via the transformation

Fig. 3: Daily cumulative intra–day return (CIDR) curves constructed from the WTI commodity futures prices curves illustrated in Figure 2, using (22). Estimates for change points in the covariance function obtained using binary segmentation are also displayed.



$$r_i(t) = \log(p_i(t)) - \log(p_i(0)), \tag{22}$$

where $p_i(0)$ is the opening price at 9:00 am on day i. Figure 3 shows the CIDR curves constructed from the WTI asset price curves. We applied a series of hypothesis tests to evaluate the stationarity, normality, and serial correlation structure of these CIDR curves (see Horváth et al., 2014, Kokoszka et al., 2017, Górecki et al., 2018, and Rice et al., 2020), the results of which suggested that both series of crude oil CIDR curves evolve as approximately mean stationary, non-Gaussian, serially uncorrelated and conditionally heteroscedastic functional time series.

We applied a test of H_0 to each series based on $T_N(1/4)$ to detect potential changes in the covariance of both curve sequences. Following the settings used in the simulation study, we used the Bartlett window function and the bandwidth $h = N^{1/5}$ in the calculation of the long–run covariance operator. The p-values of these tests for each asset was 0, suggesting the presence of a change point, and the change point estimates $\hat{k}_N(1/4)$ for the series corresponded to the dates March 5th, 2020 for WTI and February 28, 2020 for Brent. We segmented each series based on these estimates and performed tests of H_0 again within each segment in order to detect additional change points. One additional change point was detected at significance level 0.05 in the first segment of each series, and these were estimated at the locations December 28th, 2020, and December 29th, 2020, for the WTI and Brent series respectively (approximate p-values of 0.041 and 0.048, respectively). No further change points were detected of notable significance. The location of these change points are illustrated in Figure 3 for the WTI series. We note that we did not take into account multiple testing when producing these p-values, and an application of for example the Holm-Bonferroni method would suggest that the second two estimated change points are not significant at the level 0.05.

It is noteworthy that the change point in covariance estimates for the WTI and Brent series are very similar. The first break estimate coincides with the begining of the COVID–19 pandemic in the US. The second detected change for each series at the end of December 2018 coincides with the largest drop in oil-prices that had been observed in the previous three years, which followed an already long decline that began in November of that year.

7. Proof of Theorems 1–3

Lemma 1. *If Assumption 1 holds, then*

$$\left(E\|\epsilon_i \otimes \epsilon_i - \epsilon_{i,\ell} \otimes \epsilon_{i,\ell}\|_2^{\nu/2}\right)^{2/\nu} \le c2^{\nu}\ell^{-\alpha},$$

where $\{\epsilon_{i,\ell}(t), -\infty < i, \ell < \infty\}$ *are defined in Assumption 1.*

Proof. Since $\epsilon_i(t)\epsilon_i(s) - \epsilon_{i,\ell}(t)\epsilon_{i,\ell}(s) = \epsilon_i(t)(\epsilon_i(s) - \epsilon_{i,\ell}(t)) + \epsilon_{i,\ell}(t)(\epsilon_i(s) - \epsilon_{i,\ell}(s))$, by the Cauchy–Schwartz inequality we have

$$\begin{aligned}E\|\epsilon_i \otimes \epsilon_i - \epsilon_{i,\ell} \otimes \epsilon_{i,\ell}\|_2^{\nu/2} &\leq 2^{\nu/2}\left(E[\|\epsilon_i\|_2^{\nu/2}\|\epsilon_i - \epsilon_{i,\ell}\|_2^{\nu/2}] + E[\|\epsilon_{i,\ell}\|_2^{\nu/2}\|\epsilon_i - \epsilon_{i,\ell}\|_2^{\nu/2}]\right)\\ &\leq 2^{\nu/2}\left((E[\|\epsilon_i\|_2^{\nu}])^{1/2}(E[\|\epsilon_i - \epsilon_{i,\ell}\|_2^{\nu}])^{1/2} + (E\|\epsilon_{i,\ell}\|_2^{\nu})^{1/2}(E\|\epsilon_i - \epsilon_{i,\ell}\|_2^{\nu})^{1/2}\right)\end{aligned}$$

Now Assumption 1 implies the result. □

Proof of Theorem 1. Under the null hypothesis we have

$$\begin{aligned}&(X_i(t) - \bar{X}_N(t))(X_i(s) - \bar{X}_N(s)) - \frac{\lfloor Nu \rfloor}{N}\sum_{i=1}^{N}(X_i(t) - \bar{X}_N(t))(X_i(s) - \bar{X}_N(s))\\ &\quad = \epsilon_i(t)\epsilon_i(s) + (\mu(t) - \bar{X}_N(t))\epsilon_i(s) + (\mu(s) - \bar{X}_N(s))\epsilon_i(t) + (\mu(t) - \bar{X}_N(t))(\mu(s) - \bar{X}_N(s)).\end{aligned}$$

It follows from Theorem 1.1 of Berkes et al. (2013) (see also Jirak (2013)) that

$$\sup_{0 \leq u \leq 1}\int\left(N^{-1/2}\sum_{i=1}^{\lfloor Nu \rfloor}\epsilon_i(t)\right)^2 dt = O_P(1).$$

which yields

$$\sup_{0<u<1}\int\int\left(\sum_{i=1}^{\lfloor Nu \rfloor}(X_i(t) - \bar{X}_N(t))(X_i(s) - \bar{X}_N(s)) - \sum_{i=1}^{\lfloor Nu \rfloor}\epsilon_i(t)\epsilon_i(s)\right)^2 dtds = O_P(1).$$

Thus we conclude

$$\sup_{0<u<1}\int\int(Z_N(u,t,s) - Z_N^*(u,t,s))^2\, dtds = o_P(1),$$

where

$$Z_N^*(u,t,s) = N^{-1/2}\left(\sum_{i=1}^{\lfloor Nu \rfloor}\epsilon_i(t)\epsilon_i(s) - \frac{\lfloor Nu \rfloor}{N}\sum_{i=1}^{N}\epsilon_i(t)\epsilon_i(s)\right). \tag{23}$$

Using again Theorem 1.1 of Berkes et al. (2013) (cf. also Jirak, 2013), we can define a sequence of Gaussian processes $\{\bar{\Gamma}_N(u,t,s), 0 \leq u,t,s \leq 1\}$ such that

$$\sup_{0<u<1}\int\int\left(N^{-1/2}\sum_{i=1}^{\lfloor Nu \rfloor}(\epsilon_i(t)\epsilon_i(s) - \mathfrak{C}(t,s)) - \bar{\Gamma}_N(u,t,s)\right)^2 dtds = o_P(1) \tag{24}$$

with $E\bar{\Gamma}_N(u,t,s) = 0$ and $E\bar{\Gamma}_N(u,t,s)\bar{\Gamma}_N(u',t',s') = \min(u,u')\mathfrak{D}(t,t',s,s')$. Defining $\Gamma_N(u,t,s) = \bar{\Gamma}(u,t,s) - u\bar{\Gamma}(1,t,s)$, the result follows from (24). □

The proof of Theorem 2 requires a Hájek–Rényi type inequality for partial sums of weakly dependent random functions, which is described by the following Lemma.

Lemma 2. *If Assumption 1 holds and* $0 \leq \kappa < 1/2$, *then*

$$\max_{1\leq k\leq N}\frac{N^{-1/2+\kappa}}{k^{\kappa}}\left\|\sum_{i=1}^{k}\epsilon_i\right\|_2 = O_P(1), \quad \max_{1\leq k\leq N}\frac{N^{-1/2+\kappa}}{k^{\kappa}}\left\|\sum_{i=1}^{k}(\epsilon_i \otimes \epsilon_i - \mathfrak{C})\right\|_2 = O_P(1), \tag{25}$$

$$\max_{1\leq k<N}\frac{N^{-1/2+\kappa}}{(N-k)^{\kappa}}\left\|\sum_{i=k+1}^{N}\epsilon_i\right\|_2 = O_P(1), \text{ and } \max_{1\leq k<N}\frac{N^{-1/2+\kappa}}{(N-k)^{\kappa}}\left\|\sum_{i=k+1}^{N}(\epsilon_i \otimes \epsilon_i - \mathfrak{C})\right\|_2 = O_P(1). \tag{26}$$

Proof. We write

$$\max_{1\le k\le N}\frac{1}{k^\kappa}\left\|\sum_{i=1}^k \epsilon_i\right\|_2 \le \max_{1\le j\le \log N} e^{-(j-1)\kappa}\max_{e^{j-1}\le k\le e^j}\left\|\sum_{i=1}^k \epsilon_i\right\|_2.$$

Berkes et al. (2013), Theorem 3.2, show that Assumption 1 implies

$$E\left\{\max_{1\le k\le M}\left\|\sum_{i=1}^k \epsilon_i\right\|_2^\nu\right\}\le c_1 M^{\nu/2}.$$

Hence by Markov's inequality we obtain

$$\begin{aligned}&P\left\{\max_{1\le k\le N}\frac{N^{-1/2+\kappa}}{k^\kappa}\left\|\sum_{i=1}^k \epsilon_i\right\|_2\ge x\right\}\le P\left\{\max_{1\le j\le \log N} e^{-(j-1)\kappa}\max_{e^{j-1}\le k\le e^j}\left\|\sum_{i=1}^k \epsilon_i\right\|_2\ge xN^{1/2-\kappa}\right\}\\ &\le\sum_{j=1}^{\log N}P\left\{\max_{e^{j-1}\le k\le e^j}\left\|\sum_{i=1}^k \epsilon_i\right\|_2\ge xe^{(j-1)\kappa}N^{1/2-\kappa}\right\}\le\sum_{j=1}^{\log N}x^{-\nu}e^{-\nu(j-1)\kappa}N^{-\nu/2+\nu\kappa}E\left\{\max_{e^{j-1}\le k\le e^j}\left\|\sum_{i=1}^k \epsilon_i\right\|_2^\nu\right\}\le\frac{c_2}{x^\nu},\end{aligned}$$

completing the proof of the first half of (25). The same arguments gives the second half of (25), and with minor modification give (26). □

Proof of Theorem 2. First we prove that

$$\sup_{1/(N+1)<u\le 1/2}\frac{1}{u^{2\kappa}}\int\!\!\int\left(\sum_{i=1}^{\lfloor Nu\rfloor}(X_i(t)-\bar X_N(t))(X_i(s)-\bar X_N(s))-\sum_{i=1}^{\lfloor Nu\rfloor}\epsilon_i(t)\epsilon_i(s)\right)^2 dtds=O_P(1) \tag{27}$$

By Lemma 2 and the Cauchy-Schwarz inequality we have

$$\max_{1\le k\le N}\frac{N^{-1+2\kappa}}{k^{2\kappa}}\int\!\!\int|\mu(t)-\bar X_N(t)|\left|\sum_{i=1}^k\epsilon_i(s)\right|dts\le\|\mu-\bar X_N\|_2N^{-1/2+\kappa}\max_{1\le k\le N}\frac{N^{-1/2+\kappa}}{k^\kappa}\left\|\sum_{i=1}^k\epsilon_i\right\|_2=O_P\left(N^{-1+\kappa}\right)$$

and

$$\max_{1\le k\le N}\frac{N^{-1+2\kappa}}{k^{2\kappa}}\int\int|\mu(t)-\bar X_N(t)||\mu(s)-\bar X_N(s)|dtds=O_P\left(N^{-2+2\kappa}\right),$$

completing the proof of (27). Minor modifications of the proof of (27) give

$$\sup_{1/2\le u<1-\frac{1}{(N+1)}}\frac{1}{(1-u)^{2\kappa}}\int\!\!\int\left(\sum_{i=1}^{\lfloor Nu\rfloor}(X_i(t)-\bar X_N(t))(X_i(s)-\bar X_N(s))-\sum_{i=1}^{\lfloor Nu\rfloor}\epsilon_i(t)\epsilon_i(s)\right)^2 dtds=O_P(1).$$

Hence we need to prove only that

$$\sup_{1/(N+1)<u<1-1/(N+1)}\frac{1}{w_\kappa^2(u)}\iint\left(Z_N^*(u,t,s)-\Gamma_N(u,t,s)\right)^2dtds=o_P(1),$$

where Z_N^* is defined in (23). It follows from Theorem 1 that

$$\sup_{\delta\le u\le 1-\delta}\frac{1}{w_\kappa^2(u)}\iint\left(Z_N^*(u,t,s)-\Gamma_N(u,t,s)\right)^2dtds=o_P(1)$$

for all $0<\delta<1/2$. Let $\bar\nu=\nu/2$. According to the proof of Lemma 2, for $x>0$

$$P\left\{\max_{1\le k\le\lfloor N\delta\rfloor}\frac{N^{-1/2+\kappa}}{k^\kappa}\left\|\sum_{i=1}^k(\epsilon_i\otimes\epsilon_i-\mathfrak{C})\right\|_2\ge x\right\}\le\frac{c_3}{x^\nu}N^{-\bar\nu/2+\kappa\bar\nu}\sum_{j=1}^{\log\lfloor N\delta\rfloor}e^{-j\bar\nu\kappa+j\bar\nu/2}\le c_4\delta^{\nu(1/2-\kappa)}$$

and therefore

$$\lim_{\delta\to 0}\limsup_{N\to\infty} P\left\{\max_{1\le k\le \lfloor N\delta\rfloor} \frac{N^{-1/2+\kappa}}{k^{\kappa}}\left\|\sum_{i=1}^{k}(\epsilon_i\otimes\epsilon_i-\mathfrak{C})\right\|_2 \ge x\right\}=0$$

for all $x>0$. The distribution of $\bar{\Gamma}_N(u,t,s)$ does not depend on N and let $\bar{\Gamma}(u,t,s)$ be a Gaussian process distributed as $\bar{\Gamma}_N(t,s)$. It is easy to see that

$$\sup_{0<u\le\delta}\frac{u}{u^{\kappa}}\left(\int\int\bar{\Gamma}^2(1,t,s)dtds\right)^{1/2} \to 0 \quad \text{a.s., as } \delta\to 0.$$

Let $\lambda_1\ge\lambda_2\ge\ldots$ and $\phi_1(t,s),\phi_2(t,s),\ldots$ be the corresponding eigenvalues and eigenfunctions of $\mathfrak{D}(t,s,t',s')$ as in (5). Checking the covariance function one can verify that

$$\{\bar{\Gamma}(u,t,s), 0\le u,t,s\le 1\}\stackrel{\mathcal{D}}{=}\sum_{i=1}^{\infty}\lambda_i^{1/2}W_i(u)\phi_i(t,s),$$

where $\{W_1(u), 0\le u\le\},\{W_2(u),0\le u\le\},\ldots$ are independent Wiener processes. Using the orthonormality of the eigenfunctions we have

$$\int\int\Gamma^2(u,t,s)dtds\stackrel{\mathcal{D}}{=}\sum_{i=1}^{\infty}\lambda_i W_i^2(u).$$

According to Garsia (1972), there are independent and identically distributed random variables $\{\Xi_i, i\ge 1\}$ with finite moments such that

$$|W_i(u)|\le \Xi_i u^{1/2}(\log(1/u))^{1/2}, \quad 0\le s\le 1,$$

and therefore

$$E\sup_{0<u\le\delta}\frac{1}{u^{2\kappa}}\sum_{i=1}^{\infty}\lambda_i W_i^2(u)\le c_5\delta^{1-2\kappa}\log(1/\delta).$$

Thus we get

$$\sup_{0<u\le\delta}\frac{1}{u^{2\kappa}}\int\int\bar{\Gamma}^2(u,t,s)dtds\xrightarrow{P}0, \quad \text{as } \delta\to 0,$$

resulting in

$$\sup_{0<u\le\delta}\frac{1}{u^{2\kappa}}\int\int\Gamma_N^2(u,t,s)dtds\xrightarrow{P}0, \quad \text{as } \delta\to 0$$

for each N, on account of

$$\{\Gamma_N(u,t,s), 0\le u,t,s\le 1\}\stackrel{\mathcal{D}}{=}\{\bar{\Gamma}(t,t,s)-u\bar{\Gamma}(1,t,s), 0\le u,t,s\le 1\}.$$

Similar arguments give for each N

$$\sup_{1-\delta\le u<1}\frac{1}{(1-u)^{\kappa}}\int\int\Gamma_N^2(u,t,s)dtds\xrightarrow{P}0, \quad \text{as } \delta\to 0,$$

completing the proof. of (3). Similar arguments give (4) and therefore the details are omitted. □

Proof of Theorem 3. First we note that by the proofs of Theorems 1 and 2, the result in (3) is proven if we show

$$\frac{1}{N(\theta_N(1-\theta_N))^{2-2\kappa}\|\mathfrak{C}_\Delta\|_2^2}\times\max_{1\le k<N}\frac{N^{-1+4\kappa}}{(k(N-k))^{2\kappa}}\left\|\sum_{i=1}^{k}\hat{X}_i\otimes\hat{X}_i-\frac{k}{N}\sum_{i=1}^{N}\hat{X}_i\otimes\hat{X}_i\right\|_2^2 \xrightarrow{P} 1 \tag{28}$$

with

$$\hat{X}_i(t)=X_i(t)-\mu(t). \tag{29}$$

We have under the alternative

$$\sum_{i=1}^{k} \hat{X}_i(t)\hat{X}_i(s) = \begin{cases} \displaystyle\sum_{i=1}^{k} \epsilon_i(t)\epsilon_i(s), & 1 \le k \le k^* \\ \displaystyle\sum_{i=1}^{k^*} \epsilon_i(t)\epsilon_i(s) + \sum_{k^*+1}^{k} \epsilon_{i,A}(t)\epsilon_{i,A}(s), & k^*+1 \le k \le N \end{cases}$$

and therefore

$$\sum_{i=1}^{k} \hat{X}_i(t)\hat{X}_i(s) - \frac{k}{N}\sum_{i=1}^{N} \hat{X}_i(t)\hat{X}_i(s) = R_k(t,s) + g_k(t,s)$$

with

$$R_k(t,s) = \begin{cases} \displaystyle\sum_{i=1}^{k}(\epsilon_i(t)\epsilon_i(s) - \mathfrak{C}(t,s)) - \frac{k}{N}\left(\sum_{i=1}^{k^*}(\epsilon_i(t)\epsilon_i(s) - \mathfrak{C}(t,s)) + \sum_{i=k^*+1}^{N}(\epsilon_{i,A}(t)\epsilon_{i,A}(s) - \mathfrak{C}_A(t,s))\right), \\ \qquad 1 \le k \le k^*, \\ \displaystyle\sum_{i=1}^{k^*}(\epsilon_i(t)\epsilon_i(s) - \mathfrak{C}(t,s)) + \sum_{i=k^*+1}^{k}(\epsilon_{i,A}(t)\epsilon_{i,A}(s) - \mathfrak{C}_A(t,s) - \frac{k}{N}\Bigg(\sum_{i=1}^{k^*}(\epsilon_i(t)\epsilon_i(s) - \mathfrak{C}(t,s)) \\ \qquad + \displaystyle\sum_{i=k^*+1}^{N}(\epsilon_{i,A}(t)\epsilon_{i,A}(s) - \mathfrak{C}_A(t,s))\Bigg), \quad k^*+1 \le k \le N \end{cases}$$

and

$$g_k(t,s) = \begin{cases} \dfrac{k}{N}(N-k^*)\mathfrak{C}_\Delta(t,s), & 1 \le k \le k^*, \\ \dfrac{k^*}{N}(N-k)\mathfrak{C}_\Delta(t,s), & k^*+1 \le k \le N. \end{cases} \tag{30}$$

It follows from Berkes et al. (2013), as in the proof of Theorem 2,

$$\max_{1\le k<N} \frac{N^{-1+4\kappa}}{(k(N-k))^{2\kappa}} \|R_k\|_2^2 = O_P(1).$$

Elementary algebra gives

$$\max_{1\le k\le N} \frac{N^{-1+4\kappa}}{(k(N-k))^{2\kappa}} \|g_k\|_2^2 = N(\theta_N(1-\theta_N))^{2-2\kappa}\|\mathfrak{C}_\Delta\|_2^2(1+o(1)),$$

completing the proof of (28). Similar arguments give (4) and therefore the proof is omitted. □

8. Proof of Theorem 4

Theorems 1 and 2 imply the order of the weighted partial sums of the $\epsilon_i(t)\epsilon_i(s)$'s when $0 \le \kappa < 1/2$. Next we get a bound when $\kappa = 1/2$ and obtain the same results for the weighted sums of the $\epsilon_{i,A}(t)\epsilon_{i,A}(s)$'s.

Lemma 3. *If Assumptions 1 and 4 hold, then*

$$\max_{1\le k<N} \frac{N^{1/2}}{(k(N-k))^{1/2}} \left\| \sum_{i=1}^{k}(\epsilon_i \otimes \epsilon_i - \mathfrak{C}) - \frac{k}{N}\sum_{i=k+1}^{N}(\epsilon_i \otimes \epsilon_i - \mathfrak{C}) \right\|_2 = O_P((\log N)^{2/\nu}) \tag{31}$$

$$\max_{1\le k<N} \frac{N^{-1/2+2\kappa}}{(k(N-k))^{\kappa}} \left\| \sum_{i=1}^{k}(\epsilon_{i,A} \otimes \epsilon_{i,A} - \mathfrak{C}_A) - \frac{k}{N}\sum_{i=k+1}^{N}(\epsilon_{i,A} \otimes \epsilon_{i,A} - \mathfrak{C}_A) \right\|_2 = O_P(1) \tag{32}$$

$$\max_{1\le k<N}\frac{N^{1/2}}{(k(N-k))^{1/2}}\left\|\sum_{i=1}^{k}(\epsilon_{i,A}\otimes\epsilon_{i,A}-\mathfrak{C}_A)-\frac{k}{N}\sum_{i=k+1}^{N}(\epsilon_{i,A}\otimes\epsilon_{i,A}-\mathfrak{C}_A)\right\|_2=O_P((\log N)^{2/\nu}). \tag{33}$$

If $M>a>1$, then for all $x>0$

$$P\left\{\max_{1\le k\le M-a}\frac{1}{M-k}\left\|\sum_{i=k+1}^{M}(\epsilon_i\otimes\epsilon_i-\mathfrak{C})\right\|_2>xa^{-1/2}\right\}\le\frac{c_1}{x^{\nu/2}}, \tag{34}$$

and

$$P\left\{\max_{1\le k\le M-a}\frac{1}{M-k}\left\|\sum_{i=k+1}^{M}(\epsilon_{i,A}\otimes\epsilon_{i,A}-\mathfrak{C}_A)\right\|_2>xa^{-1/2}\right\}\le\frac{c_2}{x^{\nu/2}} \tag{35}$$

with some constant c_1 and c_2.

Proof. Let

$$\zeta_i(t,s)=\epsilon_i(t)\epsilon_i(s)-\mathfrak{C}(t,s)\quad\zeta_{i,A}(t,s)=\epsilon_{i,A}(t)\epsilon_{i,A}(s)-\mathfrak{C}_A(t,s). \tag{36}$$

We write

$$\max_{1\le k\le N}\frac{1}{k^{1/2}}\left\|\sum_{i=1}^{k}\zeta_i\right\|_2\le\max_{1\le j\le N}\max_{e^{j-1}\le k\le e^j}\frac{1}{k^{1/2}}\left\|\sum_{i=1}^{k}\zeta_i\right\|_2\le\max_{1\le j\le N}\max_{e^{j-1}\le k\le e^j}e^{-(j-1)/2}\left\|\sum_{i=1}^{k}\zeta_i\right\|_2.$$

It is shown Berkes et al. (2013) that

$$\max_{1\le k\le m}\left\|\sum_{i=1}^{k}\zeta_i(t,s)\right\|_2^{\nu/2}\le c_3m^{\nu/4} \tag{37}$$

and therefore

$$P\left\{\max_{1\le k\le N}\frac{1}{k^{1/2}}\left\|\sum_{i=1}^{k}\zeta_i\right\|_2\ge x(\log N)^{2/\nu}\right\}\le\frac{c_3}{x^{\nu/2}\log N}\sum_{j=1}^{\log N}e^{-\nu(j-1)/4}e^{\nu j/4}\le\frac{c_4}{x^{\nu/2}} \tag{38}$$

resulting in

$$\max_{1\le k\le N}\frac{1}{k^{1/2}}\left\|\sum_{i=1}^{k}\zeta_i\right\|_2=O_P\left((\log N)^{2/\nu}\right).$$

Similar arguments give

$$\max_{1\le k<N}\frac{1}{(N-k)^{1/2}}\left\|\sum_{i=k+1}^{N}\zeta_i\right\|_2=O_P\left((\log N)^{2/\nu}\right).$$

Hence (31) is proven. Assumption 4 and (37) yield as before that

$$\max_{1\le k\le m}\left\|\sum_{i=1}^{k}\zeta_{i,A}\right\|_2^{\nu/2}\le c_5m^{\nu/4}. \tag{39}$$

Repeating the proof of (31), one can establish (32) and (33), we only need to replace (37) with (39). Proceeding along the lines of (38)

$$P\left\{\max_{1\le k\le M-a}\frac{1}{M-k}\left\|\sum_{i=k+1}^{M}\zeta_i\right\|_2>xa^{-1/2}\right\}=P\left\{\max_{a\le\ell\le M-1}\frac{1}{\ell}\left\|\sum_{i=M-\ell}^{M}\zeta_i\right\|_2>xa^{-1/2}\right\}\le\frac{c_1}{x^{\nu/2}},$$

and therefore (34) holds. The same argument gives (35). □

Proof of Theorem 4. We use the decomposition

$$\sum_{i=1}^{k}(X_i(t)-\bar{X}_N(t))(X_i(s)-\bar{X}_N(s))-\frac{k}{N}\sum_{i=1}^{N}(X_i(t)-\bar{X}_N(t))(X_i(s)-\bar{X}_N(s))$$
$$=V_{k,1}(t,s)-V_{k,2}(t)(\bar{X}_N(s)-\mu(s))-V_{k,2}(s)(\bar{X}_N(t)-\mu(t))+g_k(t,s),$$

where

$$V_{k,1}(t,s)=\begin{cases}\displaystyle\sum_{i=1}^{k}\zeta_i(t,s)-\frac{k}{N}\left(\sum_{i=1}^{k^*}\zeta_i(t,s)+\sum_{i=k^*+1}^{N}\zeta_{i,A}(t,s)\right), & \text{if}\quad 1\le k\le k^*,\\[2ex] \displaystyle-\sum_{i=k+1}^{N}\zeta_{i,A}(t,s)+\left(1-\frac{k}{N}\right)\left(\sum_{i=1}^{k^*}\zeta_i(t,s)+\sum_{i=k^*+1}^{N}\zeta_{i,A}(t,s)\right), & \text{if}\quad k^*+1\le k\le N,\end{cases}$$

$$V_{k,2}(t)=\begin{cases}\displaystyle\sum_{i=1}^{k}\epsilon_i(t)-\frac{k}{N}\left(\sum_{i=1}^{k^*}\epsilon_i(t)+\sum_{i=k^*+1}^{N}\epsilon_{i,A}(t)\right), & \text{if}\quad 1\le k\le k^*,\\[2ex] \displaystyle-\sum_{i=k^*+1}^{N}\epsilon_{i,A}(t)+\left(1-\frac{k}{N}\right)\left(\sum_{i=1}^{k^*}\epsilon_i(t)+\sum_{i=k^*+1}^{N}\epsilon_{i,A}(t)\right), & \text{if}\quad k^*+1\le k\le N,\end{cases}$$

with $g_k(t,s)$, $\zeta_i(t,s)$ and $\zeta_{A,i}(t,s)$ defined in (30) and (36). We assume that $0\le\kappa<1/2$. It follows from Assumption 3 that $(k(N-k))^{-\kappa}\|g_k\|_2$ reaches its largest value at k^* and for all $0<\alpha<\theta<\beta<1$,

$$\frac{1}{N^{1/2}}\min_{N\alpha\le k\le N\beta}\|g_k\|_2\ \to\ \infty. \tag{40}$$

The proof of Theorem 2 (cf. Lemma 2) yields

$$\max_{1\le k<N}\left(\frac{N^2}{k(N-k)}\right)^{\kappa}\left\|\sum_{i=1}^{k}\zeta_i-\frac{k}{N}\sum_{i=1}^{N}\zeta_i\right\|_2=O_P(N^{1/2}), \tag{41}$$

and Lemma 3 imply

$$\max_{1\le k<N}\left(\frac{N^2}{k(N-k)}\right)^{\kappa}\left\|\sum_{i=1}^{k}\zeta_{i,A}-\frac{k}{N}\sum_{i=1}^{N}\zeta_{i,A}\right\|_2=O_P(N^{1/2}). \tag{42}$$

Similar arguments give

$$\max_{1\le k<N}\left(\frac{N^2}{k(N-k)}\right)^{\kappa}\left\|\sum_{i=1}^{k}\epsilon_i-\frac{k}{N}\sum_{i=1}^{N}\epsilon_i\right\|_2=O_P(N^{1/2}) \tag{43}$$

and

$$\max_{1\le k<N}\left(\frac{N^2}{k(N-k)}\right)^{\kappa}\left\|\sum_{i=1}^{k}\epsilon_{i,A}-\frac{k}{N}\sum_{i=1}^{N}\epsilon_{i,A}\right\|_2=O_P(N^{1/2}). \tag{44}$$

Putting together (41)–(44) we conclude

$$\max_{1\le k<N}\left(\frac{N^2}{k(N-k)}\right)^{\kappa}\left\|V_{k,1}-V_{k,2}\otimes(\bar{X}_N-\mu)-V_{k,2}\otimes(\bar{X}_N-\mu)\right\|_2=O_P(N^{1/2})$$

and therefore (40) implies

$$|\hat{k}_N(\kappa)-k^*|=o_P(N). \tag{45}$$

Next we show

$$\|\mathfrak{C}_\Delta\|_2^2 \left|\hat{k}_N - k^*\right| = o_P(N). \tag{46}$$

To do this we require a more detailed decomposition. We note that

$$\hat{k}_N(\kappa) = \operatorname{argmax}_{1\le k\le N} \int\int Q_k(t,s)dtds,$$

where

$$\begin{aligned}
Q_k(t,s) =& \left(\frac{N}{k(N-k)}\right)^{2\kappa} \left(V_{k,1}(t,s) - V_{k,2}(t)(\bar{X}_N(s)-\mu(s)) - V_{k,2}(s)(\bar{X}_N(t)-\mu(t)) + g_k(t,s)\right)^2 \\
&- \left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa} \left(V_{k^*,1}(t,s) - V_{k^*,2}(t)(\bar{X}_N(s)-\mu(s)) - V_{k^*,2}(s)(\bar{X}_N(t)-\mu(t)) \ +g_{k^*}(t,s)\right)^2 \\
=& Q_{k,1}(t,s) + \ldots + Q_{k,8}(t,s)
\end{aligned}$$

with

$$\begin{aligned}
Q_{k,1}(t,s) &= \left(\frac{N}{k(N-k)}\right)^{2\kappa} (V_{k,1}(t,s) + g_k(t,s))^2 - \left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa} (V_{k^*,1}(t,s) + g_{k^*}(t,s))^2, \\
Q_{k,2}(t,s) &= \left(\frac{N}{k(N-k)}\right)^{2\kappa} V_{k,2}^2(t)(\bar{X}_N(s)-\mu(s))^2 - \left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa} V_{k^*,2}^2(t)(\bar{X}_N(s)-\mu(s))^2 \\
Q_{k,3}(t,s) &= Q_{k,2}(s,t), \\
Q_{k,4}(t,s) &= -2\left(\frac{N}{k(N-k)}\right)^{2\kappa} V_{k,1}(t,s)V_{k,2}(t)(\bar{X}_N(s)-\mu(s)) + 2\left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa} V_{k^*,1}(t,s)V_{k^*,2}(t)(\bar{X}_N(s)-\mu(s)), \\
Q_{k,5}(t,s) &= Q_{k,4}(s,t), \\
Q_{k,6}(t,s) &= 2\left(\frac{N}{k(N-k)}\right)^{2\kappa} V_{k,2}(t)(\bar{X}_N(s)-\mu(s))V_{k,2}(s)(\bar{X}_N(t)-\mu(t)) \\
&\quad - \left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa} 2V_{k^*,2}(t)(\bar{X}_N(s)-\mu(s))V_{k^*,2}(s)(\bar{X}_N(t)-\mu(t)), \\
Q_{k,7}(t,s) &= -2\left(\frac{N}{k(N-k)}\right)^{2\kappa} V_{k,2}(t)(\bar{X}_N(s)-\mu(s))g_k(t,s) + 2\left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa} V_{k^*,2}(t)(\bar{X}_N(s)-\mu(s))g_{k^*}(t,s), \\
Q_{k,8}(t,s) &= Q_{k,7}(s,t).
\end{aligned}$$

We provide details for $1 \le k \le k^*$. Using (44), we only need to investigate when $N\alpha \le k \le k^* - a$, where $0 < \alpha < \theta$ and $a = C/\|\mathfrak{C}_\Delta\|_2^2$. We start with $Q_{k,1}(t,s)$. We write $Q_{k,1}(t,s)$ as $Q_{k,1}(t,s) = Q_{k,1,1}(t,s) + \ldots + Q_{k,1,8}(t,s)$, where

$$\begin{aligned}
Q_{k,1,1}(t,s) &= \left\{\left(\frac{N}{k(N-k)}\right)^{2\kappa} - \left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa}\right\} \left(\sum_{i=1}^{k} \zeta_i(t,s) - \frac{k}{N}\left(\sum_{i=1}^{k^*} \zeta_i(t,s) \ + \sum_{i=k^*+1}^{N} \zeta_{i,A}(t,s)\right)\right)^2, \\
Q_{k,1,2}(t,s) &= -2\left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa} \left(\sum_{i=k+1}^{k^*} \zeta_i(t,s)\right) \left\{\sum_{i=1}^{k} \zeta_i(t,s) - \frac{k}{N}\left(\sum_{i=1}^{k^*} \zeta_i(t,s)\right.\right. \\
&\qquad \left.\left. + \sum_{i=k^*+1}^{N} \zeta_{i,A}(t,s)\right) + \sum_{i=1}^{k^*} \zeta_i(t,s) - \frac{k^*}{N}\left(\sum_{i=1}^{k^*} \zeta_i(t,s) + \sum_{i=k^*+1}^{N} \zeta_{i,A}(t,s)\right)\right\}, \\
Q_{k,1,3}(t,s) &= 2\left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa} \frac{k^*-k}{N} \left(\sum_{i=1}^{k^*} \zeta_i(t,s) + \sum_{i=k^*+1}^{N} \zeta_{i,A}(t,s)\right) \left\{\sum_{i=1}^{k} \zeta_i(t,s)\right.
\end{aligned}$$

$$-\frac{k}{N}\left(\sum_{i=1}^{k^*}\zeta_i(t,s)+\sum_{i=k^*+1}^{N}\zeta_{i,A}(t,s)\right)+\sum_{i=1}^{k^*}\zeta_i(t,s)\;-\frac{k^*}{N}\left(\sum_{i=1}^{k^*}\zeta_i(t,s)+\sum_{i=k^*+1}^{N}\zeta_{i,A}(t,s)\right)\Bigg\},$$

$$Q_{k,1,4}(t,s)=-2\left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa}(g_k(t,s)-g_{k^*}(t,s))$$

$$Q_{k,1,5}(t,s)=-\left\{\left(\frac{N}{k(N-k)}\right)^{2\kappa}-\left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa}\right\}\left(\sum_{i=1}^{k^*}\zeta_i(t,s)-\frac{k^*}{N}\left(\sum_{i=1}^{k^*}\zeta_i(t,s)\;+\sum_{i=k^*+1}^{N}\zeta_{i,A}(t,s)\right)\right)\times g_{k^*}(t,s),$$

$$Q_{k,1,6}(t,s)=2\left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa}\frac{k-k^*}{N}\left(\sum_{i=1}^{k^*}\zeta_i(t,s)+\sum_{i=k^*+1}^{N}\zeta_{i,A}(t,s)\right)g_{k^*}(t,s),$$

$$Q_{k,1,7}(t,s)=-2\left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa}\left(\sum_{i=1}^{k}\zeta_i(t,s)-\sum_{i=1}^{k^*}\zeta_i(t,s)\right),$$

$$Q_{k,1,8}(t,s)=\left(\frac{N}{k(N-k)}\right)^{2\kappa}g_k^2(t,s)-\left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa}g_{k^*}^2(t,s).$$

Using the mean value theorem we get

$$\max_{N\alpha\le k\le k^*}\frac{1}{k^*-k}\left|\left(\frac{N}{k(N-k)}\right)^{2\kappa}-\left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa}\right|=O(N^{-1-2\kappa})$$

and therefore by Theorem 2 and (32)

$$\max_{N\alpha\le k\le k^*}\frac{1}{N^{1-2\kappa}\|\mathfrak{C}_\Delta\|_2^2(k-k^*)}\int|Q_{k,1,1}(t,s)|dtds=O_P\left(\frac{N^{-1-2\kappa}N}{N^{1-2\kappa}\|\mathfrak{C}_\Delta\|_2^2}\right)=O_P\left(\frac{1}{N\|\mathfrak{C}_\Delta\|_2^2}\right)=o_P(1). \tag{47}$$

By the Cauchy–Schwartz inequality for integrals and Lemma 3 we have

$$\max_{N\alpha\le k\le k^*-a}\frac{1}{k^*-k}\int\!\!\int|Q_{k,1,2}(t,s)|dtds=O_P\left(N^{-2\kappa-1/2}\right)\max_{N\alpha\le k\le k^*-a}\frac{1}{k^*-k}\left\|\sum_{i=k+1}^{k^*}\zeta_i\right\|_2.$$

So by (34) we get

$$\max_{N\alpha\le k\le k^*}\frac{1}{N^{1-2\kappa}\|\mathfrak{C}_\Delta\|_2^2(k-k^*)}\int|Q_{k,1,2}(t,s)|dtds=O_P\left(\frac{N^{-2\kappa+1/2}}{N^{1-2\kappa}\|\mathfrak{C}_\Delta\|_2^2}\|\mathfrak{C}_\Delta\|_2\right)=o_P(1). \tag{48}$$

Lemma 3 implies

$$\max_{N\alpha\le k\le k^*}\frac{1}{N^{1-2\kappa}\|\mathfrak{C}_\Delta\|_2^2(k-k^*)}\int|Q_{k,1,3}(t,s)|dtds=O_P\left(\frac{1}{N\|\mathfrak{C}_\Delta\|_2^2}\right)=o_P(1). \tag{49}$$

Similarly to (47)

$$\max_{N\alpha\le k\le k^*}\frac{1}{N^{1-2\kappa}\|\mathfrak{C}_\Delta\|_2^2(k-k^*)}\int|Q_{k,1,4}(t,s)|dtds=O_P\left(\frac{1}{N^{1/2}\|\mathfrak{C}_\Delta\|_2}\right)=o_P(1), \tag{50}$$

$$\max_{N\alpha\le k\le k^*}\frac{1}{N^{1-2\kappa}\|\mathfrak{C}_\Delta\|_2^2(k-k^*)}\int|Q_{k,1,5}(t,s)|dtds=O_P\left(\frac{1}{N^{1/2}\|\mathfrak{C}_\Delta\|_2}\right)=o_P(1) \tag{51}$$

and

$$\max_{N\alpha\le k\le k^*}\frac{1}{N^{1-2\kappa}\|\mathfrak{C}_\Delta\|_2^2(k-k^*)}\int|Q_{k,1,6}(t,s)|dtds=O_P\left(\frac{1}{N^{1/2}\|\mathfrak{C}_\Delta\|_2}\right)=o_P(1). \tag{52}$$

Using (34) and the definition of a we get that

$$\max_{N\alpha\le k\le k^*}\frac{1}{N^{1-2\kappa}\|\mathfrak{C}_\Delta\|_2^2(k-k^*)}\int|Q_{k,1,7}(t,s)|dtds=\frac{1}{C^{1/2}}O_P(1).$$

and therefore

$$\lim_{C\to\infty}\limsup_{N\to\infty}P\left\{\max_{N\alpha\le k\le k^*}\frac{1}{N^{1-2\kappa}\|\mathfrak{C}_\Delta\|_2^2(k-k^*)}\int|Q_{k,1,7}(t,s)|dtds>x\right\}=0. \tag{53}$$

for any $x>0$. Elementary algebra yields that with some $c_1>0$ and $c_2>0$

$$-c_1N^{1-2\kappa}\|\mathfrak{C}_\Delta\|_2^2(k^*-k)\le\iint Q_{k,1,8}(t,s)dtds\le -c_2N^{1-2\kappa}\|\mathfrak{C}_\Delta\|_2^2(k^*-k), \tag{54}$$

if $N\alpha\le k\le k^*$. We then write $Q_{k,2}(t,s)=Q_{k,2,1}(t,s)+Q_{k,2,2}(t,s)$, where

$$Q_{k,2,1}(t,s)=\left\{\left(\frac{N}{k(N-k)}\right)^{2\kappa}-\left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa}\right\}V_{k,2}^2(t)(\bar X_N(s)-\mu(s))^2$$

and

$$Q_{k,2,2}(t,s)=\left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa}\left[V_{k,2}^2(t)(\bar X_N(s)-\mu(s))^2-V_{k^*,2}^2(t)(\bar X_N(s)-\mu(s))^2\right].$$

Assumption 1 implies with Theorem 3.2 of Berkes et al. (2013) (see also and Jirak, 2013) that

$$\max_{1\le k\le N}\|V_{k,2}\|_2=O_P(N^{1/2})\quad\text{and}\quad\max_{1\le k\le N}\|\bar X_N-\mu\|_2=O_P(N^{-1/2}), \tag{55}$$

so the mean value theorem gives

$$\max_{N\alpha\le k\le k^*-a}\frac{1}{N^{1-2\kappa}(k^*-k)\|\mathfrak{C}_\Delta\|_2^2}\iint|Q_{k,2,1}(t,s)|dtds=O_P\left(\frac{1}{N\|\mathfrak{C}_\Delta\|_2^2}\right)=o_P(1).$$

Elementary algebra yields

$$Q_{k,2,2}(t,s)=\left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa}[V_{k,2}(t)-V_{k^*,2}(t)]\,[V_{k,2}(t)+V_{k^*,2}(t)]\,(\bar X_N(s)-\mu(s))^2$$

and therefore

$$\max_{N\alpha\le k\le k^*-a}\frac{1}{N^{1-2\kappa}(k^*-k)\|\mathfrak{C}_\Delta\|_2^2}\iint|Q_{k,2,2}(t,s)|dtds=O_P\left(\frac{1}{N^{3/2}\|\mathfrak{C}_\Delta\|_2^2}\right)\max_{N\alpha\le k\le k^*-a}\frac{1}{k^*-k}\left\|\sum_{i=k+1}^{k^*}\zeta_i\right\|_2.$$

Hence Lemma 3 implies $\max_{N\alpha\le k\le k^*-a}[N^{1-2\kappa}(k^*-k)\|\mathfrak{C}_\Delta\|_2^2]^{-1}\iint|Q_{k,2,2}(t,s)|dtds=o_P(1)$, and thus

$$\max_{N\alpha\le k\le k^*-a}\frac{1}{N^{1-2\kappa}(k^*-k)\|\mathfrak{C}_\Delta\|_2^2}\iint|Q_{k,2}(t,s)|dtds=o_P(1). \tag{56}$$

Repeating the proof of (56), one can verify

$$\max_{N\alpha\le k\le k^*-a}\frac{1}{N^{1-2\kappa}(k^*-k)\|\mathfrak{C}_\Delta\|_2^2}\iint|Q_{k,i}(t,s)|dtds=o_P(1),\quad i=3,4,5. \tag{57}$$

We may also write $Q_{k,6}(t,s)=2Q_{k,6,1}(t,s)+2Q_{k,6,2}(t,s)+2Q_{k,6,3}(t,s)$, where

$$Q_{k,6,1}(t,s)=\left\{\left(\frac{N}{k(N-k)}\right)^{2\kappa}-\left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa}\right\}V_{k,2}(t)(\bar X_N(s)-\mu(s))V_{k,2}(s)(\bar X_N(t)-\mu(t)),$$

$$Q_{k,6,2}(t,s)=\left(\frac{N}{k^*(N-k^*)}\right)^{2\kappa}(V_{k,2}(t)-V_{k^*,2}(t))(\bar{X}_N(s)-\mu(s))V_{k,2}(s)(\bar{X}_N(t)-\mu(t))$$

and $Q_{k,6,3}(t,s)=Q_{k,6,2}(s,t)$. Similarly to (55) $\max_{1\le k\le N}\|V_{k,2}(t)\|_2=O_P(N^{1/2})$, and hence by the mean value theorem we have

$$\max_{N\alpha\le k\le k^*-a}\frac{1}{N^{1-2\kappa}(k^*-k)\|\mathfrak{C}_\Delta\|_2^2}\iint|Q_{k,6,1}(t,s)|dtds=o_P(1).$$

This and Lemma 3 yield

$$\max_{N\alpha\le k\le k^*-a}\frac{1}{N^{1-2\kappa}(k^*-k)\|\mathfrak{C}_\Delta\|_2^2}\iint|Q_{k,6,2}(t,s)|dtds=O_P\left(\frac{1}{N^{1/2}\|\mathfrak{C}_\Delta\|_2}\right)=o_P(1).$$

Thus we conclude

$$\max_{N\alpha\le k\le k^*-a}\frac{1}{N^{1-2\kappa}(k^*-k)\|\mathfrak{C}_\Delta\|_2^2}\iint|Q_{k,6}(t,s)|dtds=o_P(1).\tag{58}$$

Using the same arguments as above one can show that

$$\max_{N\alpha\le k\le k^*-a}\frac{1}{N^{1-2\kappa}(k^*-k)\|\mathfrak{C}_\Delta\|_2^2}\iint|Q_{k,i}(t,s)|dtds=o_P(1),\quad i=7,8.\tag{59}$$

Putting together (47)–(59) for all $K<0$ we have

$$\lim_{C\to\infty}\limsup_{N\to\infty}P\left\{\max_{N\alpha\le k\le k^*-C/\|\mathfrak{C}_\Delta\|_2^2}\iint\left(\sum_{\ell=1}^{8}Q_{k,\ell}(t,s)\right)dtds>K\right\}=0,$$

By symmetry, for all $K<0$ and $\theta<\beta$

$$\lim_{C\to\infty}\limsup_{N\to\infty}P\left\{\max_{k^*-C/\|\mathfrak{C}_\Delta\|_2^2\le k\le N\beta}\iint\left(\sum_{\ell=1}^{8}Q_{k,\ell}(t,s)\right)dtds>K\right\}=0,$$

which completes the proof of (46).

With minor modifications of the arguments used in (47)–(59), one can prove that for all $C>0$

$$\max_{|k^*-k|\le C/\|\mathfrak{C}_\Delta\|_2^2}\frac{1}{N^{1-2\kappa}}\iint\sum_{\ell=2}^{8}|Q_{k,\ell}(t,s)|dtds=o_P(1)\tag{60}$$

and

$$\max_{|k^*-k|\le C/\|\mathfrak{C}_\Delta\|_2^2}\frac{1}{N^{1-2\kappa}}\iint\sum_{\ell=1}^{6}|Q_{k,1,\ell}(t,s)|dtds=o_P(1).\tag{61}$$

Thus we proved that it is enough to investigate the properties of $Q_{k,1,7}(t,s)$ and $Q_{k,1,8}(t,s)$, when $|k-k^*|=O(1/\|\mathfrak{C}_\Delta\|_2^2)$. According to the definitions of $Q_{k,1,7}(t,s)$ and $Q_{k,1,8}(t,s)$

$$\frac{1}{N^{1-2\kappa}}\iint Q_{k,1,7}(t,s)ds=\begin{cases}-2\left(\dfrac{k^*(N-k^*)}{N^2}\right)^{1-2\kappa}\|\mathfrak{C}_\Delta\|_2\displaystyle\sum_{i=k+1}^{k^*}\iint\zeta_i(t,s)\mathfrak{C}_N^*(t,s)dtds, & \text{if } k<k^*,\\ 0,\quad \text{if } k=k^*, & \\ 2\left(\dfrac{k^*(N-k^*)}{N^2}\right)^{1-2\kappa}\|\mathfrak{C}_\Delta\|_2\displaystyle\sum_{i=k^*+1}^{k}\iint\zeta_{i,A}(t,s)\mathfrak{C}_N^*(t,s)dtds, & \text{if } k>k^*.\end{cases}$$

with $\mathfrak{C}_N^*(t,s) = \mathfrak{C}_\Delta(t,s)/\|\mathfrak{C}_\Delta\|_2$, and

$$Q_{k,1,8}(t,s) = \left(\frac{N}{k(N-k)}\right)^\kappa g_k^2(t,s) - \left(\frac{N}{k^*(N-k^*)}\right)^\kappa g_{k^*}^2(t,s).$$

Using Assumptions 4, 5 and (42) we get for all $C > 0$

$$\sup_{0\le u\le C} \|\mathfrak{C}_\Delta\|_2 \left| \sum_{i=k^*+1}^{k^*+u\tau^2/\|\mathfrak{C}_\Delta\|_2} \iint (\zeta_{i,A}(t,s) - \zeta_i(t,s))\mathfrak{C}_N^*(t,s)dtds \right| = o_P(1).$$

Due to Assumptions 1 and 4 it is straightforward to show that

$$\left\{ \frac{1}{N^{1-2\kappa}} \iint Q_{k^*+u\tau^2/\|\mathfrak{C}_\Delta\|_2^2,1,7}(t,s)dtds,\ u<0 \right\} \text{ and } \left\{ \frac{1}{N^{1-2\kappa}} \iint Q_{k^*+u\tau^2/\|\mathfrak{C}_\Delta\|_2^2,1,7}(t,s)dtds,\ u\ge 0 \right\},$$

are asymptotically independent. Thus we have for all $C > 0$

$$\frac{1}{N^{1-2\kappa}} \iint Q_{k^*+u\tau^2/\|\mathfrak{C}_\Delta\|_2^2,1,7}(t,s)dtds \overset{\mathcal{D}[-C,C]}{\longrightarrow} 2(\theta(1-\theta))^{1-2\kappa}\tau^2 W(u), \tag{62}$$

where $\{W(u), -\infty < u < \infty\}$ is the two sided Wiener process of (20). Elementary arguments give

$$\sup_{|u|\le C} \left| \frac{1}{N^{1-2\kappa}} \iint Q_{k^*+u\tau^2/\|\mathfrak{C}_\Delta\|_2^2,1,8}(t,s)dtds + 2(\theta(1-\theta))^{1-2\kappa}\tau^2 m_\kappa(u) \right| = o(1), \tag{63}$$

where $m_\kappa(u)$ is defined in (19). Since we can take C as large as we wish, Theorem 4 is proven when $0 \le \kappa < 1/2$. The second part of Theorem 4 is based on (31) and (33). Replacing (41)–(44) with (31) and (33), one can verify that (45) holds when $\kappa = 1/2$. The rest of the proof remains the same. □

9. Acknowledgments

We wish to thank guest editor for this special issue, and two anonymous referees, whose detailed and constructive comments led to important improvements to this work. We also wish to give special thanks to the Editor-in-Chief, who provided numerous helpful suggestions to improve the presentation of our paper.

References

[1] Aneiros, G., Cao, R., Fraiman, R. Genest, C. and Vieu, P.: Recent advances in functional data analysis and high-dimensional statistics. *Journal of Multivariate Analysis* **170**(2019), 3-9.

[2] Antoch, J., Hušková, M.. and Prášková, Z.: Effect of dependence on statistics for determination of change. *Journal of Statistical Planning and Inference* **60**(1997), 291–310.

[3] Antoch, J., Hušková, M. and Veraverbeke, N.: Change–point problem and bootstrap. *Journal of Nonparametric Statistics* **5**(1995), 123–144.

[4] Aston, J.A.D. and Kirch, C.: Evaluating stationarity via change–point alternatives with applications to fMRI data. *Annals of Applied Statistics* **6**(2012a), 1906–1948.

[5] Aston, J.A.D. and Kirch, C.: Detecting and estimating changes in dependent functional data. *Journal of Multivariate Analysis* **109**(2012b), 204–220.

[6] Aue, A., Gabrys, R., Horváth, L. and Kokoszka, P.: Estimation of a change–point in the mean function of functional data. *Journal of Multivariate Analysis*, **100**(2009), 2254–2269.

[7] Aue, A., Hörmann, S., Horváth, L. and M. Hušková.: Dependent functional linear models with applications to monitoring structural change. *Statistica Sinica* **24**(2014), 1043–1073.

[8] Aue, A., Hörmann, S., Horváth, L. and Reimherr, M.: Break detection in the covariance structure of multivariate time series models. *Annals of Statistics* **37**(2009), 4046–4087.

[9] Aue, A., Horváth, L. and Pellatt, D.: Functional generalized autoregressive conditional heteroscedasticity. *Journal of Time Series Analysis* **38**(2017), 3–21.

[10] Aue, A., Rice, G. and Sönmez, O.: Detecting and dating structural breaks in functional data without dimension reduction. *Journal of the Royal Statistical Society, Series B* **80**(2018), 509–529.

[11] Aue, A. and Rice, G. and Sönmez, O.: Structural break analysis for spectrum and trace of covariance operators, *Environmetrics*, **31**, (2020).

[12] Bardsley, P., Horváth, L., Kokoszka, P. and Young, G.: Change point tests in functional factor models with application to yield curves. *Econometrics Journal* **20**(2017), 373–403.

[13] Berens, T., Weiss, G.N.F. and Wied, D.: Testing for structural breaks in correlations: Does it improve Value–at–Risk forecasting? *Journal of Empirical Finance* **32**(2015), 135-152
[14] Berkes, I., Gabrys, R., Horváth, L. and Kokoszka, P.: Detecting changes in the mean of functional observations. *Journal of the Royal Statatistical Society, Series B* **71**(2009), 927–946.
[15] Berkes, I., Horváth, L. and Rice, R.: Weak invariance principles for sums of dependent random functions. *Stochastic Processes and Their Applications* **123**(2013), 385–403.
[16] Berkes, I., Horváth, L. and Rice, R.: On the asymptotic normality of kernel estimators of the long run covariance of functional time series. *Journal of Multivariate Analysis*, **144**(2016), 150–175.
[17] Bosq, D.: *Linear Processes in Function Spaces* Springer, New York, NY, USA, 2000.
[18] Cerovecki, C., Francq, C., Hörmann S. and Zakoıan J.-M.: Functional GARCH models: the quasi–likelihood approach and its applications. *Journal of Econometrics* **209**(2019), 353–375.
[19] Cuevas, A.: A partial overview of the theory of statistics with functional data *Journal of Statistical Planning and Inference* **147**(2014), 1–23.
[20] Dette H. and Kutta T. Detecting structural breaks in eigensystems of functional time series. 2019. preprint, available on arxiv.
[21] Dümbgen. L.: The asymptotic behavior of some nonparametric change–point estimators. *Annals of Statistics* **19**(1991), 1471–1495.
[22] Ferraty, V. and Vieu, P.: *Nonparametric Functional Data Analysis: Theory and Practice.* Springer, 2006.
[23] Fraiman, R., Justel. A., Liu, R. and Llop, P.: Detecting trends in time series of functional data: A study of Antarctic climate change. *The Canadian Journal of Statistics* **42**(2014), 597–6-9.
[24] Galeano, P. and Wied, D.: Multiple break detection in the correlation structure of random variables. *Computational Statistics & Data Analysis*, **76**(2014), 262–282.
[25] Garsia, A.M.: Continuity properties of Gaussian processes with multidimensional time parameter. *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability* Volume **2**, University of California Press, 1972, 369–374.
[26] Goia, A., and Vieu, P.: An introduction to recent advances in high/infinite dimensional statistics. *Journal of Multivariate Analysis* **146**(2016), 1-6.
[27] Górecki, T., Hörmann, S., Horváth, L. and Kokoszka, P.: Testing normality of functional time series. *Journal of Time Series Analysis* **39**(2018), 471–487.
[28] Guo, H. and Kliesen, K.L.: Oil price volatility and US macroeconomic activity. *Review-Federal Reserve Bank of Saint Louis 87*(2005), 669.
[29] Hörmann S., and Kokoszka P.: Functional time series. In C. R. Rao and T. S. Rao, editors, Time Series, volume 30 of Handbook of Statistics. Elsevier, 2012.
[30] Horváth, L. and Kokoszka, P.: *Inference for Functional Data with Applications.* Springer, New York, 2012.
[31] Horváth, L., Kokoszka, P. and Rice, G.: Testing stationarity of functional time series. *Journal of Econometrics* **179**(2014), 66–82.
[32] Horváth, L., Miller, C. and Rice, G.: A new class of change point test statistics of Rényi type. *Journal of Business and Economic Statistics* (2020) To appear.
[33] Horváth, L. and Rice, G.: Extensions of some classical methods in change point analysis (with discussions). *TEST* 23(2014), 219–290.
[34] Horváth, L. Rice, G., and Zhao, Y.: Supplement to "Change point analysis of covariance functions: a weighted cumulative sum approach". *Journal of Multivariate Analysis* (2021+).
[35] Inclán, C. and Tiao, G.C.: Use of cumulative sums of squares for retrospective detection of change of variance. *Journal of the American Statistical Association* 89 (1994), 913–923.
[36] Jaruškova, D.: Testing for a change in the covariance operator, *Journal of Statistical Planning and Inference*, 143 (2013), 1500–1511.
[37] Jiao, S., Frostig, R.D. and Ombao, H.: Break point detection for functional covariance. Preprint, 2020.
[38] Jirak, M.: On weak invariance principles for sums of dependent random functionals. *Statistics & Probability Letters* **83**(2013), 2291–2296.
[39] Kim, J. and Pollard, D.: Cube Root Asymptotics. *Annals of Statistics***18** (1990) 191–219.
[40] Kokoszka, P.: Dependent functional data *International Scholarly Research Notices*, 2012.
[41] Kokoszka, P. Reimherr, M.: *Introduction to Functional Data Analysis.* Chapman and Hall/CRC, 2017.
[42] Kokoszka, P., Rice, G. and Shang, H.L.: Inference for the autocovariance of a functional time series under conditional heteroscedasticity. *Journal of Multivariate Analysis* **162** (2017), 32–50.
[43] Küchnert, S.: Functional ARCH and GARCH models: A Yule–Walker approach. *Electronic Journal of Statistics* **14**(2020), 4321—4360.
[44] Liu, T. and Gong, X.: Analyzing time-varying volatility spillovers between the crude oil markets using a new method. *Energy Economics* **87**(2020), 104711.
[45] Ramsey, J.O. and Silverman, B.W.: *Functional Data Analysis.* Springer, 2002.
[46] Rice, G., Wirjanto, T. and Zhao, Y.: Tests for conditional heteroscedasticity of functional data. *Journal of Time Series Analysis* In press, 2020.
[47] Sharif, A., Aloui, C. and Yarovaya, L.: COVID–19 pandemic, oil prices, stock market, geopolitical risk and policy uncertainty nexus in the US economy: Fresh evidence from the wavelet–based approach. *International Review of Financial Analysis*, 2020, 101496.
[48] Sharipov O., Tewes, J. and Wendler, M.: Sequential block bootstrap in a hilbert space with application to change point analysis. *Canadian Journal of Statistics*, **44** 300–322, (2016).
[49] Sharipov O., and Wendler, M.: Bootstrapping covariance operators of functional time series *Journal of Nonparametric Statistics*, **32** (2020) 1-19.
[50] Stoehr, C., Aston, J.A.D. and Kirch, C.: Detecting changes in the covariance structure of functional time series with application to fMRI data *Econometrics and Statsitics*, 2020 (to appear).
[51] Wied, D.: A nonparametric test for a constant correlation matrix. *Econometric Reviews* **36**(2017), 1157–1172.
[52] Wied, D., Krämer, W. and Dehling, H.: Testing for a change in correlation at an unknown point in time using an extended functional delta method. *Econometric Theory* **28**(2012), 570–589.