mathematics

Article

Multiple Change-Point Detection in a Functional Sample via the $\mathcal{G}$-Sum Process

Tadas Danielius * and Alfredas Račkauskas

Institute of Applied Mathematics, Vilnius University, 03225 Vilnius, Lithuania; alfredas.rackauskas@mif.vu.lt
* Correspondence: tadas.danielius@mif.vu.lt

Abstract: We first define the $\mathcal{G}$-CUSUM process and investigate its theoretical aspects including asymptotic behavior. By choosing different sets $\mathcal{G}$, we propose some tests for multiple change-point detections in a functional sample. We apply the proposed testing procedures to the real-world neurophysiological data and demonstrate how it can identify the existence of the multiple change-points and localize them.

Keywords: p-variation; functional data; functional change-point detection; functional principal component analysis

MSC: 62R10

Citation: Danielius, T.; Račkauskas, A. Multiple Change-Point Detection in a Functional Sample via the $\mathcal{G}$-Sum Process. *Mathematics* **2022**, *10*, 2294. https://doi.org/10.3390/math10132294

Academic Editor: Jin-Ting Zhang

Received: 8 June 2022
Accepted: 27 June 2022
Published: 30 June 2022

1. Introduction

Consider a second-order stationary sequence of stochastic processes $Y_i = (Y_i(t), t \in [0,1]), i \in \mathbb{N}$, defined on a probability space $(\Omega, \mathcal{F}, P)$, having zero mean and covariance function $\gamma = \{\gamma(s,t), s,t \in [0,1]\}$. For a given functional sample $X_1(t), \ldots, X_n(t), t \in [0,1]$, consider the model:

$$X_k(t) = g(k/n, t) + Y_k(t), \quad t \in [0,1], \quad k = 1, \ldots, n, \tag{1}$$

where the function $g : [0,1] \times [0,1] \to \mathbb{R}$ is deterministic, but unobserved. Our main aim is in testing the hypothesis:

$$H_0 : \quad g = 0 \;\text{ versus }\; H_1 : g \neq 0$$

with emphasis on a case of change-point detection, which corresponds to a piecewise-constant function g with respect to the first argument.

This model covers a broad range of real-world problems such as climate change detection, image analysis, analysis of medical treatments, especially magnetic resonance images of brain activities, and speech recognition, to name a few. Besides, the change-point detection model (1) can be used for knot selection in spline smoothing as well for trend changes in functional time series analysis.

There is a huge list of references on testing for change-points or structural changes for a sequence of independent random variables/vectors. We refer to Csörgő and Horváth [1], Brodsky and Darkhovsky [2], Basseville and Nikiforov [3], and Chen and Gupta [4] for accounts of various techniques.

Within the functional data analysis literature, change-point detection has largely focused on one change-point problem. In Berkes et al. [5], a cumulative sum (CUSUM) test was proposed for independent functional data by using projections of the sample onto some principal components of covariance γ. Later, the problem was studied in Aue et al. [6], where its asymptotic properties were developed. This test was extended to weakly dependent functional data and epidemic changes by Aston and Kirch [7]. Aue et al. [8] proposed a fully functional method for finding a change in the mean without losing information due to dimension reduction, T. Harris, Bo Li, and J. D. Tucker [9] propose the multiple change-point isolation method for detecting multiple changes in the mean and covariance of a functional process.

The methodology we propose is based on some measures of variation of the process:

$$W_n(s) = \sum_{k=1}^{\lfloor ns \rfloor}(X_k - \overline{X}_n) + (ns - \lfloor ns \rfloor)(X_{\lfloor ns \rfloor + 1} - \overline{X}_n), \ \ s \in [0,1],$$

where $\overline{X}_n = n^{-1}(X_1 + \cdots + X_n)$.

Since this process is infinite-dimensional, we used the projections technique to reduce the dimension. To this aim, we assumed that Y_i is mean-squared continuous and jointly measurable and that γ has finite trace: $\mathrm{tr}(\gamma) = \int_0^1 \gamma(t,t)\, \mathrm{d}t < \infty$. In this case, Y_i is also an $L_2(0,1)$-valued random element, where $L_2 := L_2(0,1)$ is a Hilbert space of Lebesgue square integrable functions on $[0,1]$ endowed with the inner product $\langle f, g \rangle = \int_0^1 f(t)g(t)\, \mathrm{d}t$ and the norm $\|f\| := \sqrt{\langle f, f \rangle}$.

In the case where the number of change-points is known to be no bigger than m, our test statistics are constructed from (m,p)-variation (see the definition below) of the processes $(\langle W_n(s), \psi \rangle, s \in [0,1])$, where $\psi \in \Psi \subset L_2(0,1)$ runs through a finite set Ψ of possibly random directions in $L_2(0,1)$. In particular, Ψ consists of estimated principal components. If the number of change-points is unknown, we consider the p-variation of the processes $(\langle W_n(s), \psi \rangle, s \in [0,1]), \psi \in \Psi$ and estimate the possible number of change-points.

The paper is organized as follows. In Section 2, $\mathcal{G}$-sum and $\mathcal{G}$-CUSUM processes are defined and their asymptotic behavior is considered in a framework of the $\ell^\infty(\mathcal{G})$ space. The results presented in this section are used to derive the asymptotic distributions of the test statistics presented in Section 3. Section 4 is devoted to simulation studies of the proposed test algorithms. Section 5 contains a case study. Finally, Section 6 is devoted to the proofs of our main theoretical results.

2. $\mathcal{G}$-Sum Process and Its Asymptotic

Let $\mathcal{Q}$ be the set of all probability measures on $([0,1], \mathcal{B}_{[0,1]})$. For any $Q \in \mathcal{Q}$ and Q-integrable function f, $Qf := \int_0^1 f\, \mathrm{d}Q$. As usual, $L_2([0,1], Q)$ is a set of measurable functions on $[0,1]$, which are square-integrable for the measure Q, and $L_2([0,1], Q)$ is an associated Hilbert space endowed with the inner product:

$$\langle f, g \rangle_Q = \int_0^1 f(t)g(t)Q(dt), \ \ f, g \in L_2([0,1], Q)$$

and corresponding distance $\rho_Q(f,g)$, $f, g \in L_2([0,1], Q)$. We abbreviate $L_2([0,1], \lambda)$ to L_2 and $\langle \cdot, \cdot \rangle_\lambda$ to $\langle \cdot, \cdot \rangle$ for Lebesgue measure λ. We use the norm $\|f\| := \sqrt{\langle f, f \rangle}$ and the distance $\rho(f,g) = \|f - g\|$ for the elements $f, g \in L_2$. On the set $L_2 \times L_2$, we use the inner product:

$$\langle (f,g), (f', g') \rangle_2 = \langle f, f' \rangle + \langle g, g' \rangle$$

and the corresponding distance:

$$\rho_2((f,g),(f',g')) = \left(\|f - f'\|^2 + \|g - g'\|^2 \right)^{1/2}, \ \ f, f', g, g' \in L_2.$$

For two given sets $\mathcal{F}, \Psi \subset L_2$, we consider the $\mathcal{F} \times \Psi$-sum process:

$$\nu_n = \Big(\sum_{k=1}^{n} \nu_{nk}(f, \psi), f \in \mathcal{F}, \psi \in \Psi \Big),$$

where $\nu_{nk}(f,\psi) = \langle X_k, \psi\rangle \lambda_{nk}(f)$, λ_{nk} is a uniform probability on the interval $[(k-1)/n, k/n]$ and $\lambda_{nk}(f) = \int_0^1 f(t)\,\mathrm{d}\lambda_{nk}(t)$. A natural framework for stochastic process ν_n is the space $\ell^\infty(\mathcal{G})$, where $\mathcal{G} = \mathcal{F} \times \Psi$. Recall for a class $\mathcal{G}$ that $\ell^\infty(\mathcal{G})$ is a Banach space of all uniformly bounded real-valued functions μ on $\mathcal{G}$ endowed with the uniform norm:

$$\|\mu\|_{\mathcal{G}} := \sup\{|\mu(g)| : g \in \mathcal{G}\}.$$

Given a pseudometric d on $\mathcal{G}$, $UC(\mathcal{G}, d)$ is a set of all $\mu \in \ell^\infty(\mathcal{G})$, which are uniformly d-continuous. The set $UC(\mathcal{G}, d)$ is a separable subspace of $\ell^\infty(\mathcal{G})$ if and only if $(\mathcal{G}, d)$ is totally bounded. The pseudometric space $(\mathcal{G}, d)$ is totally bounded if $N(\varepsilon, \mathcal{G}, d)$ is finite for every $\varepsilon > 0$, where $N(\varepsilon, \mathcal{G}, d)$ is the minimal number of open balls of d-radius ε, which are necessary to cover $\mathcal{G}$.

It is worth noting that the process ν_n is continuous when $\mathcal{F} \times \Psi$ is endowed with the metric ρ_2. Indeed,

$$\begin{aligned}
|\nu_{nk}(f,\psi) - \nu_{nk}(f',\psi')| &\le |\langle Y_k, \psi\rangle\lambda_{nk}(f) - \langle Y_k, \psi'\rangle\lambda_{nk}(f')| \\
&\le |\lambda_{nk}(f)|\langle Y_k, \psi - \psi'\rangle + \langle Y_k, \psi'\rangle\lambda_{nk}(f - f')| \\
&\le \|Y_k\|[\sqrt{n}\|f\| \cdot \|\psi - \psi'\| + \|\psi'\| \cdot \|f - f'\|] \\
&\le \sqrt{2}\|Y_k\| \max\{\sqrt{n}\|f\|, \|\psi\|\}\rho_2((f,\psi),(f',\psi')),
\end{aligned}$$

since $|\lambda_{nk}(f)| \le \sqrt{n}\|f\|$ for every $f \in L_2$. If both sets $\mathcal{F}$ and Ψ are totally bounded, then the process ν_n is uniformly continuous so that ν_n takes values in the subspace $UC(\mathcal{G})$.

Next, we specify the set $\mathcal{F} \subset L_2$. To this aim, we recall some definitions. For a function $f : [0,1] \to \mathbb{R}$, a positive number $0 < p < \infty$, and an integer $m \in \mathbb{N}$, the (m,p)-variation of f on the interval $[0,t]$ is

$$v_{m,p}(f;[0,t]) := \sup\Big\{\sum_{j=1}^{m} |f(t_j) - f(t_{j-1})|^p\Big\},$$

where the supremum is taken over all partitions $0 = t_0 < t_1 < \cdots < t_m = t$, of the interval $[0,t]$. We abbreviate $v_{m,p}(f) := v_{m,p}(f;[0,1])$. If $v_p(f) := \sup_{m\ge 1} v_{m,p}(f) < \infty$, then we say that f has finite p-variation and $\mathcal{W}_p[0,1]$ is the set of all such functions. The set $\mathcal{W}_p[0,1]$, $p \ge 1$, is a (non-separable) Banach space with the norm:

$$||f||_{[p]} := \sup_{0\le t\le 1} |f(t)| + v_p^{1/p}(f).$$

The embedding $\mathcal{W}_p[0,1] \hookrightarrow \mathcal{W}_q[0,1]$ is continuous and

$$v_q^{1/q}(f) \le v_p^{1/p}(f), \quad \text{for } 1 \le p < q.$$

For more information on the space $\mathcal{W}_p[0,1]$, we refer to [10].

The limiting zero mean Gaussian process $\nu_\gamma = (\nu(f,\psi), f \in \mathcal{F}, \psi \in \Psi)$ is defined via covariance:

$$E\nu_\gamma(f,\psi)\nu_\gamma(f',\psi') = \mathcal{K}_\gamma((f,\psi),(f',\psi,)) := \langle\Gamma\psi,\psi'\rangle\langle f, f'\rangle, \quad \psi,\psi', f, f' \in L_2, \tag{2}$$

where $\Gamma : L_2 \to L_2$ is the covariance operator corresponding to the kernel γ. The function $\mathcal{K}_\gamma : \mathcal{G} \times \mathcal{G} \to \mathbb{R}$ is positive definite:

$$\sum_{k,j=1}^{m} c_j c_k \mathcal{K}_\gamma((f_j,\psi_j),(f_k,\psi_k)) \ge 0, \tag{3}$$

for all $c_1, \dots, c_m \in \mathbb{R}$, $(f_1,\psi_1), \dots, (f_m,\psi_m) \in \mathcal{G}$, and $m \ge 1$. Indeed, if we denote by $\mathcal{W} = (\mathcal{W}(f), f \in L_2)$ the isonormal Gaussian process on the Hilbert space L_2, we see that

$$\mathcal{K}_\gamma((f_j, \psi_j), (f_k, \psi_k)) = E\langle Y, \psi_j\rangle\langle Y, \psi_k\rangle E\mathcal{W}(f_j)\mathcal{W}(f_k);$$

hence,

$$\sum_{k,j=1}^m c_j c_k \mathcal{K}_\gamma((f_j, \psi_j), (f_k, \psi_k)) = E\Big(\sum_{k=1}^m c_k\langle Y, \psi_k\rangle\mathcal{W}(f_k)\Big)^2$$

and (3) follows. This justifies the existence of the process ν_γ.

Throughout, we shall exploit the following.

Assumption 1. *Random processes $Y, Y_1, Y_2, \dots$ are i.i.d. mean square continuous, jointly measurable, with mean zero and covariance γ such that $\int_0^1 \gamma(t,t)\,\mathrm{d}t < \infty$.*

For the model (1), we consider null hypothesis $H_0 : \ g = 0$ and two possible alternatives:

$$H_A : \ g = g_n = u_n q_n, \ \text{where} \ u_n \to u \ \text{in} \ \mathcal{W}_2[0,1], \ \sqrt{n}q_n \to q \ \text{in} \ L_2,$$

and

$$H'_A : \ g = g_n = u_n q_n, \ \text{where} \ u_n \to u \ \text{in} \ \mathcal{W}_2[0,1], \ \sqrt{n}\sup_{\psi\in\Psi}|\langle q_n, \psi\rangle| \to \infty.$$

In both alternatives, the function u_n is responsible for the configuration of a drift within the sample, whereas the function q_n estimates a magnitude of the drift.

Our main theoretical results are Theorems 1 and 3, which are proven in Section 6.

Theorem 1. *Let the random processes (X_k) be defined by (1), where $Y, Y_1, Y_2, \dots$ satisfy Assumption 1. Assume that, for some $1 \le q < 2$, the set $\mathcal{F} \subset \mathcal{W}_q[0,1]$ is bounded and the set $\Psi \subset L_2$ satisfies*

$$\int_0^1 \sqrt{\log N(\varepsilon, \Psi, \rho)}\,\mathrm{d}\varepsilon < \infty. \tag{4}$$

Then, there exists a version of a Gaussian process ν_γ on $L_2 \times L_2$ such that its restriction on $\mathcal{F} \times \Psi$, $(\nu_\gamma(f, \psi), f \in \mathcal{F}, \psi \in \Psi)$ is continuous and the following hold:

(1a) Under H_0:

$$n^{-1/2}\nu_n \xrightarrow[n\to\infty]{\mathcal{D}} \nu_\gamma \ \text{in} \ \ell^\infty(\mathcal{F} \times \Psi). \tag{5}$$

(1b) Under H_A,

$$n^{-1/2}\nu_n \xrightarrow[n\to\infty]{\mathcal{D}} \nu_\gamma + \Delta, \ \text{in} \ \ell^\infty(\mathcal{F} \times \Psi), \tag{6}$$

where

$$\Delta(f, \psi) = \langle u, f\rangle\langle q, \psi\rangle.$$

If $u(s) = 1, s \in [0,1]$, then the alternative H_A corresponds to the presence of a signal in a noise. In this case, $\Delta(f, \psi) = \lambda(f)\langle q, \psi\rangle$. Therefore, the use of this theorem for testing a signal in a noise is meaningful provided $\langle q, \psi\rangle \ne 0$.

As a corollary, Theorem 1 combined with the continuous mapping theorem gives the following result.

Theorem 2. *Assume that conditions of Theorem 1 are satisfied. Then, the following hold:*

(2a) Under H_0

$$\sup_{\psi\in\Psi, f\in\mathcal{F}} |n^{-1/2}\nu_n(f, \psi)| \xrightarrow[n\to\infty]{\mathcal{D}} \sup_{\psi\in\Psi, f\in\mathcal{F}} |\nu_\gamma(f, \psi)|.$$

(2b) Under H_A,

$$\sup_{\psi\in\Psi, f\in\mathcal{F}} |n^{-1/2}\nu_n(f, \psi)| \xrightarrow[n\to\infty]{\mathcal{D}} \sup_{\psi\in\Psi, f\in\mathcal{F}} |\nu_\gamma(f, \psi) + \langle u, f\rangle\langle q, \psi\rangle|.$$

$(2c)$ *Under* H_A',

$$\sup_{\psi\in\Psi, f\in\mathcal{F}} |n^{-1/2}\nu_n(f,\psi)| \xrightarrow[n\to\infty]{\mathrm{P}} \infty. \tag{7}$$

Proof. Since both (2a) and (2b) are by-products of Theorem 1 and continuous mappings, we need to prove only (2c). First, we observe that

$$\begin{aligned}\sup_{\psi\in\Psi, f\in\mathcal{F}} |\nu_n(f,\psi)| &\geq \sup_{\psi\in\Psi, f\in\mathcal{F}} \Big| \sum_{k=1}^n \Big(\langle Y_k,\psi\rangle + \langle q_n,\psi\rangle u_n(k/n)\Big)\lambda_{nk}(f)\Big| \\ &\geq \sup_{\psi\in\Psi, f\in\mathcal{F}} \Big| \sum_{k=1}^n u_n(k/n)\lambda_{nk}(f)\Big| \cdot |\langle q_n,\psi\rangle| - O_P(\sqrt{n}),\end{aligned}$$

by (2a). Consider

$$I_n(f) := \Big| \sum_{k=1}^n u_n(k/n)\lambda_{nk}(f)\Big|.$$

We have

$$\begin{aligned} I_n(f) &= n\Big| \sum_{k=1}^n u_n(k/n) \int_{(k-1)/n}^{k/n} f(t)\,\mathrm{d}t\Big| \\ &\geq n\Big| \sum_{k=1}^n \int_{(k-1)/n}^{k/n} u_n(t)f(t)\,\mathrm{d}t\Big| - n\Big| \sum_{k=1}^n \int_{(k-1)/n}^{k/n} (u_n(t)-u_n(k/n))f(t)\,\mathrm{d}t\Big| \\ &:= I_n'(f) - I_n''(f).\end{aligned}$$

By the Hölder inequality,

$$\begin{aligned} I_n''(f) &\leq n\sum_{k=1}^n \Big(\int_{(k-1)/n}^{k/n} (u_n(t)-u_n(k/n))^2\,\mathrm{d}t\Big)^{1/2}\Big(\int_{(k-1)/n}^{k/n} f^2(t)\,\mathrm{d}t\Big)^{1/2} \\ &\leq n\Big(\sum_{k=1}^n \int_{(k-1)/n}^{k/n} (u_n(t)-u_n(k/n))^2\,\mathrm{d}t\Big)^{1/2}\Big(\sum_{k=1}^n\int_{(k-1)/n}^{k/n} f^2(t)\,\mathrm{d}t\Big)^{1/2} \\ &\leq n\Big(n^{-1}\sum_{k=1}^n v_2(u_n,[(k-1)/n,k/n])\Big)^{1/2}\|f\| \leq \sqrt{n}v_2^{1/2}(u_n)\|f\|.\end{aligned}$$

Since $I_n'(f) = n|\langle u_n, f\rangle|$, we deduce

$$I_n(\psi,f) \geq n|\langle u_n,f\rangle| - \sqrt{n}v_2^{1/2}(u_n)\|f\|.$$

Hence,

$$n^{-1/2}\sup_{\psi\in\Psi, f\in\mathcal{F}} |\nu_n(f,\psi)| \geq \sqrt{n}\sup_{\psi\in\Psi, f\in\mathcal{F}} |\langle u_n,f\rangle|\cdot|\langle q_n,\psi\rangle| - O_P(1)$$

and this completes the proof of (2c). □

Next, we consider $\mathcal{G}$-sum process $\mu_n = (\mu_n(f,\psi), f\in\mathcal{F}, \psi\in\Psi)$ defined by

$$\mu_n(f,\psi) = \sum_{k=1}^n \langle X_k - \overline{X}_n, \psi\rangle\lambda_{nk}(f),$$

where $\overline{X}_n = n^{-1}(X_1+\cdots+X_n)$. Its limiting zero mean Gaussian process μ_γ is defined via covariance:

$$E\mu_\gamma(f,\psi)\mu_\gamma(f',\psi') = \langle\Gamma\psi,\psi'\rangle[\langle f,f'\rangle - \lambda(f)\lambda(f')],\ \ \psi,\psi',f,f'\in L_2. \tag{8}$$

The existence of Gaussian process μ_γ can be justified as that of ν_γ above. Just notice that

$$\langle f, f' \rangle - \lambda(f)\lambda(f') - E(\mathcal{W}(f) - \lambda(f)\mathcal{W}(1))(\mathcal{W}(f') - \lambda(f')\mathcal{W}(1)),$$

where $1(t) = 1, t \in [0,1]$.

Theorem 3. *Assume that the conditions of Theorem 1 are satisfied. Then, there exists a version of the Gaussian process μ_γ on $L_2(0,1) \times L_2(0,1)$ such that its restriction on $\mathcal{F} \times \Psi$, $(\nu(f,\psi), f \in \mathcal{F}, \psi \in \Psi)$ is continuous and the following hold:*

(3a) *Under H_0,*

$$n^{-1/2}\mu_n \xrightarrow[n\to\infty]{\mathcal{D}} \mu_\gamma \;\; in \;\; \ell^\infty(\mathcal{F} \times \Psi); \tag{9}$$

(3b) *Under alternative H_A,*

$$n^{-1/2}\mu_n \xrightarrow[n\to\infty]{\mathcal{D}} \mu_\gamma + \widetilde{\Delta} \; in \;\; \ell^\infty(\mathcal{F} \times \Psi), \tag{10}$$

where

$$\widetilde{\Delta}(f,\psi) = [\langle u, f \rangle - \lambda(u)\lambda(f)]\langle q, \psi \rangle.$$

We see that the limit distribution of the $\mathcal{G}$-sum process separates the null and alternative hypothesis provided $[\langle u, f \rangle - \lambda(u)\lambda(f)]\langle q, \psi \rangle \neq 0$. As a corollary, Theorem 3 combined with the continuous mapping theorem gives the following results.

Theorem 4. *Assume that the conditions of Theorem 1 are satisfied. Then, the following hold:*

(4a) *Under H_0,*

$$\sup_{\psi\in\Psi, f\in\mathcal{F}} |n^{-1/2}\mu_n(f,\psi)| \xrightarrow[n\to\infty]{\mathcal{D}} \sup_{\psi\in\Psi, f\in\mathcal{F}} |\mu_\gamma(f,\psi)|. \tag{11}$$

(4b) *Under H_A,*

$$\sup_{\psi\in\Psi, f\in\mathcal{F}} |n^{-1/2}\mu_n(f,\psi)| \xrightarrow[n\to\infty]{\mathcal{D}} \sup_{\psi\in\Psi, f\in\mathcal{F}} |\mu_\gamma(f,\psi) + \widetilde{\Delta}(f,\psi)|. \tag{12}$$

(4*c*) *Under H'_A,*

$$\sup_{\psi\in\Psi, f\in\mathcal{F}} |n^{-1/2}\mu_n(f,\psi)| \xrightarrow[n\to\infty]{\mathrm{P}} \infty. \tag{13}$$

Proof. Both (4a) and (4b) are by-products of Theorem 3 and continuous mappings, whereas the proof of (4c) follows the lines of the proof of Theorem 2 (2c). □

3. Test Statistics

Several useful test statistics can be obtained from the $\mathcal{G}$-sum process $\mu_n = (\mu_n(f,\psi), (f,\psi) \in \mathcal{G} = \mathcal{F} \times \Psi)$, by considering concrete examples of sets Ψ and $\mathcal{F}$.

Throughout this section, we assume that the sample $X_1, X_2, \ldots, X_n$ follows the model (1) and $Y, Y_1, Y_2, \ldots$ satisfies Assumption 1.

By Γ, we denote the covariance operator of Y: $\Gamma = E(Y \otimes Y)$. Recall

$$\Gamma x(t) = \int_0^1 \gamma(t,s)x(s)\,\mathrm{d}s, \;\; x \in [0,1].$$

According to Mercer's theorem, the covariance γ has then the following singular-value decomposition:

$$\gamma(s,t) = \sum_{r=1}^{m} \lambda_r \psi_r(s)\psi_r(t), \;\; t,s \in [0,1], \tag{14}$$

where $\lambda_1, \ldots, \lambda_m$ are all the decreasingly ordered positive eigenvalues of Γ and $\psi_1, \ldots, \psi_m$ are the associated eigenfunctions of Γ such that

$$\int_0^1 \psi_r^2(t)dt = 1, \ \int_0^1 \psi_r(t)\psi_\ell(t)dt = 0, \ r \neq \ell,$$

and m is the smallest integer such that, when $r > m$, $\lambda_r = 0$. If $m = \infty$, then all the eigenvalues are positive, and in this case, $\sum_r \lambda_r < \infty$. Note that $\lambda_r = E\langle Y, \psi_r \rangle^2$. Besides, we shall assume the following.

Assumption 2. *The eigenvalues λ_r satisfy, for some $d > 0$,*

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

In statistical analysis, the eigenvalues and eigenfunctions of Γ are replaced by their estimated versions. Noting that, for each k,

$$E[(X_k - E(X_k)) \otimes (X_k - E(X_k))] = \Gamma,$$

one estimates Γ by

$$\widehat{\Gamma}_n := \frac{1}{n}\sum_{i=1}^{n}[(X_i - \overline{X}_n) \otimes (X_i - \overline{X}_n)],$$

where $\overline{X}_n(s) = n^{-1}(X_1(s) + \cdots + X_n(s))$. We denote the eigenvalues and eigenfunctions of $\widehat{\Gamma}$ by $\widehat{\lambda}_{nr}$ and $\widehat{\psi}_{nr}$, $r = 1, \ldots, n-1$, respectively. In order to ensure that $\widehat{\psi}_{nr}$ may be viewed as an estimator of ψ_r rather than of $-\psi_r$, we will in the following assume that the signs are such that $\langle \widehat{\psi}_{nr}, \psi_r \rangle \geq 0$. Note that

$$\widehat{\Gamma}\widehat{\psi}_{nr} = \widehat{\lambda}_{nr}\widehat{\psi}_{nr}, \ \ r = 1, \ldots, n-1, \tag{15}$$

and

$$\widehat{\lambda}_{nr} = \frac{1}{n-1}\sum_{i=1}^{n}\langle X_i - \overline{X}_n, \widehat{\psi}_{nr} \rangle^2, \ \ r = 1, \ldots, n. \tag{16}$$

The use of the estimated eigenfunctions and eigenvalues in the test statistics is justified by the following result. For a Hilbert–Schmidt operator T on L_2, we denote by $\|T\|_{HS}$ its Hilbert–Schmidt norm.

Lemma 1. *Assume that Assumption 1 holds. Then, under H_A,*

$$\|\widehat{\Gamma}_n - \Gamma\|_{HS} \to 0 \ \text{ as } \ n \to \infty.$$

Proof. First, we observe that

$$\widehat{\Gamma}_n = \widetilde{\Gamma}_n + T_{n1} + T_{n2} + T_{n3},$$

where

$$\widetilde{\Gamma}_n = \frac{1}{n}\sum_{k=1}^{n}(Y_k - \overline{Y}_n) \otimes (Y_k - \overline{Y}_n),$$

$$T_{n1} = \frac{1}{n}\sum_{k=1}^{n}\left[u_n(k/n) - \frac{1}{n}\sum_{j=1}^{n}u_n(j/n)\right](Y_k - \overline{Y}_n) \otimes q_n,$$

$$T_{n2} = \frac{1}{n}\sum_{k=1}^{n}\left[u_n(k/n) - \frac{1}{n}\sum_{j=1}^{n}u_n(j/n)\right]q_n \otimes (Y_k - \overline{Y}_n),$$

$$T_{n3} = \frac{1}{n}\sum_{k=1}^{n}\left[u_n(k/n) - \frac{1}{n}\sum_{j=1}^{n}u_n(j/n)\right]^2 q_n \otimes q_n.$$

It is well known that $\|\widetilde{\Gamma}_n - \Gamma\|_{HS} \xrightarrow[n\to\infty]{\text{a.s.}} 0$ as $n \to \infty$. By the moment inequality for sums of independent random variables, we deduce

$$E\|T_{ni}\|_{HS}^2 \le cn^{-2} \sum_{k=1}^{n} \Big[u_n(k/n) - \frac{1}{n}\sum_{j=1}^{n} u_n(j/n)\Big]^2 E\|Y\|^2 \|q_n\|^2,$$

for both $i = 1, 2$. This yields $T_{ni} \xrightarrow[n\to\infty]{\text{P}} 0$. Next, we have

$$\|T_{n3}\|_{HS} = \frac{1}{n}\sum_{k=1}^{n} \Big[u_n(k/n) - \frac{1}{n}\sum_{j=1}^{n} u_n(j/n)\Big]^2 \|q_n\|^2 \le \frac{1}{n}\sum_{k=1}^{n} u_n^2(k/n)\|q_n\|^2 \to 0$$

as $n \to \infty$ due to assumption H_A. This completes the proof. □

Lemma 2. *Assume that Assumptions 1 and 2 for some finite d hold and* $E(\|Y\|^4) < \infty$. *Then, under* H_0, *as well as under* H_A*:*

$$n^{1/2}|\widehat{\lambda}_{nj} - \lambda_j| = O_P(1), \ \text{ and } \ n^{1/2}\|\widehat{c}_j\widehat{\psi}_{nj} - \psi_j\| = O_P(1)$$

for each $1 \le j \le d$, *where* $\widehat{c}_{nj} = \langle \widehat{\psi}_{nj}, \psi_j \rangle$.

Proof. If the null hypothesis is satisfied, then $\widehat{\Gamma}_n = \widetilde{\Gamma}_n$ and the asymptotic results for the eigenvalues and eigenfunctions of $\widetilde{T}_n$ are well known (see, e.g., [11]). Under alternative H_A, the results follow from Lemma 1 and Lemmas 2.2 and 2.3 in [11]. □

Next, we consider separately the test statistics for at most one, at most m, and for an unknown number of change-points.

3.1. Testing at Most One Change-Point

Define for $d > 0$,

$$T_{n,1}(d) := \max_{1\le j\le d} \frac{1}{\sqrt{\lambda_j}} \max_{1\le k\le n} \Big| \sum_{i=1}^{k} \langle X_i - \overline{X}_n, \psi_j \rangle \Big|. \tag{17}$$

This statistic is designed for at most one change-point alternative. Its limiting distribution is established in the following theorem.

Theorem 5. *Let random functional sample* (X_k) *be defined by* (1) *where* $Y, Y_1, Y_2, \ldots$ *satisfies Assumptions 1 and 2. Then,*

(a) Under H_0, *it holds that*

$$n^{-1/2}T_{n,1}(d) \xrightarrow[n\to\infty]{\mathcal{D}} \max_{1\le k\le d}\max_{0\le t\le 1} |B_k(t)|,$$

where $B_1, \ldots, B_d$ *are independent standard Brownian bridge processes;*

(b) Under H_A, *it holds that*

$$n^{-1/2}T_{n,1}(d) \xrightarrow[n\to\infty]{\mathcal{D}} \max_{1\le k\le d}\max_{0\le t\le 1} |B_k(t) + \Delta(t)\langle q, \psi_k/\sqrt{\lambda_k}\rangle|,$$

where

$$\Delta(t) = \int_0^t u(s)\,\mathrm{d}s - t\int_0^1 u(s)\,\mathrm{d}s, \ \ t \in [0,1]. \tag{18}$$

(c) Under H_A', *it holds that*

$$n^{-1/2}T_{n,1}(d) \xrightarrow[n\to\infty]{\text{P}} \infty.$$

Proof. Consider the sets

$$\Psi_{d,\gamma} := \left\{ \frac{\psi_1}{\sqrt{\lambda_1}}, \ldots, \frac{\psi_d}{\sqrt{\lambda_d}} \right\}, \text{ and } \mathcal{F}_1 = \{\mathbf{1}_{[0,t]}, t \in [0,1]\}. \tag{19}$$

Observing that

$$T_{n,1}(d) = \sup_{\psi \in \Psi_{d,\gamma}, f \in \mathcal{F}_1} |\mu_n(f, \psi)|$$

and $\mathcal{F}_1$ is a bounded set in $\mathcal{W}_q$, we complete the proof by applying Theorem 3. □

Based on this result, we construct the testing procedure in a classical way. Choose for a given $\alpha \in (0,1)$, $C_\alpha > 0$ such that

$$P(\max_{1\le k\le d} \max_{0\le t\le 1} |B_k(t)| > C_\alpha) = \alpha.$$

According to Theorem 5, the test:

$$T_{n,1}(d) \ge \sqrt{n} C_\alpha \tag{20}$$

will have asymptotic level α. Under the alternative H_A, we have

$$\begin{aligned} \lim_{n\to\infty} P(n^{-1/2} T_{n,1}(d) \ge C_\alpha) &\ge P\Big(\max_{1\le k\le d} \max_{0\le t\le 1} |B_k(t)| \le \max_{1\le k\le d} \max_{0\le t\le 1} |\Delta(t)\langle q, \psi_k/\lambda_k\rangle - C_\alpha\Big) \\ &\ge 1 - \alpha \end{aligned}$$

when

$$\max_{1\le k\le d} \max_{0\le t\le 1} |\Delta(t)\langle q, \psi_k/\lambda_k\rangle| \ge 2C_\alpha. \tag{21}$$

Hence, if $g(s,t) = g_n(s,t) = u_n(s) q_n(t)$ and

$$\sqrt{n} \max_{1\le k\le d} \max_{0\le t\le 1} |\langle g_n(t,\cdot), \psi_k/\sqrt{\lambda_k}\rangle| \to \infty$$

as $n \to \infty$, then the test (20) is asymptotically consistent.

Let us note that, due to the independence of Brownian bridges $B_k, k = 1, \ldots, d$, we have

$$1 - \alpha = P(\max_{1\le k\le d} \max_{0\le t\le 1} |B_k(t)| \le C_\alpha) = P^d(\max_{0\le t\le 1} |B_1(t)| \le C_\alpha).$$

This yields

$$P(\max_{0\le t\le 1} |B_1(t)| \le C_\alpha) = (1-\alpha)^{1/d}.$$

Hence, C_α is the $(1-\alpha)^{1/d}$-quantile of the distribution of $\sup_{0\le t\le 1} |B_1(t)|$. This observation simplifies the calculations of critical values C_α.

In particular, if there is $s^* \in (0,1)$ such that $u(s) = \mathbf{1}_{[0,s^*]}(s), s \in [0,1]$, then we have one change-point model:

$$X_k(t) = \mathbf{1}_{[0,s^*]}(k/n) q_n(t) + Y_k(t), \quad t \in [0,1].$$

In this case, $\Delta(t) = \Delta^*(t) := \min\{t, s^*\} - ts^*, \quad t \in [0,1]$.

Figure 1 below shows generated density functions of $\max_{1\le k\le d} \max_{0\le t\le 1} |B_k(t)|$ and $\max_{1\le k\le d} \max_{0\le t\le 1} |B_k(t) + \Delta^*(t)\langle q, \psi_k/\sqrt{\lambda_k}\rangle|$ for $d = [1, 10, 30]$, $s^* \in \{1/4, 1/2, 3/4\}$ where $q = a\psi_k\sqrt{\lambda_k}$ for a fixed k.

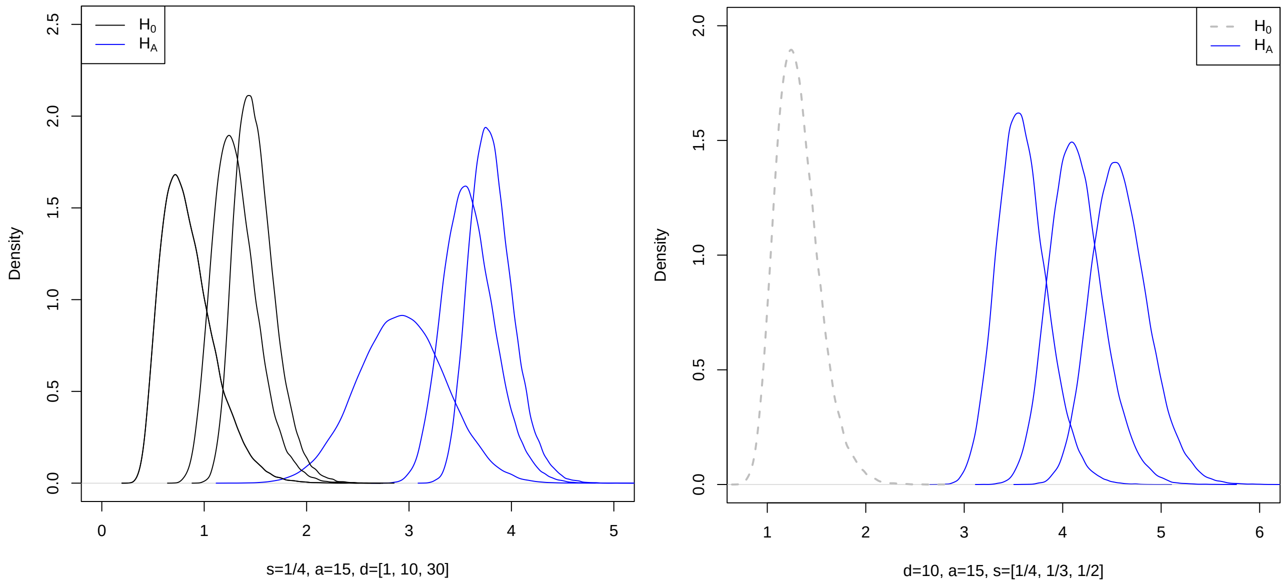


Figure 1. Density functions.

Let us observe that test statistic $T_{n,1}(d)$ tends to infinity when $d \to \infty$. On the other hand, with larger d, the approximation of X_j by series $\sum_{k=1}^{d} \langle X, \psi_j \rangle \psi_j$ is better and leads to better testing power. The following result establishes the asymptotic distribution of $T_{n,1}(d)$ as $d \to \infty$.

Theorem 6. *Let random functional sample (X_k) be defined by (1) where $Y, Y_1, Y_2, \ldots$ satisfies Assumption 1. Then, under H_0,*

$$\lim_{d\to\infty} \lim_{n\to\infty} P\Big(n^{-1/2} T_{n,1}(d) \le \frac{x}{a_d} + b_d\Big) = \exp\{-e^{-x}\}, \quad x \ge 0, \tag{22}$$

where

$$a_d = (8 \ln d)^{1/2}, \quad b_d = \frac{1}{4} a_d + \frac{\ln \ln d}{a_d}. \tag{23}$$

Proof. By Theorem 5, the proof reduces to

$$\lim_{d\to\infty} P(\max_{1\le j\le d} \|B_j\|_\infty \le x/a_d + b_d) = \exp\{-e^{-x}\}, \quad x \ge 0. \tag{24}$$

It is known that

$$P(\|B_j\|_\infty > u) = 2e^{-2u^2} u^2 (1 + o(1)), \quad u \to \infty.$$

Since Brownian bridges $B_j, 1 \le j \le d$ are independent, we have

$$\begin{aligned} P(\max_{1\le j\le d} \|B_j\|_\infty \le x/a_d + b_d) &= P^d(\|B_1\|_0 \le x/a_d + b_d) \\ &= \Big(1 - P(\|B_1\|_0 \ge x/a_d + b_d)\Big)^d \end{aligned}$$

and

$$\lim_{d\to\infty} dP(\|B_1\|_\infty \ge x/a_d + b_d) = e^{-x}.$$

This proves (24). □

When d is large, the test (20) becomes

$$T_{n,1} \ge \sqrt{n}\Big[\frac{1}{a_d} \ln\Big(\frac{1}{\ln(1/\alpha)}\Big) + b_d\Big] \tag{25}$$

and has asymptotic level α as n and d tend to infinity.

The dependence on d of critical values of the tests (20) and (25) is shown in Figure 2. A comparison was made for asymptotic level $\alpha = 0.05$. From Figure 2, we see that the critical values in (25) are smaller than those in (20). This means that the error of the first kind is more likely with the test (25), rather than with (36). This is confirmed by simulations.

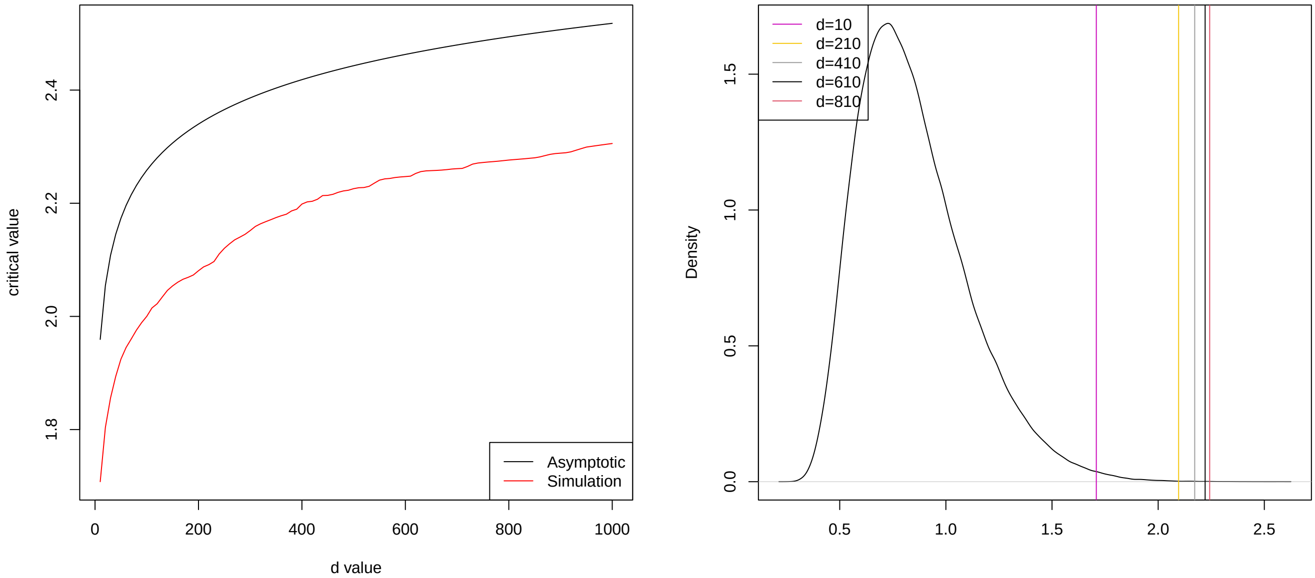


Figure 2. Comparison of the critical values in (20) and (25) with $\alpha = 0.05$ and the density function of $T_{n,1}(d)$.

If the eigenfunctions (ψ_k) are unknown, we use the statistics:

$$\widehat{T}_{n,1}(d) := \max_{1\le j\le d} \frac{1}{\sqrt{\widehat{\lambda}_j}} \max_{1\le k\le n} \Big| \sum_{i=1}^{k} \langle X_i - \overline{X}_n, \widehat{\psi}_j \rangle \Big|. \tag{26}$$

Theorem 7. *Let random functional sample (X_k) be defined by (1), where $Y, Y_1, Y_2, \ldots$ satisfies Assumptions 1 and 2. Then:*

(a) *Under H_0,*

$$n^{-1/2}\widehat{T}_{n,1}(d) \xrightarrow[n\to\infty]{\mathcal{D}} \max_{1\le k\le d} \max_{0\le t\le 1} |B_k(t)|,$$

where $B_1, \ldots, B_d$ are independent standard Brownian bridge processes;

(b) *Under H_A, if $\mathrm{E}\|Y\|^4 < \infty$, it holds that*

$$n^{-1/2}\widehat{T}_{n,1}(d) \xrightarrow[n\to\infty]{\mathcal{D}} \max_{1\le k\le d} \max_{0\le t\le 1} |B_k(t) + \Delta(t)\langle q, \psi_k/\sqrt{\lambda_k}\rangle|,$$

where $\Delta(t) = \int_0^t u(s)\,\mathrm{d}s - t\int_0^1 u(s)\,\mathrm{d}s, \quad t \in [0,1]$.

(c) *Under H_A', if $\mathrm{E}\|Y\|^4 < \infty$, it holds that*

$$n^{-1/2}\widehat{T}_{n,1}(d) \xrightarrow[n\to\infty]{\mathrm{P}} \infty.$$

Proof. The result follows from Theorem 5 if we show that

$$D_n := n^{-1/2}|T_{n,1}(d) - \widehat{T}_{n,1}(d)| \xrightarrow[n\to\infty]{\mathrm{P}} 0. \tag{27}$$

On the set $\max_{1\le j\le d}|\lambda_j-\widehat{\lambda}_{nj}|+\max_{1\le j\le d}\|\psi_j-\widehat{c}_j\widehat{\psi}_{nj}\|\le An^{-1/2}$ and for $n\ge N_0$ such that $An^{-1/2}<\lambda_d/2$, simple algebra gives $D_n\le D_{n1}+D_{n2}$, where

$$\begin{aligned} D_{n1} &= \max_{1\le j\le d}\left|\frac{1}{\lambda_j}-\frac{1}{\widehat{\lambda}_{nj}}\right|\max_{1\le k\le 1}\left|\sum_{i=1}^{k}\langle X_i-\overline{X}_n,\psi_j\rangle\right| \\ &\le \frac{2}{\lambda_d}\max_{1\le j\le n}|\widehat{\lambda}_{nj}-\lambda_j|n^{-1/2}T_{n,1}(d)\to 0 \ \text{ as } \ n\to\infty, \end{aligned}$$

and

$$\begin{aligned} D_{n2} &\le n^{-1/2}\frac{2}{\varepsilon}\max_{1\le k\le n}\left\|\sum_{i=1}^{k}[X_i-\overline{X}_n]\right\|\max_{1\le j\le d}\|\widehat{\psi}_{nj}-\psi_j\| \\ &\le \frac{2A}{\varepsilon}n^{-1}\max_{1\le k\le n}\left\|\sum_{i=1}^{k}[X_i-\overline{X}_n]\right\|\to 0 \ \text{ as } \ n\to\infty \end{aligned}$$

by the law of large numbers. Lemma 2 concludes the proof. □

Test (20) now becomes

$$\widehat{T}_{n,1}(d)\ge\sqrt{n}C_\alpha \tag{28}$$

and has asymptotic level α by Theorem 7.

3.2. Testing at Most m Change-Points

For $m>1$, let $\mathcal{N}_m$ be a set of all partitions $\kappa=(k_i,i=0,1,\ldots,m)$ of the set $\{0,1,\ldots,n\}$ such that $0=k_0<k_1<\cdots<k_{m-1}<k_m=n$. Next, consider for fixed integers d, $1\le m<n$ and real $p>2$,

$$T_{n,m}(d,p):=\max_{1\le j\le d}\frac{1}{\sqrt{\lambda_j}}\max_{\kappa\in\mathcal{N}_m}\left\{\sum_{k=1}^{m}\left|\sum_{k_{i-1}+1}^{k_i}\langle X_k-\overline{X}_n,\psi_j\rangle\right|^p\right\}^{1/p}. \tag{29}$$

The statistics $T_{n,m}(d,p)$ are designed for testing at most m change-points in a sample.

Theorem 8. *Let the random sample $(X_i,i=1,\ldots,n)$ be as in Theorem 1. Then:*

(a) *Under H_0,*

$$n^{-1/2}T_{n,m}(d,p)\xrightarrow[n\to\infty]{\mathcal{D}}\max_{1\le j\le d}v_{m,p}^{1/p}(B_j),$$

where $B_1,\ldots,B_d$ are independent standard Brownian bridges.

(b) *Under H_A,*

$$n^{-1/2}T_{n,m}(d,p)\xrightarrow[n\to\infty]{\mathcal{D}}\max_{1\le j\le d}v_{m,p}^{1/p}(B_j+\Delta\langle q,\psi_j/\sqrt{\lambda_j}\rangle),$$

where $\Delta(t),t\in[0,1]$ is as defined in Theorem 2.

(c) *Under H'_A,*

$$n^{-1/2}T_{n,m}(d,p)\xrightarrow[n\to\infty]{\mathrm{P}}\infty.$$

Proof. For $1\le m\le n$ and $q=p/(p-1)$, set

$$\mathcal{F}_{m,q}:=\left\{\sum_{j=1}^{m}b_j\mathbf{1}_{(t_{j-1},t_j]}:\ \sum_{j=1}^{m}|b_j|^q\le 1,0=t_0<t_1<\cdots<t_m=1\right\}. \tag{30}$$

It is easy to check that $\mathcal{F}_{m,q}\subset\mathcal{W}_q[0,1]$. Since

$$\sup\left\{\left|\sum_{k=1}^{m}a_kb_k\right|:\ \sum_{k=1}^{m}|b_k|^q\le 1\right\}=\left(\sum_{k=1}^{m}|a_k|^p\right)^{1/p},$$

we have

$$T_{n,m}(d) = \max_{\psi \in \Psi_{d,\gamma}} \max_{f \in \mathcal{F}_{m,q}} |\mu_n(\psi, f)|,$$

and the results follow from Theorem 2. □

In particular, if there is $s_1^*, s_2^* \in (0,1)$ such that $u(s) = \mathbf{1}_{[s_1^*, s_2^*]}(s), s \in [0,1]$, then (1) corresponds to the so-called changed segment model. In this case, we have $\Delta(t) = \Delta_2^*(t) := \max\{0, \min\{t, s_2^*\} - s_1^*\} - t(s_2^* - s_1^*)$, $t \in [0,1]$. Figure 3) shows the generated density functions of $\max_{1 \le k \le d} v_{4,p}^{1/p}(B_k)$ and $\max_{1 \le k \le d} v_{4,p}^{1/p}(B_k + a_k \Delta_2^*)$ for different values of $d \ge 1$, $0 < s_1^* < s_2 < 1$, and $p > 2$. The numbers $a_1, \dots, a_d$ were sampled from the uniform distribution on $[0, 15]$.

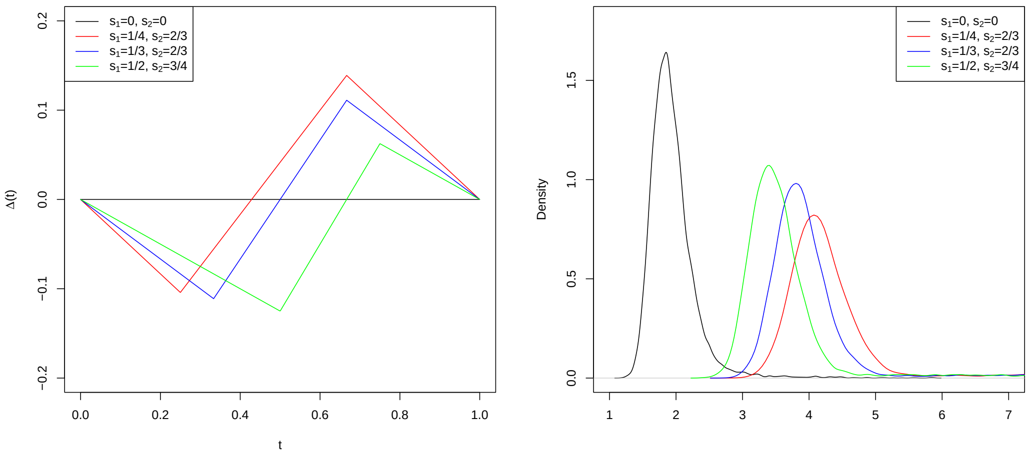


Figure 3. Functions Δ_2^* and density functions.

With the estimated eigenvalues and eigenfunctions, we define

$$\widehat{T}_{n,m}(d,p) := \max_{1 \le j \le d} \frac{1}{\sqrt{\widehat{\lambda}_{nj}}} \max_{\kappa \in \mathcal{N}_m} \left\{ \sum_{k=1}^{m} \left| \sum_{k_{i-1}+1}^{k_i} \langle X_k - \overline{X}_n, \widehat{\varphi}_{nj} \rangle \right|^p \right\}^{1/p}. \quad (31)$$

Theorem 9. *Let the functional sample $(X_k, k = 1, \dots, n)$ be defined by (1) where $Y, Y_1, Y_2, \dots$ satisfies Assumptions 1 and 2. Then:*

(a) *Under H_0,*

$$n^{-1/2} \widehat{T}_{n,m}(d,p) \xrightarrow[n\to\infty]{\mathcal{D}} \max_{1 \le j \le d} v_{m,p}^{1/p}(B_j),$$

where $B_1, \dots, B_d$ are independent standard Brownian bridges.

(b) *Under H_A,*

$$n^{-1/2} \widehat{T}_{n,m}(d,p) \xrightarrow[n\to\infty]{\mathcal{D}} \max_{1 \le j \le d} v_{m,p}^{1/p}(B_j + \Delta \langle q, \psi_j / \sqrt{\lambda_j} \rangle),$$

where $\Delta(t), t \in [0,1]$ is as defined in Theorem 2.

(c) *Under H'_A,*

$$n^{-1/2} \widehat{T}_{n,m}(d,p) \xrightarrow[n\to\infty]{\mathrm{P}} \infty.$$

Proof. This goes along the lines of the proof of Theorem 7. □

According to Theorems 8 and 9, the tests:

$$T_{n,m}(d,p) \ge \sqrt{n} C_\alpha(m,d,p) \;\text{ and }\; \widehat{T}_{n,m}(d,p) \ge \sqrt{n} C_\alpha(m,d,p) \quad (32)$$

respectively, will have asymptotic level α, if $C_\alpha(m, d, p)$ is such that

$$P(v_{m,p}^{1/p}(B) \leq C_\alpha(m, d, p)) = (1-\alpha)^{1/d}.$$

3.3. Testing Unknown Number of Change-Points

Next, consider for fixed integers d as above and real $p > 2$,

$$T_n(d, p) := \max_{1\leq j\leq d} \frac{1}{\sqrt{\lambda_j}} \max_{1\leq m\leq n} \max_{\kappa\in\mathcal{N}_m} \Big\{ \sum_{k=1}^{m} \Big| \sum_{k_{i-1}+1}^{k_i} \langle X_k - \overline{X}_n, \psi_j\rangle \Big|^p \Big\}^{1/p}. \tag{33}$$

The statistics $T_n(d, p)$ are designed for testing an unknown number of change-points in a sample.

Theorem 10. *Let random sample $(X_i, i = 1, \ldots, n)$ be as in Theorem 1. Then:*

(a) *Under H_0,*

$$n^{-1/2} T_n(d, p) \xrightarrow[n\to\infty]{\mathcal{D}} \max_{1\leq j\leq d} v_p^{1/p}(B_j),$$

where $B_1, \ldots, B_d$ are independent standard Brownian bridges.

(b) *Under H_A,*

$$n^{-1/2} T_n(d, p) \xrightarrow[n\to\infty]{\mathcal{D}} \max_{1\leq j\leq d} v_p^{1/p}(B_j + \Delta\langle q, \psi_j/\sqrt{\lambda_j}\rangle),$$

where $\Delta(t), t \in [0, 1]$ is as defined in Theorem 1.

(c) *Under H_A',*

$$n^{-1/2} T_n(d, p) \xrightarrow[n\to\infty]{\mathrm{P}} \infty.$$

Proof. For $q = p/(p-1)$, set

$$\mathcal{F}_q := \Big\{ \sum_{j=1}^{m} b_j \mathbf{1}_{(t_{j-1}, t_j]} : \sum_{j=1}^{\infty} |b_j|^q \leq 1, 0 = t_0 < t_1 < \cdots < t_m = 1, m \geq 1 \Big\}. \tag{34}$$

It is easy to check that $\mathcal{F}_q \subset \mathcal{W}_q[0, 1]$. Since

$$\sup\Big\{ \Big| \sum_{k=1}^{\infty} a_k b_k \Big| : \sum_{k=1}^{\infty} |b_k|^q \leq 1 \Big\} = \Big(\sum_{k=1}^{\infty} |a_k|^p \Big)^{1/p},$$

we have

$$T_n(d) = \max_{\psi\in\Psi_{d,\gamma}} \max_{f\in\mathcal{F}_{m,q}} |\mu_n(\psi, f)|,$$

and both statements (a) and (b) follow from Theorem 1. □

With the estimated eigenvalues and eigenfunctions, we define:

$$\widehat{T}_n(d, p) := \max_{1\leq j\leq d} \frac{1}{\sqrt{\widehat{\lambda}_{nj}}} \max_{1\leq m\leq n} \max_{\kappa\in\mathcal{N}_m} \Big\{ \sum_{k=1}^{m} \Big| \sum_{k_{i-1}+1}^{k_i} \langle X_k - \overline{X}_n, \widehat{\psi}_{nj}\rangle \Big|^p \Big\}^{1/p}. \tag{35}$$

Theorem 11. *Let random sample (X_i) be as in Theorem 1. Then:*

(a) *Under H_0,*

$$n^{-1/2} \widehat{T}_n(d, p) \xrightarrow[n\to\infty]{\mathcal{D}} \max_{1\leq j\leq d} v_p^{1/p}(B_j),$$

where $B_1, \ldots, B_d$ are independent standard Brownian bridges.

(b) *Under H_A,*

$$n^{-1/2}\widehat{T}_n(d,p) \xrightarrow[n\to\infty]{\mathcal{D}} \max_{1\le j\le d} v_p^{1/p}(B_j + \Delta\langle q, \psi_j/\sqrt{\lambda_j}\rangle),$$

where $\Delta(t), t \in [0,1]$ is as defined in Theorem 1.

(c) *Under H'_A,*

$$n^{-1/2}\widehat{T}_n(d,p) \xrightarrow[n\to\infty]{\mathrm{P}} \infty.$$

Proof. This goes along the lines of the proof of Theorem 7. □

According to Theorems 10 and 11, the tests:

$$T_n(d,p) \ge \sqrt{n}C_\alpha(d,p) \text{ and } \widehat{T}_n(d,p) \ge \sqrt{n}C_\alpha(d,p) \tag{36}$$

respectively, will have asymptotic level α, if $C_\alpha(d,p)$ is such that

$$P(v_p^{1/p}(B) \le C_\alpha(d,p)) = (1-\alpha)^{1/d}.$$

The quantiles of distribution function of $v_p^{1/p}(B)$ were estimated in [12].

4. Simulation Results

We examined the above-defined test statistics in a Monte Carlo simulation study. In the first subsection, we describe the simulated data under consideration. The statistical power analysis of the tests (36) and (32) is presented in Section 4.2.

4.1. Data

We used the following three scenarios:

(S1) Let (ξ_{jk}) be i.i.d. symmetrized Pareto random variables with index p (we used $p = 5$). Set

$$Y_j(t) = \sum_{k=1}^{d} \xi_{jk} \frac{\sqrt{2}\cos(k\pi t)}{k\sigma}, \quad t \in [0,1], \ j \ge 1, \tag{37}$$

where $\sigma^2 = E\xi_{11}^2$. Under the null hypothesis, we take $X_k = Y_k, k = 1,2,\ldots,n$. Under the alternative, we consider

$$X_j(t) = u_n(j/n)\sum_{k=1}^{d} a_{nk}\cos(k\pi t) + Y_j, \quad t \in [0,1], j = 1,\ldots,n,$$

where the function u_n defines the change-points' configuration and the coefficients (a_{nk}) are subject to choice.

(S2) We start with discrete observations $(x_{ij}, j = 0,1,\ldots,M), i = 1,\ldots,n$, by taking $x_{ij} = X_i(\tau_j)$, where the random sample $(X_j, j = 1,\ldots,n)$ is generated as in scenario (S1). Discrete observations are converted to the functional data $(X_j, j = 1,\ldots,n)$ by using B-spline bases.

(S3) Discrete observations $(i/M, y_{ij}), i = 0,1,\ldots,M, j = 1,\ldots,n$, are generated by taking

$$y_{ij} = M^{-1/2}\sum_{k=1}^{i} \xi_{kj},$$

so that y_{ij} can be interpreted as the observation of a standard Wiener process at i/M. From $(y_{ij}, i = 1,\ldots,M)$, the function Y_j is obtained using the B-spline smoothing technique. During the simulation, we used $M = 1000$ and $D = 50$ B-spline functions, thus obtaining $n = 500$ functions $Y_1, \ldots, Y_n$. Then, we define for $j = 1, \ldots, n$,

$$X_j = \begin{cases} Y_j, & \text{under null} \\ u_n(j/n)q_n + Y_j, & \text{under alternative} \end{cases}$$

and consider different configurations u_n of change-points and $q_n(t) = a_n\sqrt{Mt}$, $t \in [0,1]$.

We mainly concentrated on two possible change-point alternatives. The first is obtained with $u_n(t) = \mathbf{1}_{[0,\theta]}(t)$ and corresponds to one change-point alternative. Another is for the epidemic-type alternative, for which we take $u_n(t) = \mathbf{1}_{[\theta_1,\theta_2]}(t)$.

Scenario (S1) is used as an optimal case situation where the actual eigenvalues and eigenfunctions are known. In this case, we are not required to approximate discrete functions, thus avoiding any data loss or measurement errors. The second scenario continues with the same random functional sample, but goes through extra steps such us taking function values at discrete data points and reconstructing the random functional sample on a different set of basis functions. The aim of this exercise is to measure the impact when some information could be lost due to measurements taken at discrete points and smoothing. The simulation results show that, even after the reconstruction of the random functional sample, the performance of the test does not suffer too much.

Our simulation starts with the generation process of the random functional sample Y_j as described in the first scenario with $d = 30$.

First of all, we can compare the true eigenfunctions of covariance operator $\Gamma = E[Y_j \otimes Y_j]$ with the eigenfunctions of estimated operator $\widehat{\Gamma}_n$ (see Figure 4).

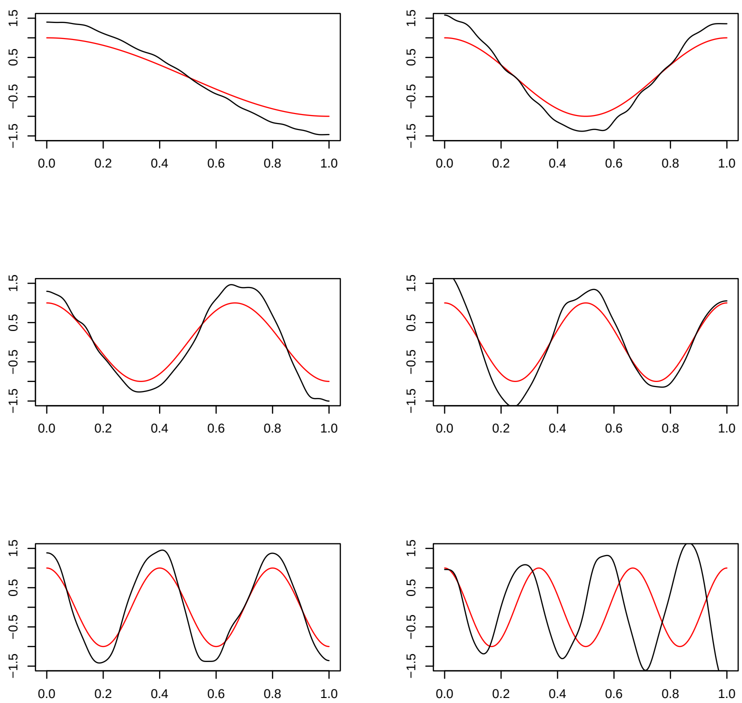


Figure 4. True basis functions (red) and "reconstructed" basis functions (black) using fPCA method.

We see that the estimated harmonics has almost the same shape; only every second, the estimated eigenfunctions are phase shifted.

Next, for both scenarios (S1) and (S2), the density functions of the test statistic $T_{n,1}(d)$ (17) were estimated using Monte Carlo with 10,000 repetitions (see Figure 5). It shows four density plots: the red density functions of $T_{n,1}(d)$ are calculated using the true eigenfunctions and eigenvalues, while the black curves show the density of $\widehat{T}_{n,1}(d)$ (26) using the estimated eigenfunctions and eigenvalues. The left side density plots were estimated from the samples under the null hypothesis, while the plots on the right side show the density of $T_{n,1}(d)$ and $\widehat{T}_{n,1}(d)$ with the sample:

$$X_j(t) = \sum_{k=1}^{d} (\xi_{jk} + \mathbf{1}_{(\tau>j)} a) \frac{\sqrt{2}\cos(k\pi t)}{k\sigma}, t \in [0,1], \tau = 250, j = 1 \ldots 500 \tag{38}$$

with added drift $a = 0.2$.

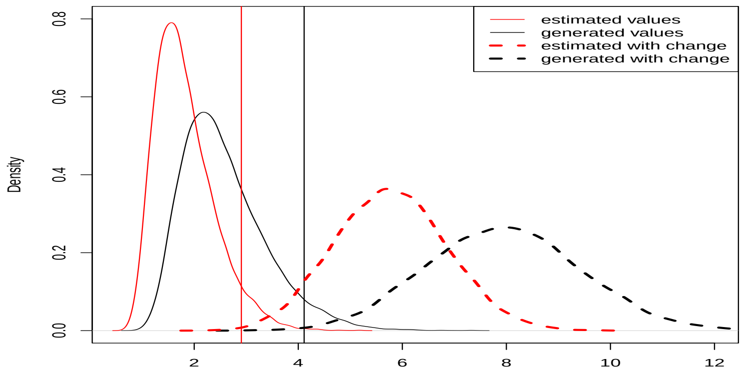


Figure 5. Density plots of the test statistic $T_{n,1}(d)$ and $\widehat{T}_{n,1}(d)$.

Since functional Principal Component Analysis (fPCA) represents a functional data sample in the most parsimonious way, we can see that the density of the test statistics in scenario (S2) is more on the left side and more concise. Critical values $c_d(\alpha)$ with $\alpha = 0.05$ of the statistics $T_{n,1}(d)$ and $\widehat{T}_{n,1}(d)$ were also calculated and are shown in Figure 5.

4.2. Statistical Power Analysis

First, we compared the statistical power of the test (20) with statistic $T_{n,1}(d)$ of the scenario (S1) and scenario (S2) with statistic $\widehat{T}_{n,1}(d)$. To this aim, we used sample $(X_j, j = 1, \ldots, n)$ defined in (38), where $\tau = 250$, which is in the middle of the sample. We started with the no drift $a = 0$ (corresponding to the null hypothesis) and increasing the drift amount a by 0.03 up to the point when $a > 0.3$. At each a value, we repeated the simulation 1000 times. This gives a good indication of the statistical power with the amount of the added drift. The statistical power is illustrated in Figure 6. Based on the simulation results, we can see that, even if the random functional sample is approximated from the discrete data points, it still holds the same statistical power and the performance does not suffer from the information loss due to smoothing and fPCA. These are important results, because, normally, in observed real-world data, the true functions are unknown and have to be approximated, which almost always introduces measurement errors.

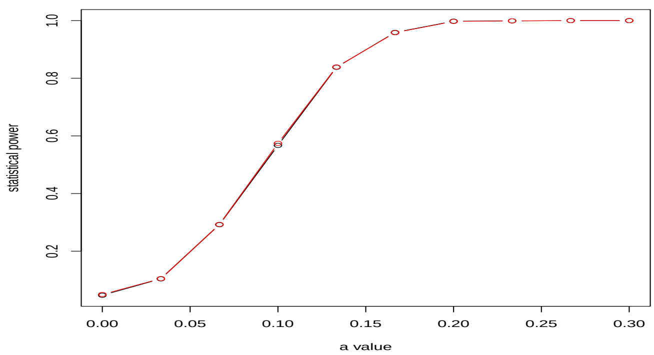
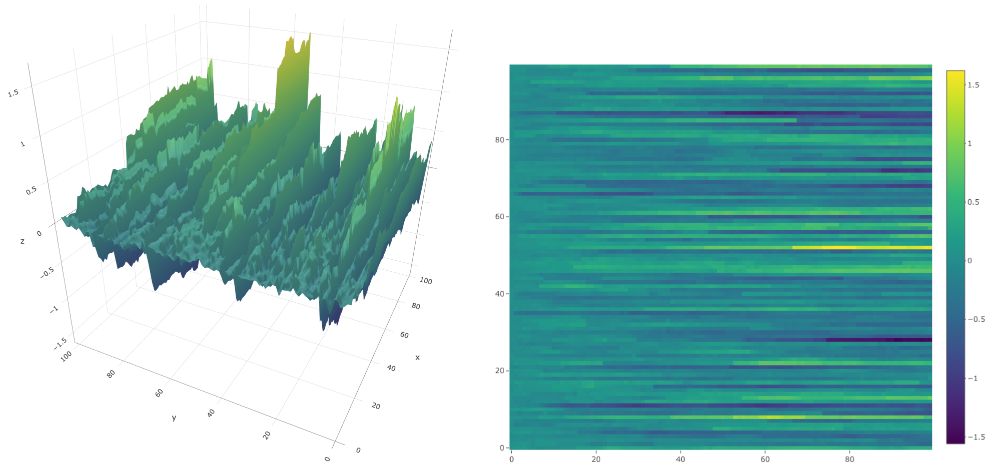


Figure 6. The comparison of the statistical power of scenario (S1) and (S2).

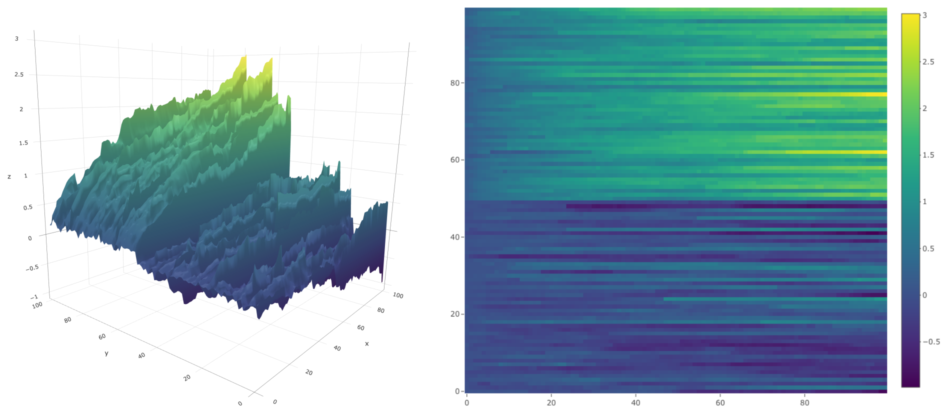
Next, we focus on the power tests (36) and (32) used directly on the functional data sets simulated in scenario (S3). Figures 7 and 8 present the clear opposites of the functional data sets with respect to the change-point. The changes can be easily observed. However, especially working with functional data sets, the changes may not be that obvious. As an example, Figure 9 illustrates another functional data set with the change-point, where the presence of the change-point is not visible, but Monte Carlo experiments show that, with the same magnitude of change, for almost 80% of the cases, H_0 was correctly rejected.

The density of the limiting distribution and asymptotic critical values were estimated using the Monte Carlo technique by simulating a Brownian bridge with 1000 points and running 100,000 replications.



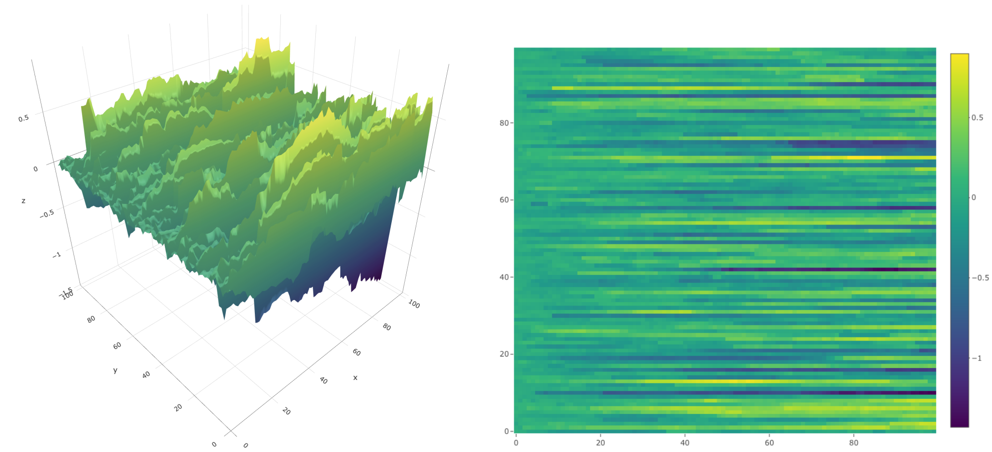
(**a**) 3D plot (**b**) Heat map plot

Figure 7. Functional data set (Y_j).



(**a**) 3D plot (**b**) Heat map plot

Figure 8. Functional sample (X_j) with one change-point.



(**a**) 3D plot of sample with change (**b**) Heat map plot of sample with change

Figure 9. Sample with introduced drift of magnitude $a = 0.004$ after the change-point.

In the power studies, we tested two variants of the random functional samples, one with a single change-point in the middle of the functional sample and the second with the two change-points forming epidemic change. In the first case, the functional sample $X_1, \ldots, X_n$ is constructed from $n = 1000$ random functions where 500 curves are changed in order to violate the null hypothesis. The model that violates the null hypothesis is defined as $X_k(t) = \Delta(t)\mathbf{1}\{i > n/2\} + Y_i(t)$, $\Delta(t) = a\sqrt{Mt}$, $t \in [0,1]$, $M = 1000$, and the parameter a is used to control the magnitude of the drift after the change-point. In the second case, during each iteration, $n = 1500$ random functions are generated, where 500 curves in the middle were modified by taking $X_k(t) = \Delta(t)\mathbf{1}\{2n/3 > i > n/3\} + Y_i(t)$, $t \in [0,1]$. During each repetition, twostatistics are calculated: $\widehat{T}_n(d,p)$ (35) and $\widehat{T}_{n,1}(d)$ (26) in the single change-point simulation. For the epidemic change simulation $\widehat{T}_{n,m}(d,p)$ (31), $m = 2$ statistic is calculated. We set the p-variation p parameter to 3. We also tested with different p-values, but this did not have any impact on the overall performance.

Figure 10 presents the results of the statistical power simulation. From the results, we can see that epidemic change has weaker statistical power. On the other hand, when restricting the partition count, we observed one benefit, that the locations of the partitions in many cases match or are very close to the actual locations of the change-point.

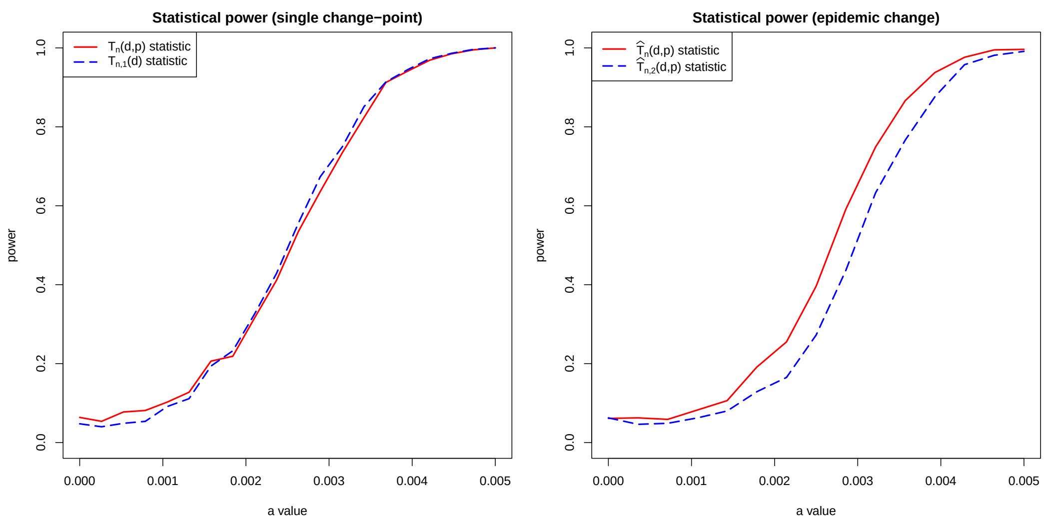


Figure 10. Power curves.

5. Application to Brain Activity Data

The findings of real data analysis to show the performance of the proposed test are demonstrated in this section. The data were collected during a long-term study on voluntary alcohol- consuming rats following chronic alcohol experience. The data consist of two sets: neurophysiological activity from the two brain centers (the dorsal and ventral striatum) and data from the lickometer device. The lickometer devices were used to monitor the drinking bouts. During the single trial, two locations of the brain were monitored for each rat. Rats were given two drinking bouts, one with alcohol and the other with water. Any time, they were able to freely choose what to drink. Electrodes were attached to the brains, and neurophysiological data were sampled at 1kHz intervals. It was not the goal of this study to confirm nor reject the findings, but to show the advantages of the functional approach for change-point detection. For this reason, the data are well suited to illustrate the behavior of the test in real-world settings.

In our analysis, we took the first alcohol drinking event, which lasted around 27 s. We also included 10 s before the drinking event and 10 s after the event. The total time was 47 s long. The time series was broken down into processes of 100 ms. Each process had 100 data points.

$$A_{470,100} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,100} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,100} \\ \vdots & \vdots & \ddots & \vdots \\ a_{470,1} & a_{470,2} & \cdots & a_{470,100} \end{pmatrix}$$

All the processes were smoothed to the functions using 50 B-spline basis functions. The overall functional sample contained 470 functions $\widehat{F} = [f_1, f_2, \ldots, f_{470}]$. The functional sample was separated into sub-samples $\widehat{F}_i = [f_1, f2, \ldots, f_{20+i}]$, $i = 0, 1, \ldots, 450$. For each sub-sample $\widehat{F}_i$, two statistics were calculated ($\widehat{T}_n(d, p)$ (35) and $\widehat{T}_{n,m}(d, p)$ (31), $m = 2$).

The results are visualized in Figure 11. We can see that tests with statistics $\widehat{T}_n(d, p)$ and $\widehat{T}_{(n,m)}(d, p)$ strongly rejected the null hypothesis at around 2 s and onward after the rat started to consume the alcohol, which suggests that the changes in the brain activity can be observed. However, the changes appear to happen only for the CPu brain region. Interestingly, the statistic $\widehat{T}_{n,m}(d, p)$ has much larger volatility compared to the unrestricted $\widehat{T}_n(d, p)$ in the Nacc brain region before the drinking event and lower volatility just after the drinking event started. However, it is not fully clear if this is the expected behavior or a Type I error.

Finally, the locations of the restricted ($m = 2$) p-variation partition points nearly matched the beginning and the end of the drinking period. In Figure 11, the gray vertical dashed lines indicate the actual beginning and the actual end of the drinking period measured by the lickometer and the black vertical lines indicate the location of the partitions calculated from the functional sample $\widehat{F}_{450}$. The first partition is located at 10.5 s and the second partition point at 38.4 s, which aligns well with the data collected from the lickometer.

The test with a restricted partition count showed weaker statistical power, but it did help determine the location of the change-points.

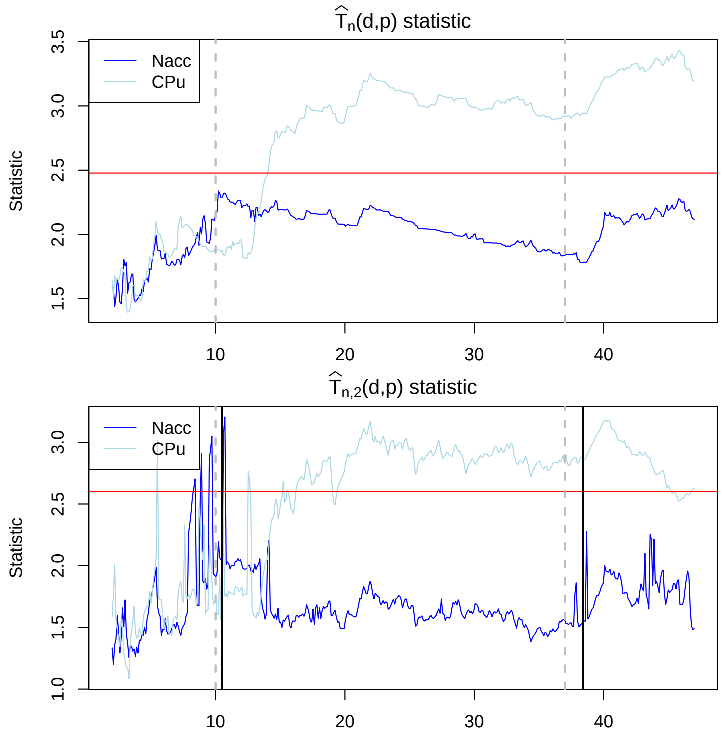


Figure 11. Statistics of the first alcohol drinking event, which lasted about 27 s. Ten seconds before and 10 s after were also included. The red horizontal line indicates the critical value with 0.95. Vertical gray dashed lines mark the beginning and the end of the drinking time. The black solid vertical lines mark the locations of the change-points detected using the restricted p-variation method. Blue and light blue colors represent different brain regions.

6. Proof of Theorems 1 and 3

The following theorem is a version of Theorem 2.11.1 in [13] adapted to the case of continuous processes.

Theorem 12. *Assume that* $\{Z_{ni}\colon\ 1 \le i \le m_n\}$ *are independent continuous stochastic processes indexed by a totally bounded semi-metric space* $(\mathcal{G}, d)$ *such that*

$$\lim_{n\to\infty} \sum_{i=1}^{m_n} E\|Z_{ni}\|_{\mathcal{F}}^2 1_{\{\|Z_{ni}\|_{\mathcal{F}}>\eta\}} = 0 \quad \textit{for every } \eta > 0, \tag{39}$$

$$\lim_{n\to\infty} \sup_{d(f,g)<\delta_n} \sum_{i=1}^{m_n} E\big[Z_{ni}(f) - Z_{ni}(g)\big]^2 = 0 \quad \textit{for every } \delta_n \downarrow 0, \tag{40}$$

$$\int_0^{\delta_n} \sqrt{\log N(\epsilon, \mathcal{G}, d_n)}\, d\epsilon \xrightarrow[n\to\infty]{\mathrm{P}} 0 \quad \textit{for every } \ \delta_n \downarrow 0, \tag{41}$$

where

$$d_n(f,g) = \Big(\sum_{k=1}^{m_n} [Z_{nk}(f) - Z_{nk}(g)]^2\Big)^{1/2}.$$

Then, the sequence $Z_n := \sum_{i=1}^{m_n}(Z_{ni} - EZ_{ni})$ *is asymptotically d-equicontinuous, that is, for every* $\varepsilon > 0$,

$$\lim_{\delta\downarrow 0} \limsup_{n\to\infty} P\big(\sup_{d(f,g)<\delta} |Z_n(f) - Z_n(g)| > \varepsilon\big) = 0.$$

Furthermore, (Z_n) *converges in law in* $\ell^\infty(\mathcal{G})$ *provided that covariances converge pointwise on* $\mathcal{G} \times \mathcal{G}$.

Proof of Theorem 1 $(1a)$**.** Without loss of generality, we assumed that $\|\psi\| \le 1$ for all $\psi \in \Psi$ and $\|f\|_{\sup} \le 1$ for all $f \in \mathcal{F}$. To prove $(1a)$, we applied Theorem 12 for $\mathcal{G} = \mathcal{F} \times \Psi$, $d = \rho_2$, and $Z_{nk} = n^{-1/2}\nu_{nk}$, $k = 1, \dots, n$, where, under H_0, $\nu_{nk}(f, \psi) = \langle Y_k, \psi \rangle \lambda_{nk}(f)$. Let us check first the conditions (39)–(41). We have

$$n^{-1/2}\|\nu_{nk}\|_{\mathcal{G}} = n^{-1/2} \sup_{\psi \in \Psi, f \in \mathcal{F}} |\langle Y_k, \psi \rangle \lambda_{nk}(f)| \le n^{-1/2}\|Y_k\| \sup_{\psi \in \Psi} \|\psi\| \sup_{f \in \mathcal{F}} \|f\|_{\sup}$$
$$\le n^{-1/2}\|Y_k\|.$$

Hence, (39) easily follows from $E\|Y\|^2 < \infty$. Since

$$\nu_{nk}(f, \psi) - \nu_{nk}(f', \psi') = \langle Y_k, \psi - \psi' \rangle \lambda_{nk}(f) + \langle Y_k, \psi' \rangle \lambda_{nk}(f - f')$$

and Y, Y_k are identically distributed, we have

$$E[\nu_{nk}(f, \psi) - \nu_{nk}(f', \psi')]^2 \le 2E\|Y\|^2[\lambda_{nk}^2(f - f')\|\psi'\|^2 + \lambda_{nk}^2(f)\|\psi - \psi'\|].$$

Summing this estimate and noting that for any $g \in L_2(0,1)$,

$$n^{-1}\sum_{k=1}^{n} \lambda_{nk}^2(g) \le n^{-1}\sum_{k=1}^{n} \lambda_{nk}(g^2) = \|g\|^2$$

by the Hölder inequality, we find

$$n^{-1}\sum_{k=1}^{n} E[\nu_{nk}(f, \psi) - \nu_{nk}(f', \psi')]^2 \le 2E\|Y\|^2[\|\psi\|^2\|f - f'\|^2 + \|f'\|\|\psi - \psi'\|^2]$$
$$\le 2E\|Y\|^2[\|\psi\|^2 + \|f'\|^2]\delta_n$$
$$\le 4E\|Y\|^2\delta_n$$

if $\rho_2((f, \psi), (f', \psi')) < \delta_n$. This estimate yields (40). To check (41), we have

$$d_n((f, \psi), (f', \psi')) = \Big(n^{-1}\sum_{k=1}^{n}[\langle Y_k, \psi \rangle \lambda_{nk}(f) - \langle Y_k, \psi' \rangle \lambda_{nk}(f')]^2\Big)^{1/2}$$
$$= \Big(n^{-1}\sum_{k=1}^{n}[\langle Y_k, \psi - \psi' \rangle \lambda_{nk}(f) + \langle Y_k, \psi' \rangle \lambda_{nk}(f - f')]^2\Big)^{1/2}$$
$$\le \Big(n^{-1}\sum_{k=1}^{n}\langle Y_k, \psi - \psi' \rangle^2 \lambda_{nk}^2(f)\Big)^{1/2} + \Big(n^{-1}\sum_{k=1}^{n}\langle Y_k, \psi' \rangle \lambda_{nk}^2(f - f')\Big)^{1/2}$$
$$\le \Big(n^{-1}\sum_{k=1}^{n}\|Y_k\|^2 \lambda_{nk}(f^2)\Big)^{1/2} \rho_{2,\lambda}(\psi, \psi') + \Big(n^{-1}\sum_{k=1}^{n}\langle Y_k, \psi' \rangle^2 \lambda_{nk}(f - f')^2\Big)^{1/2}$$
$$\le A_n[\rho_{2,\lambda}(\psi, \psi') + \rho_{2,Q}(f, f')],$$

where

$$A_n = \Big(n^{-1}\sum_{k=1}^{n}\|Y_k\|^2\Big)^{1/2}, \text{ and } Q = A_n^{-2}n^{-1}\sum_{k=1}^{n}\|Y_k\|^2\lambda_{nk}$$

Hence,

$$N(\varepsilon, \mathcal{F} \times \Psi, d_n) \le N(A_n^{-1}\varepsilon, \mathcal{F}, \rho_{2,Q})N(A_n^{-1}\varepsilon, \Psi, \rho_{2,\lambda}).$$

and the condition (41) is satisfied, provided that

$$I_1(\delta_n) := \int_0^{\delta_n} \sup_{Q \in \mathcal{Q}} \sqrt{\log N(A_n^{-1}\epsilon, \mathcal{F}, \rho_{2,Q})}\, d\epsilon \xrightarrow[n\to\infty]{\mathrm{P}} 0 \quad \text{for every } \delta_n \downarrow 0, \tag{42}$$

and

$$I_2(\delta_n) := \int_0^{\delta_n} \sqrt{\log N(B_n^{-1}\epsilon, \Psi, \rho_{2,\lambda})}\, d\epsilon \xrightarrow[n\to\infty]{\mathrm{P}} 0 \quad \text{for every } \ \delta_n \downarrow 0, \tag{43}$$

hold. Set

$$J_1(a) := \int_0^a \sup_{Q\in\mathcal{Q}} \sqrt{\log N(\epsilon, \mathcal{F}, \rho_{2,Q})}\, d\epsilon, \ \ J_2(a) := \int_0^a \sqrt{\log N(\epsilon, \Psi, \rho_{2,\lambda})}\, d\epsilon.$$

It is known (see, e.g., [14]) that $J_1(1) < \infty$. Hence, $J_1(a) \to 0$ as $a \to 0$. By the condition (4), $J_2(a) \to 0$ as $a \to 0$. Changing the integration variables gives $I_1(\delta_n) = A_n J_1(A_n^{-1}\delta_n)$ and $J_2(\delta_n) = A_n J_2(A_n^{-1}\delta_n)$.

Set $\sigma^2 := E\|Y\|^2$. By the strong law of large numbers, $A_n^2 \xrightarrow[n\to\infty]{\mathrm{P}} \sigma^2$. Choosing $\eta < 3\sigma^2/4$, we have, for any $\delta > 0$,

$$\begin{aligned}
P\Big(I_1(\delta_n) > \delta\Big) &\le P(A_n J_1(A_n^{-1}\delta_n) > \delta, |A_n^2 - \sigma^2| < \eta\Big) + P(|A_n^2 - \sigma^2| > \eta) \\
&\le P(A_n J_1(A_n^{-1}\delta_n) > \delta, A_n^2 > \sigma^2/4\Big) + P(|A_n^2 - \sigma^2| > \eta) \\
&\le P\Big(A_n J_1(\eta\delta_n/2) > \delta) + P(|A_n^2 - \sigma^2| > \eta) \to 0
\end{aligned}$$

as $n \to \infty$. Similarly, we prove $I_2(\delta_n) \xrightarrow[n\to\infty]{\mathrm{P}} 0$.

Next, we have to check the pointwise convergence of the covariances of (Z_n). Since Y_k are independent, we have

$$\begin{aligned}
E\Big(\sum_{k=1}^n \langle Y_k, \psi\rangle \lambda_{nk}(f) \sum_{k=1}^n \langle Y_k, \psi'\rangle \lambda_{nk}(f')\Big) &= \sum_{k=1}^n E\Big(\langle Y_k, \psi\rangle\langle Y_k, \psi'\rangle\Big)\lambda_{nk}(f)\lambda_{nk}(f')\Big) \\
&= (\Gamma\psi, \psi') \sum_{k=1}^n \lambda_{nk}(f)\lambda_{nk}(f').
\end{aligned}$$

We shall prove that

$$I_n := n^{-1} \sum_{k=1}^n \lambda_{nk}(f)\lambda_{nk}(f') \to \langle f, f'\rangle \ \text{ as } \ n \to \infty.$$

Set $\widetilde{I}_n := n^{-1}\sum_{k=1}^n f(k/n)f'(k/n)$. Evidently, $\lim_{n\to\infty} \widetilde{I}_n = \langle f, f'\rangle$, and we have to check

$$\lim_{n\to\infty} |I_n - \widehat{I}_n| \to 0.$$

We have

$$\Delta_n := |I_n - \widehat{I}_n| \le n^{-1} \sum_{k=1}^n [\lambda_{nk}(f)\lambda_{nk}(f') - f(k/n)f'(k/n)] \le |\Delta_n'| + |\Delta_n''|,$$

where

$$\Delta_n' := n^{-1} \sum_{k=1}^n [\lambda_{nk}(f)(\lambda_{nk}(f') - f'(k/n))], \ \ \Delta_n'' := n^{-1} \sum_{k=1}^n [f'(k/n)(\lambda_{nk}(f) - \lambda_{nk}(f'))].$$

Observing that

$$|\lambda_{nk}(f) - f(k/n)| = \Big| \int_0^1 (f(t) - f(k/n)) d\lambda_{nk}(t) \Big| \le n \Big| \int_{(k-1)/n}^{k/n} (f(t) - f(k/n)) dt \Big|$$
$$\le \int_0^1 \sup\{|f(t) - f(k/n)| : t \in [(k-1)/n, k/n]\} d\lambda_{nk}$$
$$\le \sup\{|f(t) - f(k/n)| : t \in [(k-1)/n, k/n]$$
$$\le v_2^{1/2}(f; [(k-1)/n, k/n)]),$$

we have

$$|\Delta_n| \le n^{-1} \lambda_{nk}(f) v_2^{1/2}(f, [(k-1]/n, k/n])$$
$$\le n^{-1} \Big(\sum_{k=1}^n \lambda_{nk}^2(f) \Big)^{1/2} \Big(\sum_{k=1}^n v_2(f, [(k-1)/n, k/n]) \Big)^{1/2}$$
$$\le n^{-1/2} \|f\| v_2^{1/2}(f).$$

This yields

$$\lim_{n\to\infty} E(Z_n(f,\psi) Z_n(f',\psi')) = \langle \Gamma\psi, \psi' \rangle \langle f, f' \rangle.$$

To complete the proof of (a), note that the existence of the continuous modification of Gaussian process $\nu = \nu(\psi, f), (\psi, f) \in \mathcal{G} = \Psi \times \mathcal{F})$ follows by Dudley [15], since the entropy condition $\int_0^1 \sqrt{\log N(\epsilon, \mathcal{G}, \rho_2)}\, d\epsilon < \infty$ is satisfied. □

Lemma 3. *It holds that*

$$\lim_{n\to\infty} \sup_{||g||_{(2)} \le 1} \Big| n^{-1} \sum_{k=1}^n g(k/n) - \lambda(g) \Big| = 0.$$

Proof. We have

$$I_n := \frac{1}{n} \sum_{k=1}^n g(k/n) - \lambda(g) = \sum_{k=1}^n \int_{(k-1)/n}^{k/n} [g(k/n) - g(s)]\, \mathrm{d}s$$

For every $s \in [(k-1)/n, k/n]$,

$$|g(k/n) - g(s)| \le v_2^{1/2}(g, [(k-1)/n, k/n]).$$

Hence,

$$|I_n| \le \sum_{k=1}^n \int_{(k-1)/n}^{k/n} v_2^{1/2}(g, [(k-1)/n, k/n])\, \mathrm{d}\, s$$
$$= \frac{1}{n} \sum_{k=1}^n v_2^{1/2}(g, [(k-1)/n, k/n]) \le \frac{1}{\sqrt{n}} \Big(\sum_{k=1}^n v_2(g, (k-1)/n, k/n]) \Big)^{1/2}$$
$$\le \frac{1}{\sqrt{n}} \|g\|_{(2)}$$

and this completes the proof. □

Proof of Theorem 1 $(1b)$**.** Under H_A,

$$\langle X_k, \psi \rangle \lambda_{nk}(f) = \langle Y_k, \psi \rangle \lambda_{nk}(f) + \langle g_n(k/n, \cdot), \psi \rangle \lambda_{nk}(f)$$
$$= \langle Y_k, \psi \rangle \lambda_{nk}(f) + n^{-1/2} u(k/n) [\langle a, \psi \rangle + \langle a_n, \psi \rangle].$$

Hence,

$$\nu_n(f,\psi) = \widehat{\nu}_n(f,\psi) + \Delta_n(f,\psi) + r_n(f,\psi),$$

where

$$\widehat{\mu}_n(f,\psi) = \sum_{i=1}^{n} \langle Y_i, \psi\rangle \lambda_{ni}(f),$$

$$\Delta_n(f,\psi) = n^{-1/2} \sum_{k=1}^{n} u(i/n)\lambda_{ni}(f)\langle q,\psi\rangle$$

and

$$r_n(f,\psi) = n^{-1/2} \sum_{k=1} u(i/n)]\lambda_{ni}(f)\langle q_n,\psi\rangle.$$

We have by $(1a)$

$$n^{-1/2}\widehat{\nu}_n \xrightarrow[n\to\infty]{\mathcal{D}} \nu_\gamma \text{ in the space } \ell^\infty(\mathcal{F}\times\Psi).$$

To complete the proof, we have to check

$$\lim_{n\to\infty} \sup_{f\in\mathcal{F},\psi\in\Psi} |n^{-1/2}\Delta_n(f,\psi) - \Delta(f,\psi)| = 0. \tag{44}$$

and

$$\lim_{n\to\infty} \sup_{f\in\mathcal{F},\psi\in\Psi} |n^{-1/2} r_n(f,\psi)| = 0. \tag{45}$$

To this aim, we involve lemma 3. We have

$$n^{-1/2}\Delta_n(f,\psi) - \Delta(f,\psi) = [I_{n1}(f) + I_{n2}(f)]\langle q,\psi\rangle,$$

where

$$I_{n1}(f) = \frac{1}{n}\sum_{k=1}^{n} u(k/n)[\lambda_{nk}(f) - f(k/n)],\ \ I_{n2}(f) = \frac{1}{n}\sum_{k=1}^{n} u(k/n)f(k/n) - \langle u,f\rangle.$$

By Lemma 3 applied to the function uf, we have $I_{n2}(f) \to 0$ uniformly over $f \in \mathcal{F}$. Consider I_{n1}. We have, as in the proof of Lemma 3,

$$|I_{n1}(f)| \le \sum_{k=1}^{n} |u(k/n)| \int_{(k-1)/n}^{k/n} [f(s) - f(k/n)|\, \mathrm{d}s \le n^{-1/2}\|u\|_\infty \|f\|_{(2)}.$$

Hence, $I_{n2}(f) \to 0$ uniformly over $f \in \mathcal{F}$. The convergence (45) follows by observing that

$$|n^{-1/2} r_n(f,\psi)| \le \|u\|_\infty \int_0^1 |f(s)|\,\mathrm{d}s \|\psi\| \cdot \|q_n\|.$$

This proves (45) and completes the proof of (1b). □

Proof of Theorem 3 $(3a)$**.** Consider the map $T : \ell^\infty(\mathcal{F}) \to \ell^\infty(\mathcal{F}), T(x)(f) = x(f) - x(1)\lambda(f)$. The continuity of T is easy to check. Observing that $\widehat{\nu}_n = T(\nu_n)$, the convergence (9) is a corollary of Theorem 1 and a continuous mapping theorem.

To prove $(3b)$, observe that, under H_A,

$$\begin{aligned}\langle X_k,\psi\rangle\lambda_{nk}(f) &= \langle Y_k,\psi\rangle\lambda_{nk}(f) + \langle g_n(k/n,\cdot),\psi\rangle\lambda_{nk}(f) \\ &= \langle Y_k,\psi\rangle\lambda_{nk}(f) + n^{-1/2}u(k/n)[\langle v,\psi\rangle + \langle v_n,\psi\rangle]\end{aligned}$$

hence

$$\mu_n(f,\psi) = \widehat{\mu}_n(f,\psi) + \widetilde{\Delta}_n(f,\psi) + \widetilde{r}_n(f,\psi),$$

where

$$\widehat{\mu}_n(f,\psi) = \sum_{i=1}^{n} \langle Y_i - \overline{Y}_n, \psi \rangle \lambda_{ni}(f),$$

$$\widetilde{\Delta}_n(f,\psi) = n^{-1/2} \sum_{k=1}^{n} [u(i/n) - n^{-1} \sum_{j=1}^{n} u(j/n)] \lambda_{ni}(f) \langle q, \psi \rangle$$

and

$$\widetilde{r}_n(f,\psi) = n^{-1/2} \sum_{k=1} [u(i/n) - n^{-1} \sum_{j=1}^{n} u(j/n)] \lambda_{ni}(f) \langle q_n, \psi \rangle.$$

We have by (a)

$$n^{-1/2}\widehat{\mu}_n \xrightarrow[n\to\infty]{\mathcal{D}} \mu_\gamma \text{ in the space } \ell^\infty(\mathcal{F} \times \Psi).$$

To complete the proof, we have to check

$$\lim_{n\to\infty} \sup_{f\in\mathcal{F},\psi\in\Psi} |n^{-1/2}\widetilde{\Delta}_n(f,\psi) - \widetilde{\Delta}(f,\psi)| = 0. \tag{46}$$

and

$$\lim_{n\to\infty} \sup_{f\in\mathcal{F},\psi\in\Psi} |n^{-1/2}\widetilde{r}_n(f,\psi)| = 0. \tag{47}$$

For this, we can use (44) and (45) and observe that $n^{-1}\sum_{k=1}^{n} u(k/n) \to \lambda(u)$ as $n \to \infty$. □

Author Contributions: Conceptualization, A.R.; Methodology, A.R.; Software, T.D.; Supervision, A.R.; Formal analysis, T.D. and A.R.; Writing—original draft preparation, A.R., T.D.; Investigation, T.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Csörgő, M.; Horváth, L. *Limit Theorems in Change-Point Analysis*; John Wiley & Sons: New York, NY, USA, 1997.
2. Brodsky, B.E.; Darkhovsky, B.S. *Non Parametric Methods in Change Point Problems*; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1993.
3. Basseville, M.; Nikiforov, N. *The Detection of Abrupt Changes—Theory and Applications*; Information and System Sciences Series; Prentice-Hall: Upper Saddle River, NJ, USA, 1993.
4. Chen, J.; Gupta, A.K. *Parametric Statistical Change Point Analysis*; Birkhauser Boston, Inc.: Boston, MA, USA, 2000.
5. Berkes, I.; Gabrys, R.; Horváth, L.; Kokoszka, P. Detecting changes in the mean of functional observations. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* **2009**, *71*, 927–946. [CrossRef]
6. Aue, A.; Gabrys, R.; Horváth, L.; Kokoszka, P. Estimation of a change-point in the mean function of functional data. *J. Multivar. Anal.* **2009**, *100*, 2254–2269. [CrossRef]
7. Aston, J.A.; Kirch, C. Detecting and estimating changes in dependent functional data. *J. Multivar. Anal.* **2012**, *109*, 204–220. [CrossRef]
8. Aue, A.; Rice, G.; Sönmez, O. Detecting and dating structural breaks in functional data without dimension reduction. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* **2018**, *80*, 509–529. [CrossRef]
9. Harris, T.; Li, B.; Tucker, J.D. Scalable Multiple Changepoint Detection for Functional Data Sequences. *arXiv* **2021**, arXiv:2008.01889v2.
10. Dudley, R.M.; Norvaiša, R. *Concrete Functional Analysis*; Springer: New York, NY, USA, 2011.
11. Horváth, L.; Kokoszka, P. *Inference for Functional Data with Applications*; Springer: Berlin, Germany, 2012.
12. Danielius, T.; Račkauskas, A. *p*-Variation of CUSUM process and testing change in the mean. *Commun. Stat. Simul. Comput.* **2020**, 1–13. [CrossRef]

13. der Vaart, A.W.V.; Wellner, J.A. *Weak Convergence and Empirical Processes with Applications to Statistics*; Springer: Berlin, Germany, 1996.
14. Norvaiša, R.; Račkauskas, A. Convergence in law of partial sum processes in p-variation norm. *Lith. Math. J.* **2008**, *48*, 212–227. [CrossRef]
15. Dudley, R.M. Sample functions of the gaussian process. *Ann. Probab.* **1973**, *1*, 66–103. [CrossRef]