Detection of a structural break in intraday volatility pattern

Detection of a structural break in intraday volatility pattern

Piotr Kokoszka*
Colorado State University

Tim Kutta
Colorado State University

Neda Mohammadi
Colorado State University

Haonan Wang
Colorado State University

Shixuan Wang
University of Reading

Abstract

We develop theory leading to testing procedures for the presence of a change point in the intraday volatility pattern. The new theory is developed in the framework of Functional Data Analysis. It is based on a model akin to the stochastic volatility model for scalar point-to-point returns. In our context, we study intraday curves, one curve per trading day. After postulating a suitable model for such functional data, we present three tests focusing, respectively, on changes in the shape, the magnitude and arbitrary changes in the sequences of the curves of interest. We justify the respective procedures by showing that they have asymptotically correct size and by deriving consistency rates for all tests. These rates involve the sample size (the number of trading days) and the grid size (the number of observations per day). We also derive the corresponding change point estimators and their consistency rates. All procedures are additionally validated by a simulation study and an application to US stocks.

MSC 2020 subject classifications: 62R10, 62G10, 62M10.
Keywords and phrases: Change point, Functional data, Intraday volatility, Itô integral.

1 Introduction

Consider a sample of intraday price curves $\{P_i(t), t \in [0,1]\}$, $1 \le i \le N$, where i indexes the trading day and t is intraday time normalized to the standard unit interval. For each i, we study the limits, as $\Delta \to 0$, of cumulative intraday realized volatility curves

$$\mathrm{RV}_i(\Delta)(t) = \sum_{1 \le k \le Kt} \left|\log[P_i(k\Delta)] - \log[P_i((k-1)\Delta)]\right|^2, \quad t \in [0,1]. \tag{1.1}$$

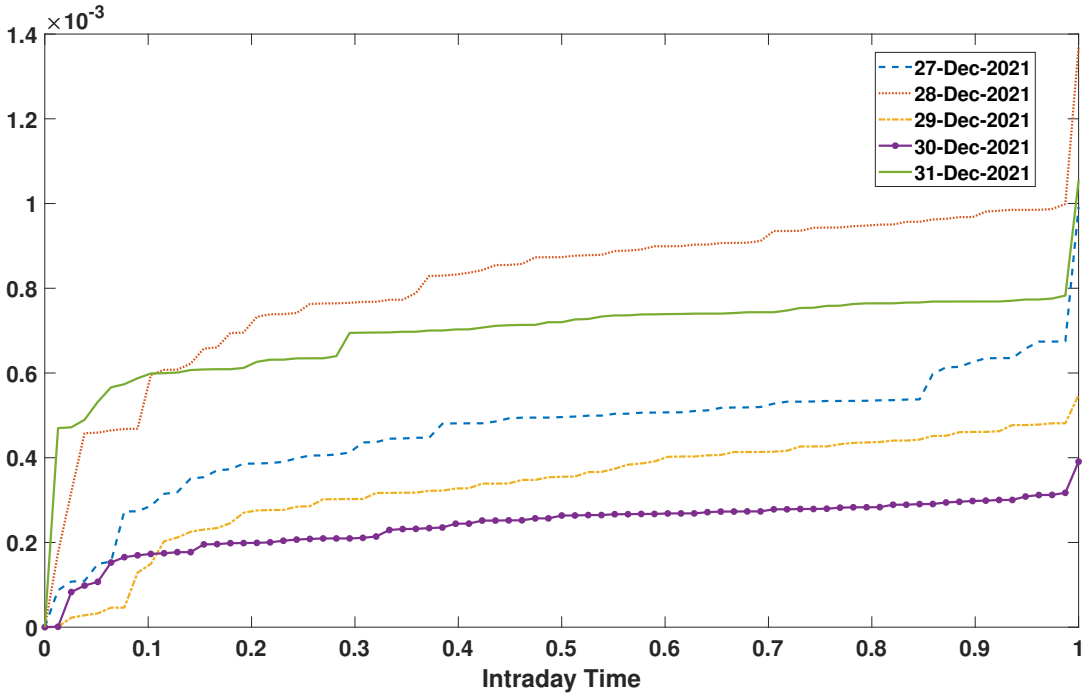

Figure 1: Five consecutive realized volatility curves computed with $K = 78$ for Tesla Inc. from Dec 27 to Dec 31, 2021.

To illustrate, five consecutive curves $\mathrm{RV}_i(\Delta)(\cdot)$ are shown in Figure 1.

Under suitable assumptions, see Section 2, for each $t \in [0, 1]$,

$$\mathrm{RV}_i(\Delta)(t) \xrightarrow{P} \int_0^t \nu_i^2(u)du, \quad \text{as } \Delta \to 0.$$

The object of our study are basically the curves ν_i, but a more precise problem statement is needed. We represent the curves ν_i as

$$\nu_i(u) = h_i\sigma_i(u), \quad u \in [0, 1], \ i \in \mathbb{Z}, \tag{1.2}$$

where the $h_i > 0$ describe the evolution of the curves from day to day (between-day volatility), while the functions σ_i quantify the residual volatility after the between days volatility has been accounted for by the sequence $\{h_i\}$. The identifiability of the components in decomposition (1.2) is addressed in Lemma 3.1. *We develop a statistical framework to test if the functions σ_i change over a time period of many days.* The volatilities h_i typically exhibit persistent magnitude clusters. We propose methodology, and supporting theory, that allows us to test the constancy of the functions σ_i in index i. We emphasize that we do not test if each σ_i is a constant function because this is well-known

*Corresponding author: Department of Statistics, Colorado State University, Fort Collins, CO 80521-1877, USA Email: Piotr.Kokoszka@colostate.edu

not to be true. The functions σ_i are approximately the derivatives of functions like those in Figure 1, divided by h_i. The shape of the functions σ_i can change, and such a change will be undetectable if one focuses only on realized daily volatilities $\mathrm{RV}_i(\Delta)$ because they are dominated by the h_i. In the following, we refer to the $\sigma_i(\cdot)$ as volatility or diffusion functions, while keeping in mind that the stochastic volatility functions are the products $h_i\sigma_i(\cdot)$.

Model (1.2) is a useful approximation that allows us to construct an effective test and justify it. Using a related perspective, [10] assume that $\nu_i(u) = h_i(u)\sigma_i(u)$ and propose a test of $H_0 : h_i(u) = h_i$, $\forall\ u \in [0,1]$. They apply it to over thousand days i and to 30 stocks, about 30,000 tests in total. There are overall more rejections than acceptances of their H_0, the results depend on the implementation of the test. The rejections dominate if the test is implemented with a good estimate of the function σ^2, under the assumption that it does not change, i.e.

$$\sigma_i^2(\cdot) = \sigma^2(\cdot), \quad \forall\ i = 1, 2, \ldots, N. \tag{1.3}$$

If condition (1.3) does not hold, any estimate of the function $\sigma^2(\cdot)$ may be meaningless. Our test is thus complementary to that of [10] and impacts its implementation; we test assumption (1.3) with an unknown $\sigma^2(\cdot)$. We evaluate and apply the test to five minute intraday returns, so we do not need to be concerned with microstructure noise and price jumps. Our objective is to provide and effective principled way of testing condition (1.3) within a broad framework of statistical change point detection and functional data analysis. Our tests will be useful in any context that requires the verification that an intraday volatility pattern remains constant over the period of many trading days.

Research on the detection and estimation of a change point in various statistical models is over 70 years old and forms a well-established subfield of statistics. Its importance stems from the fact that most statistical models assume a single data generating mechanism, so if this mechanism changes over the observational period, their application will be meaningless. There are several monographs and thousands of research papers; the paper of [21] can serve as a concise and modern introduction to the general framework of this paper. It includes change point detection for different data structures, among them functional data. A broad review of inference and estimation techniques is given in [**?**], who also consider a range of different applications. The related literature on sequential (online) tests for structural changes is reviewed in [**?**]. In the following two paragraphs, we briefly review the most closely related research. An important point to note is that in the framework of functional data analysis (FDA), the tests are $\sqrt{N}$-consistent (N is the sample size), whereas in our framework of replications of a diffusion process, we can obtain $\sqrt{NK}$-consistency, where K is the number of sampling points for each replication. Moreover, these rates depend on the type of local alternative, shape change vs. size change, and arise because we explicitly use modeling though Itô integrals. A key starting point is new concentration results, Proposition A.1 and its corollaries, that can be used in other contexts that require information about the rates, in terms of grid size, at which population volatility functions can be approximated by realized volatility curves. As far as

we know, the framework we study has been considered neither in change point research nor in intraday volatility research. We hope that the theoretical advances we make together with a comprehensive application will motivate further research at the nexus of FDA and SDEs.

Related change point research in the framework of FDA In FDA, the observations are random elements in some function space, such the space L^2 equipped with the canonical L^2-norm, or $\mathcal{C}$ (continuous functions) equipped with the supremum norm. The space L^2 has played a particularly important role in FDA since, under weak assumptions, it is a separable Hilbert space. [7] proposed a test for a change in the mean function in an L^2 setting; extensions were considered by [3], [20], [17] and [6], among others. Structural breaks of time series in the space $\mathcal{C}([0,1])$ were studied in [13], see also [36] for a more abstract context. Changes in the covariance of a functional time series were considered in [41] and in the cross-covariance operator in [37]. [12] considered inference for the covariance kernel of continuous data. More recently, [22] employed a weighted CUSUM statistic for the detection and localization of changes in the covariance operator. Tests for the stability of eigenvalues and principal components were presented in [5] and [15]. A self-normalization approach, see [39] and [38], was used in the context of change point detection in functional time series by [14]. The cited works comprise only a small fraction of the relevant literature.

Change point detection in Itô semimartingales We are not aware of any research that consideres a change point problem in a sample of trajectories, each of which is an Itô semimartingale. Research to date has focused on the detection of a change point in a single continuous–time realization. [1] modified the commonly used CUSUM approach to detect jumps in Itô semimartingales. In particular, in order to detect jumps in asset returns, they proposed a test statistic based on multiplicative difference of realized truncated p–th variation. [9] used a similar approach to detect structural changes in the volatility of Itô semimartingales. They addressed detection of jumps, the so called local changes, as well as changes in the roughness of sample path, the so called global changes. Other related papers are [19], [18], and [8] who provide further references to the general area of change point detection in Itô semimartingales.

The remainder of the paper is organized as follows. In Section 2, we collect the minimum required background on Itô integrals and the FDA. Section 3 is dedicated to the precise formulation of the problem outlined above. Testing and estimation approaches are developed in Section 4. Their finite sample properties are examined in Section 5. Section 6 contains an application to sequences of intraday returns on US stocks. Supplementary material contains proof of all results stated in Section 4, details of practical implementation of all procedures, and some additional information.

2 Mathematical preliminaries

We begin by providing some mathematical background, beginning with stochastic differential equations. We first recall the definition of the quadratic variation of a stochastic

process $\{X(t) : t \in [0,1]\}$. We assume throughout that $0 = t_0^{(K)} < t_1^{(K)} < \cdots < t_K^{(K)} = 1$ is a grid on the interval $[0,1]$ with $\max_k [t_k - t_{k-1}] \to 0$, as $K \to \infty$. Then,

$$\sum_{k=1}^{K} \left| X(t_k^{(K)}) - X(t_{k-1}^{(K)}) \right|^2 \mathbb{I}\{t_k^{(K)} \le t\}, \xrightarrow{P} [X,X]_t \quad t \in [0,1], \tag{2.1}$$

The limit $[X,X]_t$ exists for any semimartingale, and is called the the quadratic variation at time t, see e.g. Theorem 1.14 and relation (3.23) in [2]. In this work, we assume that the process X is given by the Itô integral

$$X(t) := \int_0^t \nu(u) dW(u),$$

where W is a standard Wiener process and $\nu : [0,1] \to (0,\infty)$ is a continuous function (for a detailed discussion of the existence and properties of this process, see Theorem 5.2.1 in [34]). It is well-known that for Itô integrals the quadratic variation is given by

$$[X,X]_t := \int_0^t \nu^2(u) du, \quad t \in [0,1], \tag{2.2}$$

see e.g. equation (2.1) in [25]. Moreover, in this case, (2.1) can be strengthened to

$$\sup_{0 \le t \le 1} \left| \sum_k \left| X(t_k^{(K)}) - X(t_{k-1}^{(K)}) \right|^2 \mathbb{I}\{t_k^{(K)} \le t\} - \int_0^t \nu^2(u) du \right| \xrightarrow{P} 0, \quad K \to \infty. \tag{2.3}$$

see again Theorem 1.14 and relation (3.23) in [2]. Next, we present a consequence of the Dambis–Dubins–Schwarz theorem which states that *any* continuous local martingale can be expressed as a time change of a Brownian motion, see e.g. Section 5.3.2 in [30] for a general statement. In our setting,

$$\left\{ \int_0^t \nu(u) dW(u),\ t \in [0,1] \right\} \stackrel{d}{=} \left\{ W\left(\int_0^t \nu^2(u) du \right),\ t \in [0,1] \right\}, \tag{2.4}$$

where the equality in distribution is in the space $C([0,1])$ of continuous functions, equipped with the topology of uniform convergence. Identity (2.4) entails

$$\mathbb{E}\left[\int_0^t \nu(u) dW(u) \right]^2 = \int_0^t \nu^2(u) du, \tag{2.5}$$

which is a special case of the Itô isometry for a deterministic, square integrable integrand $\nu(\cdot)$.

In identities (2.4) and (2.5), the Itô process is treated as a random function in $C([0,1])$. However, in the context of FDA, it is often useful to embed the smaller space of continuous functions in the larger Hilbert space of square integrable functions. More precisely,

we define $L^2([0,1])$ as the space of measurable functions $f : [0,1] \to \mathbb{R}$ that satisfy $\int_0^1 f^2(x)dx < \infty$. Equipped with the inner product

$$\langle f, g\rangle := \int_0^1 f(t)g(t)dt, \qquad f, g \in L^2([0,1]),$$

and the induced norm $\|\cdot\|_{L^2}$, $L^2([0,1])$ becomes a separable Hilbert space, where we identify functions equal almost everywhere. A random function X in $L^2([0,1])$ is then a measurable map $X : (\Omega, \mathcal{A}, \mathbb{P}) \to L^2([0,1])$, where $(\Omega, \mathcal{A}, \mathbb{P})$ is a probability space. If the first moment of X exists in the sense that $\mathbb{E}\|X\| < \infty$, we can define the expectation $\mu \in L^2([0,1])$ of X, which is characterized by the identity

$$\mathbb{E}\langle X, f\rangle = \langle \mu, f\rangle \qquad \forall f \in L^2([0,1]).$$

Similarly, if the second moment of X exists, $\mathbb{E}\|X\|^2 < \infty$, we can define the covariance operator $C_X : L^2([0,1]) \to L^2([0,1])$ of X by the identity

$$\langle C_X[f], g\rangle := \mathbb{E}\big[\langle X - \mu, f\rangle\langle X - \mu, g\rangle\big] \qquad \forall f, g \in L^2([0,1]).$$

It is known that C_X is a self–adjoint, positive semi-definite, Hilbert-Schmidt operator and as such it can be identified with a square integrable kernel function $c_X : [0,1]^2 \to \mathbb{R}$ via

$$C_X[f](x) := \int_0^1 c_X(x,y)f(y)dy \qquad \forall f \in L^2([0,1]).$$

Chapters 10 and 11 of [27] provide a concise introduction to the L^2 framework of FDA. For a comprehensive treatment see [23].

3 Statistical model and problem formulation

Suppressing the superscript (K), consider the grid $0 = t_0 < t_1 < \cdots < t_K = 1$ introduced in Section 2, and the cumulative returns

$$R_i(t_k) = \log[P_i(t_k)] - \log[P_i(0)].$$

The realized volatility curves (1.1) for this grid can be written as

$$\mathrm{RV}_i(t) = \sum_{k=1}^{K} |R_i(t_k) - R_i(t_{k-1})|^2 \, \mathbb{I}\{t_k \le t\}.$$

Setting $h_i = \exp(g_i)$ in (1.2), we postulate the model

$$R_i(t) = \exp(g_i) \int_0^t \sigma_i(u)dW_i(u), \quad t \in [0,1], \quad i \in \mathbb{Z}. \tag{3.1}$$

The $W_i(\cdot)$ are independent standard Wiener processes. The sequence g_i is a centered real–valued, weakly stationary time series independent of $(W_i)_{i\in\mathbb{Z}}$. The following lemma shows that for each i the volatility function $\sigma_i(\cdot)$ depends only on $R_i(\cdot)$, so g_i and $\sigma_i(\cdot)$ are identifiable.

LEMMA 3.1 *Suppose g satisfies $Eg = 0$ and is independent of the Wiener process $W(\cdot)$. Setting*

$$R(t) = e^g \int_0^t \sigma(u)dW(u), \quad t \in [0,1],$$

for a continuous function $\sigma(\cdot)$, we have

$$\int_0^t \sigma^2(u)du = \exp\left\{\mathbb{E}\log[R,R]_t\right\}.$$

PROOF. By (2.2), $[R,R]_t = \exp(2g)\int_0^t \sigma^2(u)du$, so

$$\log[R,R]_t = 2g + \log\int_0^t \sigma^2(u)du.$$

Since $\mathbb{E}(g) = 0$, $\mathbb{E}\log[R,R]_t = \log\int_0^t \sigma^2(u)du$, and the claim follows. ∎

To test for changes in the volatility functions $\sigma_i(\cdot)$, we propose the following change point model. Let $\theta \in (0,1)$ be a parameter that locates a potential change in the discrete time index i and let $\sigma_{(1)}, \sigma_{(2)} : [0,1] \to (0,\infty)$ denote two continuous volatility functions. We postulate that

$$\begin{cases} \sigma_i(\cdot) = \sigma_{(1)}(\cdot), & \text{for } i \le \lfloor N\theta \rfloor, \\ \sigma_i(\cdot) = \sigma_{(2)}(\cdot), & \text{for } i > \lfloor N\theta \rfloor. \end{cases} \tag{3.2}$$

A change occurs if $\sigma_{(1)}(\cdot) \neq \sigma_{(2)}(\cdot)$. The testing problem is thus

$$H_0 : \sigma_{(1)}(\cdot) = \sigma_{(2)}(\cdot), \qquad \text{vs.} \qquad H_A : \sigma_{(1)}(\cdot) \neq \sigma_{(2)}(\cdot). \tag{3.3}$$

A cornerstone of our statistical analysis is the translation of changes in volatility to changes of certain features in the quadratic variation process. This allows us to take advantage of regularities of the quadratic variation process compared to the process $R_i(\cdot)$. Indeed, (2.2) directly entails

$$Q_i(t) := [R_i,R_i]_t = \exp(2g_i)\int_0^t \sigma_i^2(u)du, \quad t \in [0,1], \quad i = 1,2,\ldots,N. \tag{3.4}$$

This, together with the stationarity of the time series $(g_i)_{i\in\mathbb{Z}}$, indicates that a change in volatility corresponds to a change in the distribution of the quadratic variation process $Q_i(\cdot)$, over index i.

Suppose we observe a sample $R_1,\ldots,R_N$. Reflecting practically available data, we assume that the functions R_i are observed at $K+1$ equidistant points in $[0,1]$. This means that inference will be based on the matrix of observations

$$\{R_i(k/K) : i = 1,\ldots,N,\ k = 0,\ldots,K\}. \tag{3.5}$$

In our theory, we assume that the number of grid points, $K+1$, as well as the number of curves, N, tend to infinity. In view of approximations (2.1) and (2.3), we consider the realized quadratic variation processes

$$\begin{aligned}(3.6)\qquad \widehat{Q}_i(t) &= \sum_{k=1}^{K} |R_i(k/K) - R_i((k-1)/K)|^2 \mathbb{I}\{k/K \le t\} \\ &= \exp(2g_i) \sum_{k=1}^{K} \left| \int_{(k-1)/K}^{k/K} \sigma(u) dW_i(u) \right|^2 \mathbb{I}\{k/K \le t\}, \qquad t \in [0,1].\end{aligned}$$

as estimators of the $Q_i(\cdot)$ in (3.4). Observe that $\widehat{Q}_i(t)$ is equal to the realized volatility function (1.1), with the second line reflecting the assumed model.

Assuming the g_i have exponential moments, a test could be based on the approximation

$$(3.7)\qquad \mathbb{E}[\widehat{Q}_i(t)] \approx \mathbb{E}[Q_i(t)] = \mathbb{E}[\exp(2g_i)] \cdot \int_0^t \sigma_i^2(u) du$$

which indicates that volatility function changes translate to mean changes in the realized quadratic variation process. Detecting changes in the mean of a functional time series is a well-studied problem, as discussed in Section 1. However, a test based on (3.7), requires the existence of exponential moments of the g_i and is not robust against distributional changes in g_i, which might be mistaken for changes in the volatility functions $\sigma_i(\cdot)$. Moreover, CUSUM based FDA tests are $\sqrt{N}$-consistent, but we demonstrate that against large classes of common alternatives a much stronger consistency rate of $\sqrt{NK}$ is attainable by some tests we propose. For these reasons, we present in this paper a different, more effective method to test the hypotheses (3.3). Our approach does not require exponential moments of g_i, is more stable against distributional changes (or spurious changes) in the g_i, and benefits from $\sqrt{NK}$-consistency under typical alternatives. As a first step, we express the hypothesis H_0 in (3.3) in terms of two null hypotheses, $H_0^{(1)}$ and $H_0^{(2)}$, that are together equivalent to H_0:

$$(3.8)\qquad H_0^{(1)} : \ \frac{\int_0^t \sigma_{(1)}^2(u) du}{\int_0^1 \sigma_{(1)}^2(u) du} = \frac{\int_0^t \sigma_{(2)}^2(u) du}{\int_0^1 \sigma_{(2)}^2(u) du}, \quad \forall\ t \in [0,1],$$

$$(3.9)\qquad H_0^{(2)} : \ \int_0^1 \sigma_{(1)}^2(u) du = \int_0^1 \sigma_{(2)}^2(u) du.$$

Heuristically, $H_0^{(1)}$ states that the volatility function does not change its shape, while $H_0^{(2)}$ states that the total volatility stays the same. In Section 4, we formulate statistical tests of $H_0^{(1)}$ and $H_0^{(2)}$ separately, and then combine them to test the H_0 in (3.3).

4 Change point tests

We begin by stating assumptions for our subsequent analysis.

Assumption 4.1

1. The volatility function $\sigma_{(1)}, \sigma_{(2)} : [0,1] \to (0,\infty)$ are continuous.
2. The standard Wiener processes W_i, $i \in \mathbb{Z}$, are independent.
3. The two sequences $(g_i)_{i\in\mathbb{Z}}$ and $(W_i)_{i\in\mathbb{Z}}$ are independent of each other.
4. The time series $(g_i)_{i\in\mathbb{Z}}$ is centered, weakly stationary and satisfies a weak invariance principle of the form

$$\Big\{\frac{1}{\sqrt{N}}\sum_{i=1}^{\lfloor Nx\rfloor} g_i : x \in [0,1]\Big\} \xrightarrow{d} \{\lambda W(x) : x \in [0,1]\},$$

where W is a standard Wiener process and $\lambda^2 > 0$ denotes the long-run variance.

Assumption 4.1 is satisfied in many different scenarios. It basically postulates a very general functional stochastic volatility model. Condition 1 (before and after the change) is common in the literature on diffusion processes and intuitive in our setting. Conditions 2 and 3 determine the dependence structure along our functional time series, which is moderated by the scaling factors e^{g_i}. In [26], the g_i follow an AR(p) model, but for our theory the precise dependence structure is immaterial. If the dependence is sufficiently weak, the partial sum process on the left-hand side of condition 4 converges to a Wiener process. This is true under a multitude of dependence conditions, see e.g. [31], so instead of choosing some of them, we postulate the general condition 4.

4.1 Inference for a shape change

Recall the hypothesis $H_0^{(1)}$ in (3.8). We begin with the simple observation that, according to (3.4), the functions in (3.8) can be represented by the *standardized quadratic variation* as follows:

$$F_i(t) := \frac{Q_i(t)}{Q_i(1)} = \frac{\int_0^t \sigma^2_{(j)}(u)du}{\int_0^1 \sigma^2_{(j)}(u)du}, \quad \text{with } j = \begin{cases} 1, & \text{for} \quad i = 1,\dots,\lfloor N\theta\rfloor, \\ 2, & \text{for} \quad i = \lfloor N\theta\rfloor + 1,\dots,N. \end{cases} \tag{4.1}$$

This motivates using for statistical inference the empirical versions:

$$\begin{aligned} \widehat{F}_i(t) :=& \frac{\widehat{Q}_i(t)}{\widehat{Q}_i(1)} := \frac{\sum_{k=1}^K |R_i(k/K) - R_i((k-1)/K)|^2 \,\mathbb{I}\{k/K \le t\}}{\sum_{k=1}^K |R_i(k/K) - R_i((k-1)/K)|^2} \\ =& \frac{\sum_{k=1}^K \left|\int_{(k-1)/K}^{k/K} \sigma_i(u)dW_i(u)\right|^2 \mathbb{I}\{k/K \le t\}}{\sum_{k=1}^K \left|\int_{(k-1)/K}^{k/K} \sigma_i(u)dW_i(u)\right|^2}. \end{aligned} \tag{4.2}$$

REMARK 4.1 We highlight two useful properties of $\widehat{F}_i$:

i) $\widehat{F}_i$ is monotonically increasing with $\widehat{F}_i(0) = 0$ and $\widehat{F}_i(1) = 1$ and in particular it is a random cdf (and thus measurable). It can be interpreted as a random function, mapping into the space $L^2[0,1]$ of square integrable functions on the unit interval.

ii) The functions $\widehat{F}_1, \ldots, \widehat{F}_N$ are independent, and they do not involve the g_i.

Property ii) implies that any test statistic based on the $\widehat{F}_i$s will be unaffected by the structure of the g_i, or a potentially misspecified model for them.

LEMMA 4.1 *Suppose that Conditions 1 and 2 of Assumption 4.1 hold. Then, each $\widehat{F}_i$ is a consistent estimator of the standardized quadratic variation F_i (defined in (4.1)) and satisfies a functional central limit theorem of the form*

$$\sqrt{K}\{\widehat{F}_i(\cdot) - F_i(\cdot)\} \xrightarrow{d} Z(\cdot) \tag{4.3}$$

where Z is a centered, Gaussian process in $L^2([0,1])$, with distribution depending on the volatility function σ_i.

The proof of Lemma 4.1 follows by an application of Theorem 5.3.6 in [24] together with the continuous mapping theorem. In view of the convergence in (4.3), we expect a test statistic based on $\widehat{F}_1, \ldots, \widehat{F}_N$ to have variance of order $O(1/(NK))$, or a corresponding test for $H_0^{(1)}$ to be $\sqrt{NK}$-consistent.

To test $H_0^{(1)}$, we use the CUSUM statistic

$$\widehat{S}^{(1)} := \frac{K}{N^2} \sum_{n=1}^{N} \int_0^1 \Big(\sum_{i=1}^{n} \widehat{F}_i(u) - \frac{n}{N} \sum_{i=1}^{N} \widehat{F}_i(u) \Big)^2 du. \tag{4.4}$$

In the following result, the asymptotics "$N, K \to \infty$" should be understood in terms of a sequence $K = K_N$ of natural numbers that diverges as $N \to \infty$. We do not impose any restrictions on the growth rate of K relative to N, making our method valid regardless of the interplay between K and N.

THEOREM 4.1 *Suppose that Conditions 1 and 2 of Assumption hold and that $N, K \to \infty$. Then, under $H_0^{(1)}$, the weak convergence*

$$\widehat{S}^{(1)} \xrightarrow{d} S^{(1)} := \sum_{\ell=1}^{\infty} \lambda_\ell \int_0^1 \mathbb{B}_\ell(u)^2 du \tag{4.5}$$

holds, where $(\mathbb{B}_\ell)_{\ell \in \mathbb{N}}$ is a sequence of i.i.d. Brownian bridges and $(\lambda_\ell)_{\ell \in \mathbb{N}}$ the collection of eigenvalues of the asymptotic covariance kernel

$$c_F(u,v) := \lim_{K \to \infty} K \cdot \mathbb{E}\left[\{\widehat{F}_1(u) - \mathbb{E}[\widehat{F}_1(u)]\} \{\widehat{F}_1(v) - \mathbb{E}[\widehat{F}_1(v)]\} \right]. \tag{4.6}$$

Moreover, if $H_0^{(1)}$ is violated, $\widehat{S}^{(1)} \xrightarrow{\mathbb{P}} \infty$.

Denoting for any $\alpha \in (0,1)$ the upper α-quantile of $S^{(1)}$ by $q^{(1)}_{1-\alpha}$, the decision

$$\text{reject if } \ \widehat{S}^{(1)} > q^{(1)}_{1-\alpha}$$

yields a consistent asymptotic level α test of $H^{(1)}_0$. While in practice, we do not know the distribution of $S^{(1)}$, it is uniquely determined by the eigenvalues of c_F, which can be estimated by off-the-shelf methods (we provide details in Appendix C). An explicit formula for the kernel $c_F(u,v)$ is given in Theorem B.1, but it is not needed to estimate the λ_ℓ because (4.6) is a limit of covariance kernels, and many FDA packages output their eigenvalues. The distribution of the integral in (4.5) is easy to simulate, and it is fairly well-known how to compute the approximate quantiles of the right-hand side of (4.5).

In the next theorem, we demonstrate the consistency of our test procedure against local alternatives.

THEOREM 4.2 *Suppose Conditions 1 and 2 of Assumption hold. Let $\tilde{\sigma} : [0,1] \to (0,\infty)$ be a continuous function such that $\tilde{\sigma}(\cdot)/\sigma_{(1)}(\cdot)$ is not constant and let $(a_N)_{N\in\mathbb{N}}$ be a bounded sequence of positive numbers. Then, defining $\sigma_{(2)} := \sigma_{(1)} + a_N\tilde{\sigma}$ and imposing the growth conditions $a_N\sqrt{NK} \to \infty$ and $a_N K \to \infty$, it follows that*

$$\lim_{N,K\to\infty} \mathbb{P}(\widehat{S}^{(1)} > c) = 1 \quad \forall\ c \geq 0.$$

Finally, we define the change point estimator

$$\hat{\theta}^{(1)} := \frac{1}{N} \operatorname*{argmax}_{n\in\{1,\dots,N\}} \int_0^1 \Big(\sum_{i=1}^{n} \widehat{F}_i(u) - \frac{n}{N} \sum_{i=1}^{N} \widehat{F}_i(u) \Big)^2 du. \tag{4.7}$$

If the hypothesis of no change in the shape of volatility is violated, i.e.

$$\begin{cases} F_i(\cdot) = F_{(1)}(\cdot), & \text{for } i \leq \lfloor N\theta \rfloor \\ F_i(\cdot) = F_{(2)}(\cdot), & \text{for } i > \lfloor N\theta \rfloor \end{cases}, \quad F_{(1)}(\cdot) \neq F_{(2)}(\cdot) \tag{4.8}$$

for some $\theta \in (0,1)$, we can show that the estimator $\hat{\theta}^{(1)}$ in (4.7) is consistent under local alternatives.

THEOREM 4.3 *Under the assumptions of Theorem 4.2,*

$$\hat{\theta}^{(1)} - \theta = \mathcal{O}_P\Big(\max\Big\{ \frac{a_N^{-2}}{NK}, \frac{1}{N} \Big\} \Big),$$

where $\theta \in (0,1)$ is the rescaled time of the change in (4.8).

The rate in Theorem 4.3 can be explained as follows: For a change of size a_N (potentially tending to 0) it is well-known in change point estimation that an optimal approximation rate is given by

$$\hat{\theta}^{(1)} - \theta = \mathcal{O}_P\Big(\frac{a_N^{-2}}{\textit{sample size}} \Big).$$

In our case the "sample size" is NK, yielding a rate of $\mathcal{O}_P(a_N^{-2}/(NK))$. However, since the number of curves in discrete time is N, it is also clear that a convergence rate cannot be faster than $\mathcal{O}_P(1/N)$. This limitation is simply due to the discretization of time in N steps. As a consequence, the best attainable rate is as specified in Theorem 4.3. Notice that in the special case of $a_N = O(1/\sqrt{K})$, we obtain the optimal rate $O_P(1/N)$ on the right side, the same rate as for fully observed functions, see e.g. [4].

4.2 Inference for a change in total volatility

Recall the hypothesis $H_0^{(2)}$ in (3.9). The integrals in (3.9) are closely related to the total quadratic variation $Q_i(1)$ and taking its logarithm, we obtain

$$\log(Q_i(1)) = 2g_i + \log\Big(\int_0^1 \sigma_{(j)}^2(u)du\Big), \quad \text{for} \quad \begin{cases} i = 1, \dots, \lfloor N\theta \rfloor, & j = 1, \\ i = \lfloor N\theta \rfloor + 1, \dots, N, & j = 2. \end{cases} \tag{4.9}$$

Since the g_i are centered, any change in total volatility translates into a mean change of the real–valued time series $\{\log(Q_i(1))\}$. An empirical analogue of (4.9) is

$$\log(\widehat{Q}_i(1)) = 2g_i + w_i, \tag{4.10}$$

where

$$w_i := \log\Big(\sum_{k=1}^{K} \Big| \int_{(k-1)/K}^{k/K} \sigma_i(u)dW_i(u)\Big|^2\Big). \tag{4.11}$$

This decomposition shows that the observations $\log(\widehat{Q}_1(1)), \dots, \log(\widehat{Q}_N(1))$ form (for any fixed K) a dependent time series that is stationary before and after a potential change. For the purpose of statistical inference, we use the following CUSUM statistics:

$$\widehat{S}^{(2)} := \frac{1}{N^2} \sum_{n=1}^{N} \Big(\sum_{i=1}^{n} \log(\widehat{Q}_i(1)) - \frac{n}{N} \sum_{i=1}^{N} \log(\widehat{Q}_i(1)) \Big)^2. \tag{4.12}$$

THEOREM 4.4 *If Assumption 4.1 holds and $N, K \to \infty$, then, under $H_0^{(2)}$,*

$$\widehat{S}^{(2)} \xrightarrow{d} S^{(2)} := (4\lambda) \cdot \int_0^1 \mathbb{B}(u)^2 du, \tag{4.13}$$

where $\mathbb{B}$ is a standard Brownian bridge and the long-run variance λ is defined as

$$\lambda := \sum_{i \in \mathbb{Z}} \mathbb{C}ov(g_0, g_i). \tag{4.14}$$

Moreover, if $H_0^{(2)}$ is violated, $\widehat{S}^{(2)} \xrightarrow{\mathbb{P}} \infty$.

The fact that the long-run variance λ only depends on the g_is is not an accident. As we will see in the next section, the statistic $\widehat{S}^{(2)}$ is asymptotically only dependent on the g_is and independent of the W_is. This implies that $\widehat{S}^{(1)}$ (which does not depend on the g_is) and $\widehat{S}^{(2)}$ are asymptotically independent.

Theorem 4.4 implies that if we denote by $q^{(2)}_{1-\alpha}$ the upper α-quantile of the distribution $S^{(2)}$, then the decision to

$$\text{reject if } \ \widehat{S}^{(2)} > q^{(2)}_{1-\alpha}$$

yields a consistent asymptotic level α test of the hypothesis $H^{(2)}_0$. Again, $q^{(2)}_{1-\alpha}$ cannot be directly computed, but it can be approximated, if a consistent estimator for the long-run variance is given (see Appendix C).

We now show consistency of the test against local alternatives.

THEOREM 4.5 *Suppose Assumption 4.1 holds and $(a_N)_{N\in\mathbb{N}}$ is a bounded sequence of positive numbers. Then, defining $\sigma_{(2)} := (1 + a_N)\sigma_{(1)}$ and imposing the growth conditions $a_N\sqrt{N} \to \infty$ and $a_N K \to \infty$, it follows that*

$$\lim_{N,K\to\infty} \mathbb{P}(\widehat{S}^{(2)} > c) = 1, \quad \forall\ c \geq 0.$$

Finally, with the change point estimator

$$\hat{\theta}^{(2)} := \frac{1}{N} \operatorname*{argmax}_{n\in\{1,\dots,N\}} \Big(\sum_{i=1}^{n} \log(\widehat{Q}_i(1)) - \frac{n}{N} \sum_{i=1}^{N} \log(\widehat{Q}_i(1)) \Big)^2, \tag{4.15}$$

we can localize a change in total volatility. If the hypothesis of no change in the total volatility is violated, i.e.

$$\begin{cases} \log(Q_i(1)) = \log(Q_{(1)}(1)), & \text{for } i \leq \lfloor N\theta \rfloor \\ \log(Q_i(1)) = \log(Q_{(2)}(1)), & \text{for } i > \lfloor N\theta \rfloor \end{cases}, \quad \log(Q_{(1)}(1)) \neq \log(Q_{(2)}(1)), \tag{4.16}$$

for some $\theta \in (0, 1)$, we obtain the following result.

THEOREM 4.6 *Under the assumptions of Theorem 4.5, $\hat{\theta}^{(2)} - \theta = O_P(a_N^{-2}/N)$.*

Taking $a_N = 1$, we obtain the optimal rate.

4.3 Inference for an arbitrary change

In the previous subsections, we have developed test statistics $\widehat{S}^{(1)}, \widehat{S}^{(2)}$ for the null hypotheses $H^{(1)}_0, H^{(2)}_0$ in (3.8) and (3.9), respectively. We now want to combine these two tests to yield a test for the global null hypothesis H_0 in (3.3). As a first step, we show that as $N, K \to \infty$, t he two test statistics (4.4) and (4.12) become independent of each other.

PROPOSITION 4.1 *If Assumption 4.1 and H_0 in (3.3) hold, then, as $N, K \to \infty$,*

$$\left(\widehat{S}^{(1)}, \widehat{S}^{(2)}\right) \xrightarrow{d} \left(S^{(1)}, S^{(2)}\right),$$

where $S^{(1)}, S^{(2)}$ are independent and defined in (4.5), (4.13), respectively.

In order to combine the results from both test statistics, we employ their asymptotic p-values. To be precise, if $\Lambda^{(j)}$ is the (continuous) cumulative distribution function of $S^{(j)}$, we define the p-values

$$p^{(j)} = 1 - \Lambda^{(j)}\big(\widehat{S}^{(j)}\big), \quad j = 1, 2. \tag{4.17}$$

In practice, the $\Lambda^{(j)}$ are not known, but can uniformly approximated, yielding empirical p-values. We discuss this issue in Appendix C. To combine our test statistics, we recall that under H_0, both p-values $p^{(1)}, p^{(2)}$ are asymptotically uniformly distributed on $[0, 1]$ and according to Proposition 4.1 asymptotically independent. Hence, using Fisher's method, see e.g. [35], we can combine them to

$$\widehat{S} := -2\{\log(p^{(1)}) + \log(p^{(2)})\}, \tag{4.18}$$

which then converges under H_0 to a chi-squared distribution with four degrees of freedom. Denoting the upper α-quantile of this distribution by $q_{1-\alpha}$, gives us the test decision

$$\text{reject if } \ \widehat{S} > q_{1-\alpha}. \tag{4.19}$$

We make this result precise in the following proposition.

PROPOSITION 4.2 *Under the assumptions of Proposition 4.1,*

$$\widehat{S} \xrightarrow{d} \chi^2_4,$$

where χ^2_4 is a chi-squared distribution with four degrees of freedom. If H_0 is violated, $\widehat{S} \xrightarrow{P} \infty$.

It is a simple consequence of Theorems 4.2 and 4.5 that the test (4.19) is consistent against local alternatives of shape changes and changes in total volatility, with the rates discussed in those theorems.

REMARK 4.2 In view of Proposition 4.1 there are different ways of combining the test statistics $\widehat{S}^{(1)}, \widehat{S}^{(2)}$ for a joint test, apart from our choice of $\widehat{S}$. Such combinations correspond to different rejection regions in $\mathbb{R}^2_{\geq 0}$ for $(\widehat{S}^{(1)}, \widehat{S}^{(2)})$. Generically we can define for a function $f : \mathbb{R}^2_{\geq 0} \to \mathbb{R}_{\geq 0}$ the combined statistic $\widehat{S}_f = f(\widehat{S}^{(1)}, \widehat{S}^{(2)})$. A simple choice might be a sum $f_{\text{sum}}(x, y) = x + y$, which has linear, downward sloping contour lines and thus triangular rejection regions. Our choice

$$f_{\text{Fisher}}(x, y) = -2\Big\{ \log\left(1 - \Lambda^{(1)}(x)\right) + \log\left(1 - \Lambda^{(2)}(y)\right) \Big\}$$

has astroid shaped contour lines (like a p-norm with $0 < p < 1$). Accordingly rejection regions are shaped like ellipsoids. The precise shape of the contour lines depends on the asymptotic distributions $\Lambda^{(1)}, \Lambda^{(2)}$. The function f implies how evidence against the null hypothesis is interpreted in different scenarios. Roughly speaking, f_{sum} is indifferent between large x, large y or large $x + y$. This means that more evidence against the null might come just as well from one statistic, or the other, or their sum. In contrast f_{Fisher} is largest if both x and y are large, treating evidence against the null hypothesis as strongest, when it comes from both statistics and weaker if it only comes from one.

Finally, we discuss the problem of change point localization. For this purpose, we introduce the pooled change point estimator

$$\hat{\theta} := \frac{p^{(1)}}{p^{(1)} + p^{(2)}} \hat{\theta}^{(2)} + \frac{p^{(2)}}{p^{(1)} + p^{(2)}} \hat{\theta}^{(1)}. \tag{4.20}$$

Intuitively, $\hat{\theta}$ combines information from both estimators $\hat{\theta}^{(1)}, \hat{\theta}^{(2)}$, putting priority on the one where the change is more pronounced (smaller p-value). Our proof rests on a careful investigation of the tail behavior of the distributions $\Lambda^{(1)}, \Lambda^{(2)}$, see Theorem B.3 in the Appendix. The tail behavior of these distributions determines the relative size of the p-values $p^{(1)}, p^{(2)}$ in the above weights.

PROPOSITION 4.3 *Suppose Assumption 4.1 holds, $K \to \infty$, $K/N \to 0$, and the continuous function $\tilde{\sigma} : [0,1] \to (0,\infty)$ is such that $\tilde{\sigma}(\cdot)/\sigma_{(1)}(\cdot)$ is not constant.*
(i) If only $H_0^{(2)}$ is violated with $\sigma_{(2)} = (1 + 1/\sqrt{K})\sigma_{(1)}$, then

$$|\hat{\theta} - \theta| = \mathcal{O}_P\left(\frac{K}{N}\right).$$

(ii) If, in addition, $H^{(1)}$ is violated with $\sigma_{(2)} = (1 + 1/\sqrt{K})\sigma_{(1)} + \tilde{\sigma}/\sqrt{K}$, then

$$|\hat{\theta} - \theta| = \mathcal{O}_P\left(\frac{1}{N}\right).$$

5 Finite sample properties

5.1 Empirical size

We generate data under the null hypothesis according to the Functional Stochastic Volatility Model of [26]:

$$R_i(t) = \exp(g_i) \int_0^t \sigma(u) dW_i(u), \qquad t \in [0,1], \qquad i = 1, ..., N,$$
$$g_i = \varphi g_{i-1} + \varepsilon_i, \qquad \varepsilon_i \sim i.i.d.\ \mathcal{N}(0, \sigma_\varepsilon^2).$$

There are a number of settings to be carefully chosen:

- Following [26], we set $\varphi = 0.55$ and $\sigma_\varepsilon^2 = 0.25$ in order to reflect real-world data.
- We have four settings of $\sigma(\cdot)$
  - **Flat**: $\sigma(u) = 0.2$. This is a simple case that we have the same intraday volatility throughout the day.
  - **Slope**: $\sigma(u) = 0.1+0.2u$. The is a case that the intraday volatility is increasing in a linear manner.
  - **Sine**: $\sigma(u) = 0.1\sin(2\pi u) + 0.2$. This the case we have higher volatility in the morning, but lower volatility in the afternoon.
  - **U-shape**: $\sigma(u) = (u - 0.5)^2 + 0.1145299$. This choice is the most relevant one because it reflects the stylized fact that volatility is typically higher at the beginning and the end of a trading day.

  The coefficients in $\sigma(\cdot)$ are set to ensure that the above four $\sigma(\cdot)$ have a similar scale.
- The continuous time t in $[0, 1]$ is discretized as $[t_0, t_1, ..., t_K]$, where $t_k = k\Delta$ and $k = 1, ..., K$. This is the same for all random curves.
- The number of intraday observations is $K = 26, 39, 78$, which corresponds to 15-min, 10-min 5-min sampling intervals in our data analysis respectively. Their corresponding stepsizes are $\Delta = 1/26, 1/39, 1/78$.
- The sample size is $N = 100, 200, 500$.

Details on the computation of $\int_0^t \sigma(u)dW(u)$ and $\int_0^1 \mathbb{B}^2(u)du$, both use special approaches, are presented in Section D, which also contains step-by-step formulas for the computation of the three test statistics. The long-run variance of the $\log \widehat{Q}_i(1)$ was computed using the Bartlett kernel with bandwidth selected by the procedure of [33] with prewhitening.

Table 1 provides the empirical sizes of the three tests under four different shapes of $\sigma(\cdot)$. We see that the test performs very well, even for fairly small sample sizes N and low resolution K.

One advantage of using our tests is that it is robust against changes in g_i, which should not be mistaken as changes in the volatility function $\sigma_i(\cdot)$. To verify this property, we consider

$$g_i = \begin{cases} 0.45g_{i-1} + \varepsilon_i, & \epsilon_i \sim i.i.d.\ \mathcal{N}(0, \sigma_\varepsilon^2), \qquad i = 1, \ldots, \lfloor N/2\theta \rfloor, \\ 0.65g_{i-1} + \varepsilon_i, & \epsilon_i \sim i.i.d.\ \mathcal{N}(0, \sigma_\varepsilon^2), \qquad i = \lfloor N/2\theta \rfloor + 1, \ldots, N, \end{cases}$$

and all other settings are the same as before. Table 2 presents the empirical sizes of the three tests under the U-Shaped $\sigma_i(\cdot)$. The other three shapes yield similar results. As can be seen, the empirical sizes of our three tests are not affected by the change in g_i and match their theoretical levels reasonably well.

Table 1: Empirical size

		Shape of Volatility			Total Volatility			Global		
Flat		10%	5%	1%	10%	5%	1%	10%	5%	1%
$N = 100$	$K = 26$	11.4%	5.9%	1.4%	10.5%	5.1%	0.5%	11.4%	5.4%	1.2%
	$K = 39$	10.9%	5.8%	1.2%	10.5%	4.7%	0.4%	10.6%	5.0%	1.0%
	$K = 78$	11.9%	5.9%	0.9%	9.8%	4.3%	0.4%	11.0%	5.0%	0.9%
$N = 200$	$K = 26$	10.6%	5.3%	1.1%	10.5%	5.2%	0.9%	10.8%	5.6%	1.2%
	$K = 39$	10.8%	5.5%	1.3%	10.6%	5.1%	0.7%	11.3%	5.3%	1.0%
	$K = 78$	11.8%	5.4%	1.2%	10.1%	5.0%	0.9%	11.2%	5.4%	0.9%
$N = 500$	$K = 26$	11.0%	5.6%	1.0%	10.4%	5.6%	1.1%	11.0%	5.9%	1.1%
	$K = 39$	11.2%	5.5%	1.2%	11.0%	5.1%	0.8%	11.5%	5.6%	1.0%
	$K = 78$	11.1%	5.6%	1.3%	10.2%	4.7%	0.9%	10.7%	5.5%	1.2%
Slope										
$N = 100$	$K = 26$	11.3%	5.9%	1.4%	10.4%	4.6%	0.4%	11.1%	5.3%	0.9%
	$K = 39$	10.9%	5.3%	1.2%	10.0%	4.3%	0.4%	10.7%	5.2%	1.0%
	$K = 78$	10.5%	5.6%	1.4%	9.5%	4.0%	0.6%	10.5%	5.2%	0.8%
$N = 200$	$K = 26$	10.3%	5.4%	1.1%	10.9%	5.4%	1.1%	11.3%	5.9%	1.2%
	$K = 39$	11.7%	5.9%	1.3%	9.8%	4.8%	0.7%	11.1%	5.4%	1.0%
	$K = 78$	11.2%	6.1%	1.2%	10.7%	5.1%	0.5%	11.2%	5.8%	1.1%
$N = 500$	$K = 26$	10.7%	5.1%	0.9%	10.7%	5.5%	1.2%	11.1%	5.6%	0.8%
	$K = 39$	11.3%	5.6%	1.2%	10.3%	5.0%	1.0%	11.3%	5.6%	1.1%
	$K = 78$	11.2%	5.6%	1.1%	10.2%	5.1%	0.9%	10.9%	5.4%	1.1%
Sine										
$N = 100$	$K = 26$	11.3%	5.6%	1.1%	11.0%	5.5%	0.7%	11.6%	5.3%	0.8%
	$K = 39$	11.3%	5.9%	1.2%	10.8%	5.1%	0.7%	11.6%	5.7%	1.0%
	$K = 78$	11.8%	6.4%	1.5%	9.6%	4.6%	0.6%	11.5%	5.3%	0.9%
$N = 200$	$K = 26$	10.7%	5.1%	1.3%	11.6%	6.2%	1.2%	11.5%	6.2%	1.4%
	$K = 39$	11.1%	5.8%	1.3%	10.3%	4.8%	0.8%	11.5%	5.5%	1.1%
	$K = 78$	10.8%	5.5%	1.0%	10.3%	5.0%	0.8%	11.3%	5.4%	0.8%
$N = 500$	$K = 26$	10.6%	5.2%	1.0%	11.0%	5.5%	1.0%	11.2%	5.8%	0.9%
	$K = 39$	11.4%	5.3%	1.1%	9.8%	4.9%	0.8%	11.1%	5.7%	1.2%
	$K = 78$	11.1%	5.8%	1.0%	10.0%	5.0%	0.6%	11.1%	4.9%	0.8%
U-Shape										
$N = 100$	$K = 26$	11.0%	5.7%	1.2%	10.6%	4.7%	0.5%	11.0%	5.8%	1.0%
	$K = 39$	11.3%	6.1%	1.3%	11.0%	4.9%	0.4%	11.4%	5.2%	1.1%
	$K = 78$	10.9%	5.8%	1.4%	10.1%	4.3%	0.4%	10.7%	5.4%	1.1%
$N = 200$	$K = 26$	11.0%	5.9%	1.3%	11.0%	5.4%	0.9%	11.4%	6.1%	1.2%
	$K = 39$	11.2%	6.0%	1.1%	10.5%	5.4%	1.0%	11.2%	5.8%	1.1%
	$K = 78$	11.4%	6.0%	1.4%	10.4%	5.0%	0.8%	11.2%	6.3%	1.2%
$N = 500$	$K = 26$	11.0%	5.5%	1.1%	11.3%	5.2%	0.8%	11.0%	5.8%	1.0%
	$K = 39$	9.7%	4.9%	0.8%	10.4%	5.4%	1.0%	10.3%	4.9%	1.0%
	$K = 78$	10.6%	5.2%	1.1%	10.6%	5.2%	1.1%	11.0%	5.7%	1.0%

Table 2: Empirical size under a change in g_i

		Shape of Volatility			Total Volatility			Global		
		10%	5%	1%	10%	5%	1%	10%	5%	1%
$N = 100$	$K = 26$	12.0%	6.0%	1.4%	11.9%	5.8%	0.8%	12.2%	6.0%	1.3%
	$K = 39$	11.4%	6.0%	1.2%	10.9%	5.3%	0.5%	11.9%	5.4%	1.0%
	$K = 78$	10.7%	5.6%	1.1%	11.2%	4.8%	0.5%	11.0%	5.1%	0.8%
$N = 200$	$K = 26$	11.6%	6.3%	1.2%	12.8%	6.7%	1.2%	13.3%	6.7%	1.2%
	$K = 39$	10.4%	5.1%	1.0%	12.0%	6.0%	1.2%	11.6%	5.7%	1.3%
	$K = 78$	10.6%	5.0%	0.9%	11.2%	5.3%	0.9%	10.7%	5.5%	1.1%
$N = 500$	$K = 26$	10.8%	5.3%	1.0%	12.0%	6.2%	1.2%	11.5%	5.8%	1.3%
	$K = 39$	11.6%	5.9%	1.2%	11.6%	6.5%	1.4%	12.9%	7.2%	1.3%
	$K = 78$	11.3%	5.6%	1.3%	11.6%	6.3%	1.1%	12.3%	6.3%	1.3%

5.2 Empirical power

We set the time of the change at $\theta = 0.25, 0.5, 0.75$ and consider $N = 250$ and $N = 500$. All other settings are the same as under the null.

There are unlimited possibilities for a change in $\sigma_i(\cdot)$. To focus on the scenarios emphasized in this paper, we consider the following three alternative hypotheses:

1. $H_{A,1}$: a shape change in volatility, but no change in total volatility,
2. $H_{A,2}$: a change in total volatility, but no change in the shape of volatility,
3. $H_{A,3}$: a simultaneous change in the shape of volatility and total volatility.

Under $H_{A,1}$, we have a change in $\sigma_i(\cdot)$ from the flat shape to a sine shape. We have noticed that our test is very effective in detecting changes in the shape of volatility, and it can easily get empirical power of 100%. That is why we deliberately choose a very small change in the shape in order to show the convergence of the empirical power with respect to N and K. Specifically, we set

$$\sigma_i(u) = \begin{cases} 0.2, & \text{for } i = 1, ..., \lfloor N\theta \rfloor, \\ 0.02\sin(2\pi u) + \sqrt{199/5000}, & \text{for } i = \lfloor N\theta \rfloor + 1, ..., N. \end{cases}$$

The constant in the sine function is to ensure that the total volatility before and after the change is the same, i.e. $\int_0^1 0.2^2 du = \int_0^1 \left[0.02\sin(2\pi u) + \sqrt{199/5000}\right]^2 du = 0.04$. Thus, there is a change in the shape of volatility, but no change in the total volatility.

Under $H_{A,2}$, we introduce an upward parallel shift of the flat shape:

$$\sigma_i(u) = \begin{cases} 0.2, & \text{for } i = 1, ..., \lfloor N\theta \rfloor, \\ 0.4, & \text{for } i = \lfloor N\theta \rfloor + 1, ..., N. \end{cases}$$

Note that an upward parallel shift in the other three shapes (slope, sine, U-shape) will cause a change in total volatility as well as in the shape of volatility. This is because the other three shapes are actually "compressed" due to a higher total volatility.

Under $H_{A,3}$, we have a simultaneous change in shape and total volatility:

$$\sigma_i(u) = \begin{cases} 0.2, & \text{for } i = 1, ..., \lfloor N\theta \rfloor, \\ (u - 0.5)^2 + 0.4, & \text{for } i = \lfloor N\theta \rfloor + 1, ..., N. \end{cases}$$

The shape of $\sigma_i(\cdot)$ is changed from flat to U-shape, and total volatility is changed from $\int_0^1 0.2^2 du = 0.04$ to $\int_0^1 \left[(u - 0.5)^2 + 0.3\right]^2 du = 0.1525$.

Table 3 reports the empirical power. The conclusions can be summarized as follows.

1. Under $H_{A,1}$, the empirical power of the shape test and the global test increases with of N and K, in agreement with the $\sqrt{NK}$-consistency we established theoretically. The rejection rate of total volatility test is always around 5%, as expected since there is no change in total volatility in $H_{A,1}$.
2. Under $H_{A,2}$, the empirical power of the total volatility test increases with the growth of N, not with K. This is exactly what we expected because the test on total volatility is $\sqrt{N}$-consistent. Additionally, the rejection rate of testing the shape is typically around 5%, since there is no change in the shape in $H_{A,2}$.
3. Under $H_{A,3}$, the empirical power of the shape test and the global test increases with the growth of N and K, and empirical power of the volatility test increases with the growth of N, but not with K, again as predicted by our theory.

In Section D.2, we show that the change point estimators under the three alternatives inherit the properties of the corresponding tests: the performance of θ_1 and θ improves with increasing N and K, θ_2 improves with increasing N.

6 Application to US stocks

We begin with an individual stock as a prototype analysis to showcase our developed tests. Then, there are two ways to use the developed tests on a larger scale. First, since there could be multiple changes, we use the binary segmentation to explore all changes for one stock during a sample period. Second, we apply our test procedure to a large number of stocks and present the summary of first detected changes (without the binary segmentation).

For the purpose of demonstration, we focus on Tesla Inc. (Permno: 93436) for our prototype analysis. We consider 5-min intraday prices, the sample period is from Jun 29, 2010 (the IPO date) to Dec 31, 2021, corresponding to $N = 2891$ trading days. In each trading day i, we have the opening price $P_i(t_0)$ and the subsequent 78 5-min intraday prices $P_i(t_k)$, $k = 1, ..., 78$, with the last trading price in every 5-min time interval. Thus, the equidistant grid on the unit interval is $t_k = k\Delta$, $k = 0, 1, ..., K$, where $K = 78$ and the step size $\Delta = 1/78$.

Based on the intraday price data, we calculate the cumulative intraday return (CIDR) curves as

$$R_i(t_k) = \log(P_i(t_k)) - \log(P_i(t_0)), \qquad k = 1, ..., K, \ i = 1, ..., N.$$

Table 3: Empirical power

		$\theta = 0.25$			$\theta = 0.5$			$\theta = 0.75$		
$H_{A,1}$		Shape	Total	Global	Shape	Total	Global	Shape	Total	Global
$N = 250$	$K = 26$	62.6%	5.2%	51.1%	86.8%	5.3%	78.2%	62.9%	5.6%	51.3%
	$K = 39$	82.6%	5.0%	72.0%	96.9%	5.1%	93.3%	83.5%	5.0%	72.9%
	$K = 78$	98.7%	5.1%	96.8%	100.0%	4.7%	99.9%	99.2%	4.5%	97.1%
$N = 500$	$K = 26$	92.0%	5.5%	84.5%	99.2%	5.1%	97.6%	91.7%	5.5%	83.4%
	$K = 39$	99.1%	4.6%	97.2%	100.0%	5.1%	99.9%	98.8%	5.2%	96.2%
	$K = 78$	100.0%	5.5%	100.0%	100.0%	4.9%	100.0%	100.0%	5.1%	100.0%
$H_{A,2}$										
$N = 250$	$K = 26$	5.8%	86.8%	74.5%	5.5%	98.6%	95.2%	5.7%	86.8%	73.2%
	$K = 39$	5.9%	86.1%	72.7%	6.0%	98.5%	95.1%	6.0%	86.6%	73.0%
	$K = 78$	5.9%	86.0%	72.3%	5.3%	98.3%	94.9%	5.5%	85.7%	71.6%
$N = 500$	$K = 26$	5.6%	99.6%	97.9%	6.1%	100.0%	100.0%	5.2%	99.5%	98.1%
	$K = 39$	5.4%	99.7%	98.4%	5.9%	100.0%	100.0%	5.3%	99.7%	98.3%
	$K = 78$	5.7%	99.6%	98.3%	5.9%	100.0%	100.0%	5.3%	99.6%	98.0%
$H_{A,3}$										
$N = 250$	$K = 26$	89.2%	97.4%	99.9%	99.9%	100.0%	100.0%	95.2%	97.4%	99.9%
	$K = 39$	99.3%	97.3%	100.0%	100.0%	100.0%	100.0%	99.9%	97.7%	100.0%
	$K = 78$	100.0%	97.1%	100.0%	100.0%	99.9%	100.0%	100.0%	97.3%	100.0%
$N = 500$	$K = 26$	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
	$K = 39$	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
	$K = 78$	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

By definition, the CIDR curves always start from zero, i.e. $R_i(t_0) = 0$, and are scale invariant. We also compute the cumulative intraday realized volatility (CIDRV) curves as

$$RV_i(t_k) = \sum_{k=1}^{K} |R_i(t_k) - R_i(t_{k-1})|^2 \mathbb{I}\{t_k < t\}, \qquad k = 1, ..., K, \ i = 1, ..., N.$$

In order to visualize the important functional objects, Figure 2 plots the intraday Price $P_i(t_k)$, CIDRs $R_i(t_k)$, and CIDRVs $RV_i(t_k)$ in the upper, middle, and lower panels, respectively.

We apply the tests for the whole sample period, in order to detect 1) a shape change, 2) a change in total volatility, and 3) an arbitrary change. Table 4 presents the test results. The p-value of testing H_0^1 is 0.02%, providing strong evidence of a shape change. The change point estimator $\hat{\theta}_1$ is 0.26, indicating the shape change occurred on Jul 1, 2013. As for testing H_0^2, we find strong evidence of a change in total volatility with p-value of 0.96%. The date of change in total volatility is May 12, 2014, as suggested by the $\hat{\theta}_2 = 0.34$. Combing the two tests, we have the p-value of 0.00% for the global null hypothesis H_0, with the pooled change point estimator $\hat{\theta} = 0.26$, implying that the date of arbitrary change in intraday volatility pattern is July 8, 2013.

Table 4: Test results of Tesla (note that the p-values are in percent).

	p-value	Change Point Estimator	Date of Change
Shape of Volatility (H_0^1)	0.02%	0.26	Jul 1, 2013
Total Volatility (H_0^2)	0.96%	0.34	May 12, 2014
Global (H_0)	0.00%	0.26	Jul 8, 2013

As the sample period of the Tesla analysis covers more than a decade, there could be multiple changes in the intraday volatility pattern. Thus, we use the standard binary segmentation based on the global test at the 5% significance level and the pooled change point estimator ($\hat{\theta}$). Table 5 presents the result with some associated events that could be used to validate the identified change points.

Table 5: Result of binary segmentation to test multiple changes for Tesla

p-value	Date of Change	Related News
0.00%	Jul 8, 2013	Tesla joined the Nasdaq 100 index on Jul 15, 2013
0.00%	Jul 16, 2014	Tesla announced new smaller electric vehicle named Model 3 on Jul 16, 2014
0.08%	Feb 6, 2018	Elon Musk made history launching a car into space on Feb 6, 2018
0.09%	Jan 23, 2019	Tesla posted back-to-back profits for the first time
4.79%	Dec 20, 2019	Tesla's Chinese factory delivered its first cars
0.48%	Mar 31, 2021	NHTSA confirmed no violation of Tesla's touchscreen drive selector

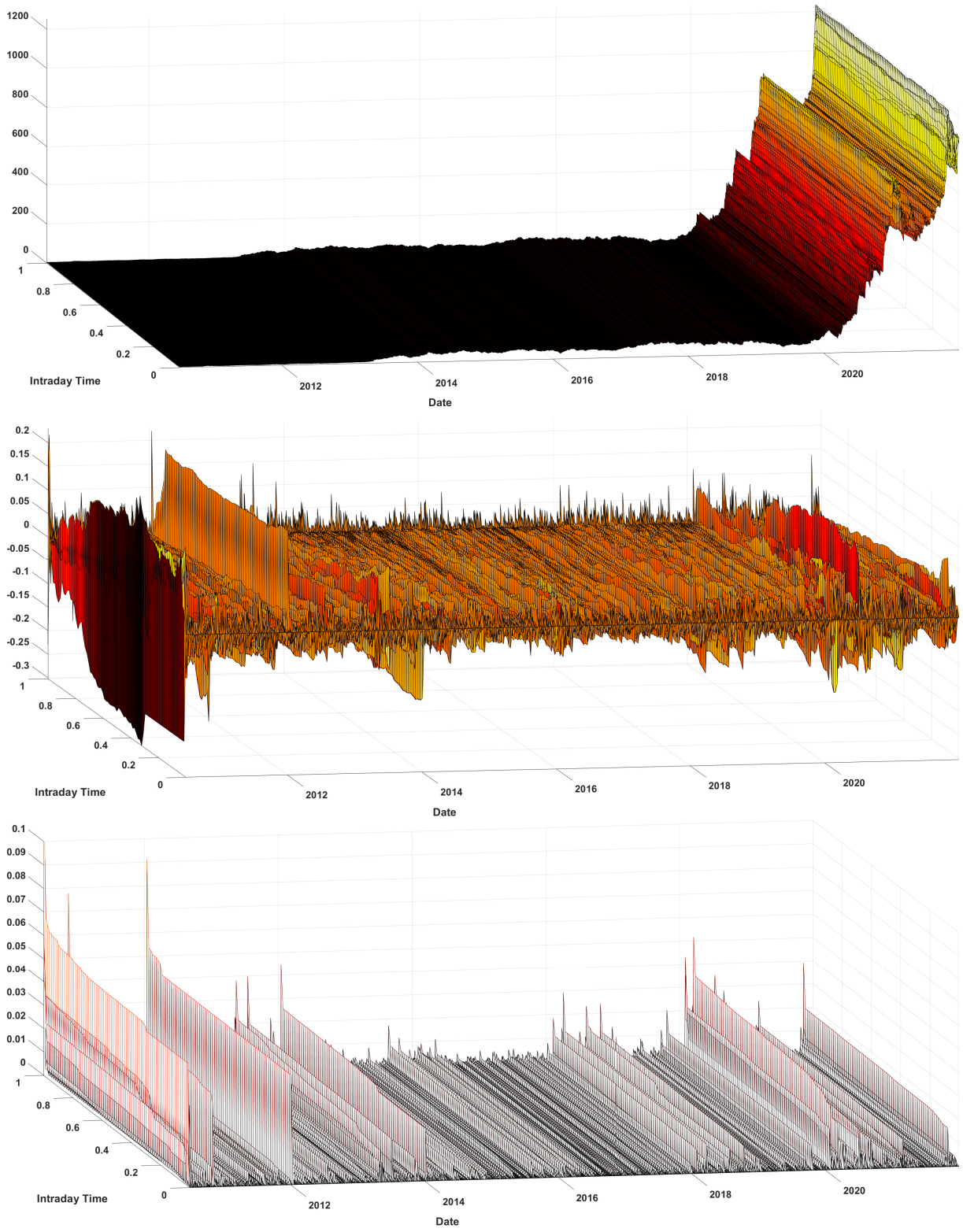

Figure 2: Time series of functional objects derived from intraday Tesla prices. Upper Panel: Intraday Price $P_i(t_k)$; Middle Panel: CIDRs $R_i(t_k)$; Lower Panel: CIDRVs $RV_i(t_k)$.

It is also interesting to examine change in the intraday volatility pattern for other stocks. Thus, we apply our test procedure to 7293 stocks in the US stock markets. To preserve space, we focus on the first change detected by the global test at 5% significance level, without using binary segmentation to find additional changes. The stocks used and the data cleaning procedure are the same as in [26]. Their sample period varies in length from 2 to 25 years. Shorter sample periods could be due to IPO dates later than Jan 3, 2006 or stocks delisted before Dec 31, 2021.

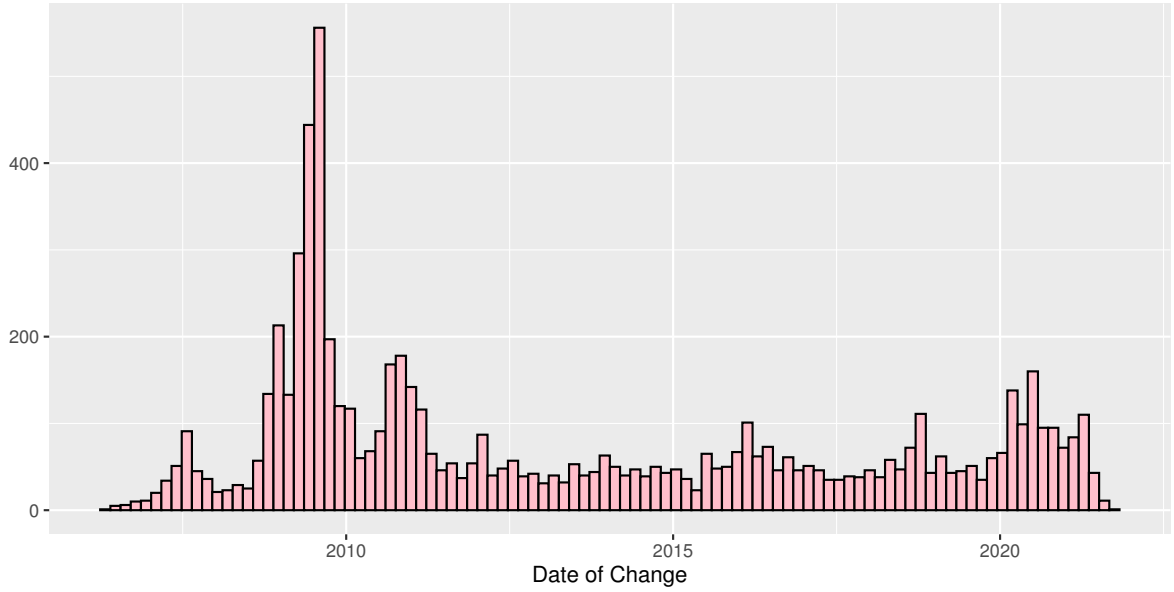

Figure 3: Dates of first change in the intraday volatility pattern for 7168 US stocks.

Our test indicates that 7168 out of 7293 stocks (98.3%) underwent at least one change in the intraday volatility pattern. This provides the evidence that change in the intraday volatility pattern is a common issue in the US stocks. To provide further insights, we plots the histogram of the first detected changes in Figure 3. We can clearly see that 1) the highest frequent changes happen during the subprime mortgage crisis in 2008, 2) the second highest frequent changes occur around the European debt crisis in the 2010s, 3) the third highest frequent changes appear after COVID in 2020. These results show that our test is able to detect change points that are consistent with well-known market events, providing additional validation on a very large data set.

Acknowledgements We thank two referees for insightful comments that helped us improve the paper. Piotr Kokoszka and Haonan Wang were partially supported by NSF grant DMS–2123761.

References

[1] Y. Aït-Sahalia and J. Jacod. Testing for jumps in a discretely observed process. *The Annals of Statistics*, 37(1):184 – 222, 2009.

[2] Yacine Aït-Sahalia and Jean Jacod. *High-Frequency Financial Econometrics*. Princeton University Press, Princeton, illustrated edition edition, 2014.

[3] M. Aston and C. Kirch. Detecting and estimating changes in dependent functional data. *Journal of Multivariate Analysis*, 109:204—-220, 2012.

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

[5] A. Aue, G. Rice, and G. Sonmez. Structural break analysis for spectrum and trace of covariance operators. *Environmetrics*, 31:e2617, 2020.

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

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

[8] M. Bibinger and M. Madensoy. Change-point inference on volatility in noisy Itó semimartingales. *Stochastic Processes and their Applications*, 129:4878–4925, 2019.

[9] M. Bibinger, J. Moritz, and M. Vetter. Nonparametric change-point analysis of volatility. *The Annals of Statistics*, 45:1542 – 1578, 2017.

[10] K. Christensen, U. Hounyo, and M. Podolskij. Is the diurnal pattern sufficient to explain intraday variation in volatility? a nonparametric assessment. *Journal of Econometric*, 205:336–362, 2018.

[11] A. DasGupta. *Probability for Statistics and Machine Learning: Fundamentals and Advanced Topics*. Springer Texts in Statistics. Springer, 2011.

[12] H. Dette and K. Kokot. Detecting relevant differences in the covariance operators of functional time series: a sup-norm approach. *Annals of the Institute of Statistical Mathematics*, 74(2):195–231, 2022.

[13] H. Dette, K. Kokot, and A. Aue. Functional data analysis in the Banach space of continuous functions. *The Annals of Statistics*, 48:1168 – 1192, 2020.

[14] H. Dette, K. Kokot, and S. Volgushev. Testing relevant hypotheses in functional time series via self-normalization. *Journal of the Royal Statistical Society Series (B)*, 82:629–660, 2020.

[15] H. Dette and T. Kutta. Detecting structural breaks in eigensystems of functional time series. *Electronic Journal of Statistics*, 15:944 – 983, 2021.

[16] M. Griebel and G. Li. On the decay rate of the singular values of bivariate functions. *SIAM Journal on Numerical Analysis*, 56:974–993, 2018.

[17] O. Gromenko, P. Kokoszka, and M. Reimherr. Detection of change in the spatiotemporal mean function. *Journal of the Royal Statistical Society (B)*, 79:29–50, 2017.

[18] M. Hoffmann and H. Dette. On detecting changes in the jumps of arbitrary size of a time-continuous stochastic process. *Electronic Journal of Statistics*, 13(2):3654 – 3709, 2019.

[19] M. Hoffmann, M. Vetter, and H. Dette. Nonparametric inference of gradual changes in the jump behaviour of time-continuous processes. *Stochastic Processes and their Applications*, 128(11):3679–3723, 2018.

[20] L. Horváth, P. Kokoszka, and G. Rice. Testing stationarity of functional time series. *Journal of Econometrics*, 179:66–82, 2014.

[21] L. Horváth and G. Rice. Extensions of some classical methods in change point analysis. *Test*, 23:219–255, 2014.

[22] L. Horváth, G. Rice, and L. Zhao. Change point analysis of covariance functions: a weighted cumulative sum approach. *Journal of Multivariate Analysis*, 189:104877, 2022.

[23] T. Hsing and R. Eubank. *Theoretical Foundations of Functional Data Analysis, with an Introduction to Linear Operators*. Wiley, 2015.

[24] J. Jacod and P. Protter. *Discretization of Processes*. Stochastic Modelling and Applied Probability. Springer Berlin Heidelberg, 2011.

[25] Ioannis Karatzas and Steven Shreve. *Brownian Motion and Stochastic Calculus*. Springer, New York, 2nd edition edition, 1991.

[26] P. Kokoszka, N. Mohammadi, H. Wang, and S. Wang. Functional diffusion driven stochastic volatility model. 2023. arXiv:2305.04112.

[27] P. Kokoszka and M. Reimherr. *Introduction to Functional Data Analysis*. Texts in Statistical Science. CRC Press, 2017.

[28] J. Kuelbs. The invariance principle for Banach space valued random variables. *Journal of Multivariate Analysis*, 3:161–172, 1973.

[29] S. Kundu, S. Majumdar, and K. Mukherjee. Central limit theorems revisited. *Statistics & Probability Letters*, 47(3):265–275, 2000.

[30] J-F. Le Gall. *Brownian Motion, Martingales, and Stochastic Calculus*. Springer, 2016.

[31] F. Merlevède, M. Peligrad, and S. Utev. Recent advances in invariance principles for stationary sequences. *Probability Surveys*, 3:1 – 36, 2006.

[32] F. A. Moricz, R. J. Serfling, and W. F. Stout. Moment and probability bounds with quasi-superadditive structure for the maximum partial sum. *The Annals of Probability*, 4:1032 – 1040, 1982.

[33] Whitney K Newey and Kenneth D West. Automatic lag selection in covariance matrix estimation. *The Review of Economic Studies*, 61(4):631–653, 1994.

[34] B. Øksendal. *Stochastic Differential Equations*. Universitext. Springer Berlin Heidelberg, 2003.

[35] G. P. Quinn and M. J. Keough. *Experimental Design and Data Analysis for Biologists*. Cambridge University Press, 2002.

[36] A. Račkauskas and C. Suquet. Testing epidemic changes of infinite dimensional parameters. *Statistical Inference for Stochastic Processes*, 9:111–134, 2006.

[37] G. Rice and M. Shum. Inference for the lagged cross-covariance operator between functional time series. *Journal of Time Series Analysis*, 40:665–692, 2019.

[38] X. Shao. Self-normalization for time series: A review of recent developments. *Journal of the American Statistical Association*, 110:1797–1817, 2015.

[39] X. Shao and X. Zhang. Testing for change points in time series. *Journal of the American Statistical Association*, 105:1228–1240, 2010.

[40] G. R. Shorack and J. A. Wellner. *Empirical Processes with Applications to Statistics*. Wiley, 1986.

[41] C. Stöhr, J. Aston, and C. Kirch. Detecting changes in the covariance structure of functional time series with application to fMRI data. *Econometrics and Statistics*, 18:44–62, 2021.

[42] L. Tolmatz. On the Distribution of the Square Integral of the Brownian Bridge. *The Annals of Probability*, 30:253–269, 2002.

[43] R. Vershynin. *High-Dimensional Probability: An Introduction with Applications in Data Science*. Cambridge University Press, Cambridge, 2018.

[44] K. Yoshihara. Moment inequalities for mixing sequences. *Kodai Math. J.*, 1:316–328, 1978.

Proofs and additional information

This supplementary material contains the proofs of our theoretical results and information about numerical implementation. In the following derivations, c denotes a generic, positive constant that may change from one line to another and is always independent of N, K and i. When needed, we use a sequence of positive constants, which we denote by $c_1, c_2, \ldots$. The supplementary material consists of three parts: In Appendix A, we gather concentration results for the realized quadratic variation process (denoted by $\widehat{V}(\cdot)$) in the absence of the random factor $\exp(g_i)$. These results are of independent interest and formulated for $K \to \infty$, since they do not require functional replications. In Appendix B, we prove our main results on the convergence of the test statistics $\widehat{S}^{(i)}$, for $i = 1, 2$, under the corresponding null hypotheses $H_0^{(i)}$ and local alternatives. Finally, Appendices C and D provide additional details, as well as pseudocode for all of our procedures.

A Uniform limits for the realized quadratic variation

In this section, we present some fundamental results concerning the limiting behavior of the realized quadratic variation process of the model (3.1) in absence of the stochastic coefficient $\exp(g_i)$. For the sake of simplicity we drop index i, and define

$$V(t) := \int_0^t \sigma^2(u)du, \quad t \in [0,1]. \tag{A.1}$$

The empirical counterpart of $V(\cdot)$ is

$$\widehat{V}(t) := \sum_{k=1}^{K} \left| \int_{(k-1)/K}^{k/K} \sigma(u)dW(u) \right|^2 \mathbb{I}\{k/K \leq t\}. \tag{A.2}$$

Most of the results stated in this section are more general than necessary for our later investigation. Yet, we consider them to be of independent interest. The bounds are formulated with constants c that are independent on σ. We make the dependence on σ explicit, as it is needed for our subsequent theory. It is convenient to state the following assumption.

Assumption A.1 For the Itô process $\int_0^t \sigma(u)dW(u)$, Condition 1 of Assumption 4.1 holds, i.e. the volatility function $\sigma : [0,1] \to (0,\infty)$ is continuous

Lemma A.1 *Under Assumption A.1, for any $M \in \mathbb{N}$,*

$$\mathbb{E}\left[\widehat{V}(1) - V(1)\right]^{2M} \leq \frac{c\|\sigma\|_\infty^{4M}}{K^M}. \tag{A.3}$$

PROOF. Recall (A.1) and (A.2). Then we have

$$\mathbb{E}\left[\widehat{V}(1)-V(1)\right]^{2M}=\mathbb{E}\left[\sum_{k=1}^{K}\left|\int_{(k-1)/K}^{k/K}\sigma(u)dW(u)\right|^{2}-V(1)\right]^{2M} \tag{A.4}$$

$$=\mathbb{E}\left[\sum_{k=1}^{K}\left(\left|\int_{(k-1)/K}^{k/K}\sigma(u)dW(u)\right|^{2}-[V(k/K)-V((k-1)/K)]\right)\right]^{2M}$$

$$=:\mathbb{E}\left[\sum_{k=1}^{K}Y_k[V(k/K)-V((k-1)/K)]\right]^{2M},$$

where

$$Y_k=\frac{\left|\int_{(k-1)/K}^{k/K}\sigma(u)dW(u)\right|^{2}}{V(k/K)-V((k-1)/K)}-1,\quad k=1,2,\ldots,K, \tag{A.5}$$

are independent centered random variables with law

$$Y_k+1\sim\chi_1^2. \tag{A.6}$$

We can now apply Theorem 1 in [44] to the right-hand side of (A.4), which yields

$$\begin{aligned}\mathbb{E}\left[\sum_{k=1}^{K}Y_k[V(k/K)-V((k-1)/K)]\right]^{2M}&\le c\Big(\sum_{k=1}^{K}[V(k/K)-V((k-1)/K)]^2\Big)^{M}\\&\le c\big(\|\sigma\|_\infty^4/K\big)^{M}.\end{aligned}$$

This completes the proof of (A.3). ∎

Proposition A.1 below extends the result of Proposition A.1 in [26] to all even moments $2M$.

PROPOSITION A.1 *Under Assumption A.1,*

$$\mathbb{E}\left[\sup_{t\in[0,1]}\left|\widehat{V}(t)-V(t)\right|\right]^{2M}\le\frac{c\|\sigma\|_\infty^{4M}}{K^M}.$$

(The constant c depends on M.)

PROOF. For $t\in[0,1]$, we have

$$\begin{aligned}\left|\widehat{V}(t)-V(t)\right|^{2M}\le& c\left|\widehat{V}(t)-V(t)-\widehat{V}(t^*(t))+V(t^*(t))\right|^{2M}\\&+c\left|\widehat{V}(t^*(t))-V(t^*(t))\right|^{2M},\end{aligned} \tag{A.7}$$

where $t^*(t) = \frac{1}{K}\lfloor tK \rfloor$, and c depends on the integer M only. Definition (A.2) implies $\widehat{V}(t^*(t)) = \widehat{V}(t)$. This leads to

$$\left|\widehat{V}(t) - V(t)\right|^{2M} \leq c\left|V(t) - V(t^*(t))\right|^{2M} + c\left|\widehat{V}(t^*(t)) - V(t^*(t))\right|^{2M}.$$

Continuity of the volatility function $\sigma(\cdot)$ implies

$$\sup_{t\in[0,1]}\left|\widehat{V}(t) - V(t)\right|^{2M} \leq cK^{-2M}\|\sigma\|_\infty^{4M} + c\max_{j\in\{1,2,\ldots,K\}}\left|\widehat{V}(j/K) - V(j/K)\right|^{2M}. \tag{A.8}$$

So, it is enough to focus on the discrete index set $\{j/K\}_{j=0}^K$. Observe that the sequence

$$\begin{aligned}\widehat{V}(j/K) - V(j/K) &= \sum_{k=1}^{j}\left|\int_{(k-1)/K}^{k/K}\sigma(u)dW(u)\right|^2 - \int_0^{j/K}\sigma^2(u)du \\ &= \sum_{k=1}^{j}\left(\left|\int_{(k-1)/K}^{k/K}\sigma(u)dW(u)\right|^2 - \int_{(k-1)/K}^{k/K}\sigma^2(u)du\right), \quad j = 1, 2, \ldots K,\end{aligned}$$

forms a martingale, see the proof of Proposition A.1 in [26] for details. We now apply Doob's maximal inequality, see [11] Theorem 14.7, to obtain

$$\mathbb{E}\left[\sup_{j\in\{1,2,\ldots,K\}}\left|\widehat{V}(j/K) - V(j/K)\right|\right]^{2M} \leq \left(\frac{2M}{2M-1}\right)^{2M}\mathbb{E}\left|\widehat{V}(1) - V(1)\right|^{2M}.$$

This, together with inequality (A.8), implies

$$\mathbb{E}\left[\sup_{t\in[0,1]}\left|\widehat{V}(t) - V(t)\right|\right]^{2M} \leq cK^{-2M}\|\sigma\|_\infty^{4M} + c\mathbb{E}\left|\widehat{V}(1) - V(1)\right|^{2M},$$

where, again, the constant c depends only on M. Applying Lemma A.1 gives

$$\mathbb{E}\left[\sup_{t\in[0,1]}\left|\widehat{V}(t) - V(t)\right|\right]^{2M} \leq cK^{-2M}\|\sigma\|_\infty^{4M} + cK^{-M}\|\sigma\|_\infty^{4M}.$$

This completes the proof. ∎

LEMMA A.2 *Under Assumption A.1, there exists an absolute constant $c_1 > 0$ such that for any $\varepsilon \in (0,1)$,*

$$\mathbb{P}(|\widehat{V}(1) - V(1)| > \varepsilon) \leq 2\exp\Big(-\frac{c_1\varepsilon^2 K}{\|\sigma\|_\infty^4}\Big).$$

PROOF. We begin by rewriting the difference

$$\begin{aligned}\widehat{V}(1) - V(1) &= \sum_{k=1}^{K} \Big\{ \Big| \int_{(k-1)/K}^{k/K} \sigma(u) dW(u) \Big|^2 - V(1) \Big\} \\ &= \sum_{k=1}^{K} \left(\left| \int_{(k-1)/K}^{k/K} \sigma(u) dW(u) \right|^2 - [V(k/K) - V((k-1)/K)] \right) \\ &=: \sum_{k=1}^{K} Y_k [V(k/K) - V((k-1)/K)],\end{aligned}$$

where $Y_1, \ldots, Y_K$ are independent centered random variables defined in (A.5) satisfying (A.6). We now define the ψ_1-norm for a subexponential random variable Y as

$$\|Y\|_{\psi_1} := \inf\{t > 0 : \mathbb{E}[\exp(|Y|/t)] \leq 2\}.$$

For details on $\|\cdot\|_{\psi_1}$ as well as the fact that it indeed constitutes a norm see, for example, [43]. Due to homogeneity of the norm, we can write

$$\Big\| Y_k [V(k/K) - V((k-1)/K)] \Big\|_{\psi_1} = \|Y_k\|_{\psi_1} [V(k/K) - V((k-1)/K)] \leq \frac{\|Y_k\|_{\psi_1} \|\sigma\|_\infty^2}{K},$$

for any K. Moreover, $\|Y_k\|_{\psi_1} = c_2 < \infty$, which follows by Lemma 2.7.6 and Exercise 2.7.10 in [43]. Now, applying Bernstein's inequality for subexponential random variables (Theorem 2.8.1 in [43]) proves the Lemma. ∎

LEMMA A.3 *Define the kernel function*

$$c_V(s,t) = \int_0^{\min(s,t)} \sigma^4(x) dx, \quad s,t \in [0,1], \tag{A.9}$$

and the sequence (indexed by the grid size K) of kernels

$$\mathbb{E}\left(\Big\{ \widehat{V}(s) - V(s) \Big\} \Big\{ \widehat{V}(t) - V(t) \Big\} \right), \quad s,t \in [0,1].$$

Then, under Assumption A.1,

$$\sup_{0 \leq s,t \leq 1} \left| K \mathbb{E}\left(\Big\{ \widehat{V}(s) - V(s) \Big\} \Big\{ \widehat{V}(t) - V(t) \Big\} \right) - c_V(s,t) \right| \to 0, \quad \text{as } K \to \infty. \tag{A.10}$$

PROOF. We start with

$$\begin{aligned}&\mathbb{E}\left[\Big\{ \widehat{V}(s) - V(s) \Big\} \Big\{ \widehat{V}(t) - V(t) \Big\} \right] \\ =&\mathbb{E}\left[\Big\{ \widehat{V}(s) - V(t^*(s)) + V(t^*(s)) - V(s) \Big\} \Big\{ \widehat{V}(t) - V(t^*(t)) + V(t^*(t)) - V(t) \Big\} \right]\end{aligned}$$

where $t^*(t) = \frac{1}{K}\lfloor tK \rfloor$, for $t \in [0,1]$. This entails

$$\begin{aligned} K\mathbb{E}\left[\left\{\widehat{V}(s) - V(s)\right\}\left\{\widehat{V}(t) - V(t)\right\}\right] =& K\mathbb{E}\left[\left\{\widehat{V}(s) - V(t^*(s))\right\}\left\{\widehat{V}(t) - V(t^*(t)\right\}\right] \\ &+ K\mathbb{E}\left\{\widehat{V}(s) - V(t^*(s))\right\}\left\{V(t^*(t)) - V(t)\right\} \\ &+ K\left\{V(t^*(s)) - V(s)\right\}\mathbb{E}\left\{\widehat{V}(t) - V(t^*(t))\right\} \\ &+ K\left\{V(t^*(s)) - V(s)\right\}\left\{V(t^*(t)) - V(t)\right\}, \end{aligned}$$

According to Proposition A.1 in [26], we have

$$\mathbb{E}\left\{\sup_{s\in[0,1]}|\widehat{V}(s) - V(t^*(s))|\right\} = O\left(K^{-\frac{1}{2}}\right).$$

Moreover, continuity of the volatility function $\sigma(\cdot)$ implies

$$\sup_{s\in[0,1]}|V(t^*(s)) - V(s)| = O\left(K^{-1}\right).$$

This gives

$$\begin{aligned} & K\mathbb{E}\left[\left\{\widehat{V}(s) - V(s)\right\}\left\{\widehat{V}(t) - V(t)\right\}\right] \\ =& K\mathbb{E}\left[\left\{\widehat{V}(s) - V(t^*(s))\right\}\left\{\widehat{V}(t) - V(t^*(t))\right\}\right] + KO\left(K^{-\frac{3}{2}}\right) \\ =& K\mathbb{E}\left[\left\{\widehat{V}(s) - V(t^*(s))\right\}\left\{\widehat{V}_i(t) - V(t^*(t))\right\}\right] + O\left(K^{-\frac{1}{2}}\right). \end{aligned} \tag{A.11}$$

We now investigate the first term in (A.11). Observe that

$$\widehat{V}(t) - V(t^*(t)) = \sum_{k=1}^{K}\left\{\left|\int_{(k-1)/K}^{k/K}\sigma(x)dW(x)\right|^2 - \int_{(k-1)/K}^{k/K}\sigma^2(x)dx\right\}\mathbb{I}\{k/K \le t\}.$$

Therefore,

$$\begin{aligned} & K\mathbb{E}\left[\left\{\widehat{V}(s) - V(t^*(s))\right\}\left\{\widehat{V}(t) - V(t^*(t))\right\}\right] \\ =& K\mathbb{E}\Bigg[\left(\sum_{k=1}^{K}\left\{\left|\int_{(k-1)/K}^{k/K}\sigma(x)dW(x)\right|^2 - [V(k/K) - V((k-1)/K)]\right\}\right)^2 \\ & \qquad \times \mathbb{I}\{k/K \le s\}\mathbb{I}\{k/K \le t\}\Bigg] \\ =:& K\mathbb{E}\left[\left(\sum_{k=1}^{K}Y_k[V(k/K) - V((k-1)/K)]\right)^2\mathbb{I}\{k/K \le \min(s,t)\}\right], \end{aligned}$$

where the Y_k are independent, centered random variables defined in (A.5) with the law specified in (A.6). Consequently,

$$
\begin{aligned}
& K\mathbb{E}\left[\left\{\widehat{V}(s)-V(t^*(s))\right\}\left\{\widehat{V}(t)-V(t^*(t))\right\}\right] \\
=& K\mathbb{E}\left[\sum_{k=1}^{K} Y_k^2[V(k/K)-V((k-1)/K)]^2\mathbb{I}\{k/K\le\min(s,t)\}\right] \\
=& K\sum_{k=1}^{K}\left(\int_{(k-1)/K}^{k/K}\sigma^2(x)dx\right)^2\mathbb{I}\{k/K\le\min(s,t)\} \\
=& K\sum_{k=1}^{K}\left(K^{-1}\sigma^2(\tilde{t}_k)\right)^2\mathbb{I}\{k/K\le\min(s,t)\} \qquad \text{(A.12)}
\end{aligned}
$$

for some $\tilde{t}_k \in [(k-1)/K, k/K]$, $k = 1, 2, \ldots, K$. So, it is enough to prove that the Riemann–type sum (A.12) converges to the kernel (A.9) uniformly. To do so, observe that

$$
\begin{aligned}
& \sup_{0\le s,t\le 1}\left|K^{-1}\sum_{k=1}^{K}\sigma^4(\tilde{t}_k)\mathbb{I}\{k/K\le\min(s,t)\}-\int_0^{\min(s,t)}\sigma^4(x)dx\right| \\
=& \sup_{0\le s,t\le 1}\left|\sum_{k=1}^{\lfloor K\cdot\min(s,t)\rfloor}\int_{\frac{k-1}{K}}^{\frac{k}{K}}\left(\sigma^4(\tilde{t}_k)-\sigma^4(x)\right)dx\right|+O(K^{-1}) \\
\le& \sup_{0\le s,t\le 1}\lfloor K\cdot\min(s,t)\rfloor K^{-1}\sup_{|x-y|\le K^{-1}}\left|\sigma^4(x)-\sigma^4(y)\right|+O(K^{-1}) \\
=& o(1), \qquad \text{(A.13)}
\end{aligned}
$$

where (A.13) is a consequence of the uniform continuity of the function $\sigma(\cdot)$. Combining (A.11) and (A.13) gives the desired convergence result (A.10). ∎

B Proofs of the results of Section 4

B.1 Proofs of Theorems 4.1, 4.4 and Propositions 4.1, 4.2 (behavior under the null hypotheses)

This section is dedicated to the analysis of the test statistics $\widehat{S}^{(i)}$, for $i = 1, 2$. We focus first on $\widehat{S}^{(1)}$ and subsequently turn to the test statistic $\widehat{S}^{(2)}$, concluding with their joint behavior.

To show weak convergence of $\widehat{S}^{(1)}$, we show a weak invariance principle for the functional partial sum process

$$
P_N^{(1)}(x,t) := \frac{1}{\sqrt{N}}\sum_{i=1}^{\lfloor xN\rfloor}\sqrt{K}\left\{\widehat{F}_i(t)-\mathbb{E}[\widehat{F}_i(t)]\right\}, \quad t\in[0,1]. \qquad \text{(B.1)}
$$

Notice that thus defined, $P_N^{(1)}(x) = P_N^{(1)}(x,\cdot)$ is for every $x \in [0,1]$ a random function.

Our first theorem, in conjunction with Lemma B.1 following it, establishes an explicit formula for the kernel $c_F(\cdot,\cdot)$ in Theorem 4.1 as well as the uniform convergence needed in subsequent proofs.

THEOREM B.1 *Under Assumption A.1,*

$$\sup_{0\le s,t\le 1}\left|K\mathbb{E}\left[\left(\widehat{F}(s)-F(s)\right)\left(\widehat{F}(t)-F(t)\right)\right]-c_F(s,t)\right|\to 0,\quad \text{as } K\to 0,$$

where

$$c_F(s,t)=\frac{1}{V^2(1)}c_V(s,t)-\frac{V(t)}{V^3(1)}c_V(s,1)-\frac{V(s)}{V^3(1)}c_V(1,t)+\frac{V(s)V(t)}{V^4(1)}c_V(1,1),\ s,t\in[0,1],$$

and where $c_V(\cdot,\cdot)$ is defined in (A.9) *and $V(\cdot)$ in* (A.1).

PROOF. First define the event $A_{K,\varepsilon}=\left\{\left|\widehat{V}(1)-V(1)\right|<\varepsilon\right\}$, for sufficiently small positive $\varepsilon\in(0,V(1))$. Then, we have

$$\begin{aligned}
&K\mathbb{E}\left[\left(\widehat{F}(s)-F(s)\right)\left(\widehat{F}(t)-F(t)\right)\right]\\
=&K\mathbb{E}\left[\left(\widehat{F}(s)-F(s)\right)\left(\widehat{F}(t)-F(t)\right)\mathbb{I}\left\{A_{K,\varepsilon}\right\}\right]\\
&+K\mathbb{E}\left[\left(\widehat{F}(s)-F(s)\right)\left(\widehat{F}(t)-F(t)\right)\mathbb{I}\left\{A_{K,\varepsilon}^c\right\}\right]\\
=&B_1(s,t)+B_2(s,t).
\end{aligned}$$

Since F, $\widehat{F}$ are cdfs, their difference is absolutely bounded by 1. Boundedness of the difference $\left(\widehat{F}(\cdot)-F(\cdot)\right)$ and Lemma A.1 entail

$$\sup_{0\le s,t\le 1}|B_2(s,t)|\le cK\mathbb{P}\left(A_{K,\varepsilon}^c\right)\le cK\frac{\mathbb{E}\left|\widehat{V}(1)-V(1)\right|^4}{\varepsilon^4}=O\left(K^{-1}\right). \tag{B.2}$$

We now investigate $B_1(\cdot,\cdot)$. Define

$$U(s):=\frac{\left[\widehat{V}(s)-V(s)\right]V(1)-V(s)\left[\widehat{V}(1)-V(1)\right]}{\widehat{V}(1)V(1)}.$$

A simple calculation entails

$$\begin{aligned}
B_1(s,t) =& K\mathbb{E}\left[U(s)U(t)\mathbb{I}\left\{A_{K,\varepsilon}\right\}\right] \\
=& K\mathbb{E}\left[\frac{\left[\widehat{V}(s)-V(s)\right]\left[\widehat{V}(t)-V(t)\right]}{\widehat{V}^2(1)}\mathbb{I}\left\{A_{K,\varepsilon}\right\}\right] \\
&- K\mathbb{E}\left[\frac{\left[\widehat{V}(s)-V(s)\right]V(t)\left[\widehat{V}(1)-V(1)\right]}{\widehat{V}^2(1)V(1)}\mathbb{I}\left\{A_{K,\varepsilon}\right\}\right] \\
&- K\mathbb{E}\left[\frac{V(s)\left[\widehat{V}(1)-V(1)\right]\left[\widehat{V}(t)-V(t)\right]}{\widehat{V}^2(1)V(1)}\mathbb{I}\left\{A_{K,\varepsilon}\right\}\right] \\
&+ K\mathbb{E}\left[\frac{V(s)V(t)\left[\widehat{V}(1)-V(1)\right]^2}{\widehat{V}^2(1)V^2(1)}\mathbb{I}\left\{A_{K,\varepsilon}\right\}\right] \\
=:& B_{11}(s,t)-B_{12}(s,t)-B_{13}(s,t)+B_{14}(s,t).
\end{aligned}$$

The term B_{11} satisfies

$$\begin{aligned}
B_{11}(s,t) =& K\mathbb{E}\left[\frac{\left[\widehat{V}(s)-V(s)\right]\left[\widehat{V}(t)-V(t)\right]}{V^2(1)}\mathbb{I}\left\{A_{K,\varepsilon}\right\}\right] \\
&+ K\mathbb{E}\left[\left[\widehat{V}(s)-V(s)\right]\left[\widehat{V}(t)-V(t)\right]\left[\frac{1}{\widehat{V}^2(1)}-\frac{1}{V^2(1)}\right]\mathbb{I}\left\{A_{K,\varepsilon}\right\}\right] \\
=:& B_{111}(s,t)+B_{112}(s,t).
\end{aligned}$$

We first explore $B_{111}(\cdot,\cdot)$. Doing so observe that

$$\begin{aligned}
&\left|B_{111}(s,t)-K\mathbb{E}\left[\frac{\left[\widehat{V}(s)-V(s)\right]\left[\widehat{V}(t)-V(t)\right]}{V^2(1)}\right]\right| \\
=& K\left|\mathbb{E}\left[\frac{\left[\widehat{V}(s)-V(s)\right]\left[\widehat{V}(t)-V(t)\right]}{V^2(1)}\mathbb{I}\left\{A_{K,\varepsilon}^c\right\}\right]\right| \\
\leq& \frac{K}{V^2(1)}\mathbb{E}\left\{\sup_{0\leq t\leq 1}\left|\widehat{V}(t)-V(t)\right|^2\right\}\mathbb{P}\left(A_{K,\varepsilon}^c\right) \\
\leq& \frac{K}{V^2(1)}\mathbb{E}\left\{\sup_{0\leq t\leq 1}\left|\widehat{V}(t)-V(t)\right|^2\right\}\frac{\mathbb{E}\left|\widehat{V}(1)-V(1)\right|^2}{\varepsilon^2} \\
=& O\left(K^{-1}\right).
\end{aligned}$$

On the other hand, Lemma A.3 implies

$$\sup_{0\leq s,t\leq 1}\left|K\mathbb{E}\left[\frac{\left[\widehat{V}(s)-V(s)\right]\left[\widehat{V}(t)-V(t)\right]}{V^2(1)}\right]-\frac{1}{V^2(1)}c_V(s,t)\right|\to 0.$$

This gives

$$\sup_{0\leq s,t\leq 1}|B_{111}(s,t)-\frac{1}{V^2(1)}c_V(s,t)|\to 0,\quad \text{as } K\to 0. \tag{B.3}$$

We now investigate $B_{112}(\cdot,\cdot)$. Using Lipschitz continuity of the map $x\mapsto x^{-1}$ on $[\varepsilon,\infty)$, we obtain

$$\begin{aligned}&\sup_{0\leq s,t\leq 1}|B_{112}(s,t)|\\ \leq& cK\mathbb{E}\left[\left[\widehat{V}(s)-V(s)\right]\left[\widehat{V}(t)-V(t)\right]\left[\widehat{V}^2(1)-V^2(1)\right]\mathbb{I}\left\{A_{K,\varepsilon}\right\}\right]\\ \leq& cK\mathbb{E}\left[\sup_{0\leq t\leq 1}|\widehat{V}(t)-V(t)|^2|\widehat{V}^2(1)-V^2(1)|\mathbb{I}\left\{A_{K,\varepsilon}\right\}\right]\end{aligned}$$

Applying Cauchy-Schwartz inequality, Proposition A.1 and Lemma A.1 we have

$$\sup_{0\leq s,t\leq 1}|B_{112}(s,t)|=KO\left(K^{-1}\right)O\left(K^{-1/2}\right)=O\left(K^{-1/2}\right). \tag{B.4}$$

Combining (B.3) and (B.4), implies

$$\sup_{0\leq s,t\leq 1}|B_{11}(s,t)-\frac{1}{V^2(1)}c_V(s,t)|\to 0,\quad \text{as } K\to 0. \tag{B.5}$$

A similar argument implies

$$\sup_{0\leq s,t\leq 1}|B_{12}(s,t)-\frac{V(t)}{V^3(1)}c_V(s,1)|\to 0,\quad \text{as } K\to 0, \tag{B.6}$$

$$\sup_{0\leq s,t\leq 1}|B_{13}(s,t)-\frac{V(s)}{V^3(1)}c_V(1,t)|\to 0,\quad \text{as } K\to 0, \tag{B.7}$$

$$\sup_{0\leq s,t\leq 1}|B_{14}(s,t)-\frac{V(s)V(t)}{V^4(1)}c_V(1,1)|\to 0,\quad \text{as } K\to 0. \tag{B.8}$$

Combining (B.5), (B.6), (B.7) and (B.8), we have

$$\sup_{0\leq s,t\leq 1}|B_1(s,t)-c_F(s,t)|\to 0,\quad \text{as } K\to 0. \tag{B.9}$$

The limiting result (B.9) together with (B.2) completes the proof. ∎

The next lemma relates $\mathbb{E}[\widehat{F}_i]$ to the deterministic function $F_i:=V_i/V_i(1)$. As before, we drop dependence on the index i for this result.

LEMMA B.1 *Under Assumption A.1,*

$$\|\mathbb{E}[\widehat{F}] - F\|_\infty = c_\sigma \; O(K^{-1}),$$

where c_σ, but not the O-term, depends on $\sigma(\cdot)$. More precisely, $c_\sigma = g(\|\sigma\|_\infty, \|\sigma^{-1}\|_\infty)$, for a function $g : (0,\infty) \times (0,\infty) \to (0,\infty)$ that is increasing in each component. (Notice that $\|\sigma\|_\infty$ and $\|\sigma^{-1}\|_\infty$ are finite by Assumption 4.1.)

PROOF. Let $0 < \varepsilon < V(1)$ and define the event

$$A_{K,\varepsilon} := \{|\widehat{V}(1) - V(1)| \leq \varepsilon\}.$$

Then we can decompose

$$\begin{aligned}
&\|\mathbb{E}[\widehat{F}] - F\|_\infty \leq \|\mathbb{E}[(\widehat{F} - F)\mathbb{I}\{A_{K,\varepsilon}\}]\|_\infty + \|\mathbb{E}[(\widehat{F} - F)\mathbb{I}\{A^c_{K,\varepsilon}\}]\|_\infty \\
\leq &\|\mathbb{E}[(\widehat{F} - F)\mathbb{I}\{A_{K,\varepsilon}\}]\|_\infty + \mathbb{P}(A^c_{K,\varepsilon}) \leq \|\mathbb{E}[(\widehat{F} - F)\mathbb{I}\{A_{K,\varepsilon}\}]\|_\infty + 2\exp\Big(-\frac{c_1\varepsilon^2 K}{\|\sigma\|^4_\infty}\Big).
\end{aligned}$$

Here, we have used the fact that F, $\widehat{F}$ are cdfs and hence their difference is absolutely bounded by 1. Moreover, in the last step, we have employed Lemma A.2 to bound from above the probability of the event $A^c_{K,\varepsilon}$. We can now focus on the first term on the right. A straightforward calculation shows

$$\begin{aligned}
&\|(\mathbb{E}[\widehat{F}] - F)\mathbb{I}\{A_{K,\varepsilon}\}\|_\infty = \left\|\mathbb{E}\left[\frac{[\widehat{V} - V]V(1) - V[\widehat{V}(1) - V(1)]}{\widehat{V}(1)V(1)}\mathbb{I}\{A_{K,\varepsilon}\}\right]\right\|_\infty \\
\leq &\left\|\mathbb{E}\left[\frac{\widehat{V} - V}{\widehat{V}(1)}\mathbb{I}\{A_{K,\varepsilon}\}\right]\right\|_\infty + \left\|\mathbb{E}\left[\frac{V[\widehat{V}(1) - V(1)]}{\widehat{V}(1)V(1)}\mathbb{I}\{A_{K,\varepsilon}\}\right]\right\|_\infty =: R_1 + R_2,
\end{aligned}$$

where R_1, R_2 are defined in the obvious way. For R_2 we furthermore observe the bound

$$R_2 \leq \frac{\|V\|_\infty}{V(1)}\left|\mathbb{E}\left[\frac{\widehat{V}(1) - V(1)}{\widehat{V}(1)}\mathbb{I}\{A_{K,\varepsilon}\}\right]\right| = \left|\mathbb{E}\left[\frac{\widehat{V}(1) - V(1)}{\widehat{V}(1)}\mathbb{I}\{A_{K,\varepsilon}\}\right]\right| \leq R_1.$$

Here we have used that $\max_t |V(t)| = V(1)$. We can thus focus our further analysis on R_1, which can be bounded from above by

$$\begin{aligned}
&\left\|\mathbb{E}\left[\frac{\widehat{V} - V}{V(1)}\mathbb{I}\{A_{K,\varepsilon}\} + [\widehat{V} - V]\left(\frac{1}{\widehat{V}(1)} - \frac{1}{V(1)}\right)\mathbb{I}\{A_{K,\varepsilon}\}\right]\right\|_\infty \\
\leq &\left\|\mathbb{E}\left[\frac{\widehat{V} - V}{V(1)}\mathbb{I}\{A_{K,\varepsilon}\}\right]\right\|_\infty + \left\|\mathbb{E}\left[[\widehat{V} - V]\left(\frac{1}{\widehat{V}(1)} - \frac{1}{V(1)}\right)\mathbb{I}\{A_{K,\varepsilon}\}\right]\right\|_\infty =: R_{1,1} + R_{1,2}.
\end{aligned}$$

Focusing on $R_{1,1}$ first, we observe that it can be bounded by

$$\begin{aligned}
R_{1,1} \leq &\frac{\|\mathbb{E}[(\widehat{V} - V)\mathbb{I}\{A_{K,\varepsilon}\}]\|_\infty}{V(1)} \\
\leq &\frac{\|\mathbb{E}[(\widehat{V} - V)]\|_\infty}{V(1)} + \frac{\mathbb{E}\|\widehat{V} - V\|_\infty \mathbb{P}(A^c_{K,\varepsilon})}{V(1)} =: R_{1,1,1} + R_{1,1,2}.
\end{aligned}$$

Jensen's inequality implies that

$$R_{1,1,2} \leq \frac{\{\mathbb{E}\|\widehat{V} - V\|_\infty^2\}^{1/2}\mathbb{P}(A_{K,\varepsilon}^c)}{V(1)}.$$

As before, we can use Lemma A.2 to bound the probability in $R_{1,1,2}$, which combined with the bound in Proposition A.1 implies

$$R_{1,1,2} \leq \frac{2c^{1/2}\|\sigma\|_\infty^2}{K^{1/2}V(1)} \exp\Big(-\frac{c_1\varepsilon^2 K}{\|\sigma\|_\infty^4}\Big).$$

The constant c here stems from Proposition A.1 (for $M = 1$) and is independent of σ. Now, we turn to $R_{1,1,1}$, for which we observe

$$R_{1,1,1} \leq \frac{\max_{u\in[0,1]}|\mathbb{E}\widehat{V}(u) - V(u)|}{V(1)} \leq \frac{\max_{k=1,\dots,K}\int_{(k-1)/K}^{k/K}\sigma^2(u)du}{V(1)} \leq \frac{\|\sigma^2\|_\infty}{V(1)K}.$$

This shows the desired rate for $R_{1,1}$. For $R_{1,2}$, we observe that it is upper bounded by

$$R_{1,2} \leq \frac{\mathbb{E}\Big[\|\widehat{V} - V\|_\infty|\widehat{V}(1) - V(1)|\mathbb{I}\{A_{K,\varepsilon}\}\Big]}{\varepsilon^2},$$

where we have used Lipschitz continuity of the map $x \mapsto x^{-1}$ on $[\varepsilon, \infty)$, with constant $1/\varepsilon^2$. Notice that here ε has to be sufficiently small, and we can choose it as $V(1)/2$ (so that $\widehat{V}(1)$ is also bounded away from 0). Now, employing the Cauchy-Schwartz inequality, together with the moment bound from Proposition A.1 yields a rate of $O(K^{-1})$ for the right side. Together, these considerations demonstrate that

$$\|\mathbb{E}[\widehat{F}] - F\|_\infty \leq \frac{c}{K}$$

for a sufficiently large constant $c > 0$.

The bounds that we have employed so far, namely from Proposition A.1 and Lemma A.2, depend only on $\|\sigma\|_\infty$ and are monotonically increasing in it. Moreover, c_σ depends on the factor $1/V(1)$ (explicitly and via $1/\varepsilon = 2/V(1)$). Since

$$\frac{1}{V(1)} \leq \frac{1}{\min_{t\in[0,1]}\sigma^2(t)} = \|\sigma^{-1}\|_\infty^2,$$

the constant $c_\sigma \in (0, \infty)$ depends only on σ via the values $\|\sigma\|_\infty$ and $\|\sigma^{-1}\|_\infty$ and this dependence is monotone in each norm. ∎

Lemma B.1 has two important implications for our analysis. In the context of this section, it establishes for the covariance operator, that F can be used as a centering term instead of $\mathbb{E}\widehat{F}$, i.e. that

$$\begin{aligned} &K\mathbb{E}\big(\{\widehat{F}(u) - \mathbb{E}\widehat{F}(u)\}\cdot\{\widehat{F}(v) - \mathbb{E}\widehat{F}(v)\}\big) \\ =&K\mathbb{E}\big(\{\widehat{F}(u) - F(u)\}\cdot\{\widehat{F}(v) - F(v)\}\big) + o(1). \end{aligned} \tag{B.10}$$

Here the o-term vanishes w.r.t. to the sup-norm and the non-negligible part on the right converges according to Theorem B.1. A second use of Lemma B.1 discussed later will be that any change in the functions F_i transpires through to the functions $\widehat{F}_i$ as a (detectable) mean change.

LEMMA B.2 *Under Assumption A.1,*

$$K\mathbb{E}\left\|\widehat{F}-\mathbb{E}(\widehat{F})\right\|^2 \leq c_\sigma. \tag{B.11}$$

The constant c_σ in (B.11) *is different from the c_σ in Lemma B.1, but admits the same representation in terms of the two norms.*

PROOF. The left hand side of (B.11) satisfies

$$\begin{aligned} K\mathbb{E}\left\|\widehat{F}-\mathbb{E}(\widehat{F})\right\|^2 \leq & 2K\mathbb{E}\left\|\widehat{F}-F\right\|^2 + 2K\left\|F-\mathbb{E}(\widehat{F})\right\|^2 \\ =: & B_1 + B_2. \end{aligned}$$

According to Lemma B.1, $B_2 = c_\sigma\, O(K^{-1})$. We now investigate B_1. Define the event $A_{K,\varepsilon} = \left\{\left|\widehat{V}(1)-V(1)\right| < \varepsilon\right\}$, for $\varepsilon = V(1)/2$ and observe that

$$\begin{aligned} B_1 =& 2K\mathbb{E}\left[\int_0^1 \left(\widehat{F}(s)-F(s)\right)^2 ds\right] \\ =& 2K\mathbb{E}\left[\int_0^1 \left(\widehat{F}(s)-F(s)\right)^2 \mathbb{I}\left\{A_{K,\varepsilon}^c\right\} ds\right] + 2K\mathbb{E}\left[\int_0^1 \left(\widehat{F}(s)-F(s)\right)^2 \mathbb{I}\left\{A_{K,\varepsilon}\right\} ds\right]. \end{aligned}$$

Therefore,

$$\begin{aligned} B_1 \leq & 2K\mathbb{P}\left(A_{K,\varepsilon}^c\right) + 2K\mathbb{E}\left[\int_0^1 \left(\widehat{F}(s)-F(s)\right)^2 \mathbb{I}\left\{A_{K,\varepsilon}\right\} ds\right] \\ =& 2K\mathbb{P}\left(A_{K,\varepsilon}^c\right) + 2K\mathbb{E}\int_0^1 \left(\frac{\widehat{V}(s)}{\widehat{V}(1)} - \frac{\widehat{V}(s)}{V(1)} + \frac{\widehat{V}(s)}{V(1)} - \frac{V(s)}{V(1)}\right)^2 \mathbb{I}\left\{A_{K,\varepsilon}\right\} ds \\ \leq & 2K\mathbb{P}\left(A_{K,\varepsilon}^c\right) + 4K\mathbb{E}\int_0^1 \left(\frac{\widehat{V}(s)}{\widehat{V}(1)} - \frac{\widehat{V}(s)}{V(1)}\right)^2 \mathbb{I}\left\{A_{K,\varepsilon}\right\} ds \\ & + 4K\mathbb{E}\int_0^1 \left(\frac{\widehat{V}(s)}{V(1)} - \frac{V(s)}{V(1)}\right)^2 \mathbb{I}\left\{A_{K,\varepsilon}\right\} ds =: B_{11} + B_{12} + B_{13}. \end{aligned}$$

An argument similar to (B.2) implies

$$B_{11} \leq \|\sigma\|_\infty^8 O\left(K^{-1}\right). \tag{B.12}$$

Regarding B_{12}, we have

$$
\begin{aligned}
B_{12} =& 4K\mathbb{E}\left[\int_0^1 \widehat{V}^2(s)\left(\frac{1}{\widehat{V}(1)}-\frac{1}{V(1)}\right)^2 \mathbb{I}\{A_{K,\varepsilon}\}\,ds\right] \\
\leq& 4K\mathbb{E}\left[\widehat{V}^2(1)\left(\frac{1}{\widehat{V}(1)}-\frac{1}{V(1)}\right)^2 \mathbb{I}\{A_{K,\varepsilon}\}\right] \qquad \text{(B.13)} \\
\leq& 16KV^2(1)\mathbb{E}\left[\left(\frac{1}{\widehat{V}(1)}-\frac{1}{V(1)}\right)^2 \mathbb{I}\{A_{K,\varepsilon}\}\right] \qquad \text{(B.14)} \\
\leq& 16KV^2(1)\mathbb{E}\left[\frac{4}{V^2(1)}\left(\widehat{V}(1)-V(1)\right)^2 \mathbb{I}\{A_{K,\varepsilon}\}\right] \qquad \text{(B.15)} \\
=& 64K\mathbb{E}\left[\left(\widehat{V}(1)-V(1)\right)^2 \mathbb{I}\{A_{K,\varepsilon}\}\right] \\
\leq& 64K\mathbb{E}\left[\left(\widehat{V}(1)-V(1)\right)^2\right] \leq c\|\sigma\|_\infty^4. \qquad \text{(B.16)}
\end{aligned}
$$

Inequality (B.13) is a consequence of monotonicity of the empirical quadratic variation process $\widehat{V}(\cdot)$. Inequality (B.14) is a result of restriction to the event $A_{K,\varepsilon}$. Inequality (B.15) is a consequence of Lipschitz continuity of the function $x \mapsto x^{-1}$ on $[V(1)/2, \infty)$, with Lipschitz constant $4/V(1)^2$. Finally, Lemma A.1 implies (B.16). We now turn to B_{13}. Observe that

$$
\begin{aligned}
B_{13} =& 4K\frac{1}{V^2(1)}\mathbb{E}\left[\int_0^1 \left(\widehat{V}(s)-V(s)\right)^2 \mathbb{I}\{A_{K,\varepsilon}\}\,ds\right] \\
=& 4K\frac{1}{V^2(1)}\mathbb{E}\left[\int_0^1 \left(\widehat{V}(s)-V(s)\right)^2 ds\right] \\
=& 4K\frac{1}{V^2(1)}\mathbb{E}\left[\sup_{0\leq s\leq 1}\left|\widehat{V}(s)-V(s)\right|^2\right] \leq c\|\sigma\|_\infty^4\|\sigma^{-1}\|_\infty^2, \qquad \text{(B.17)}
\end{aligned}
$$

where (B.17) is a consequence of Proposition A.1. Pooling (B.12), (B.16) and (B.17) completes the proof. ∎

We are now in the position to demonstrate weak convergence of the partial sum process $P_N^{(1)}$, see (B.1). In the following theorem, we invoke the notion of a Brownian motion $\mathbb{G}$ in a Hilbert space. For the definition and details, we refer to [28].

THEOREM B.2 *Suppose that Conditions 1, 2 of Assumption 4.1 hold and $N, K \to \infty$. Then, under $H_0^{(1)}$ (3.8), there exists a functional Brownian motion $\mathbb{G}$ in the Hilbert space $L^2[0,1]$ such that*

$$\{P_N^{(1)}(x,\cdot)\}_{x\in[0,1]} \xrightarrow{d} \{\mathbb{G}(x,\cdot)\}_{x\in[0,1]},$$

where $P_N^{(1)}$ is the partial sum process defined in (B.1)*. The process $\mathbb{G}$ is centered and characterized by the covariance in* (4.6)*.*

PROOF. We apply Theorem 1 in [28] (a weak invariance principle for triangular arrays of i.i.d. random elements in a Banach space). We validate the following three conditions of this theorem.

(K.1) The array $\sqrt{K}[\widehat{F}_i - \mathbb{E}[\widehat{F}_i]]$, $i = 1, \ldots, N$, consists of i.i.d. random elements for any fixed N (and K).

(K.2) $P_N^{(1)}(1)$ satisfies a central limit theorem in the space $L^2[0,1]$.

(K.3) For any $\epsilon > 0$, there exists an $h > 0$ (sufficiently small) such that

$$\limsup_N \mathbb{P}\Big(\|P_N^{(1)}(xh)\| \geq \epsilon \Big) < 1.$$

Condition (K.1) clearly holds because the Wiener processes W_i are i.i.d. Condition (K.3) follows directly from Markov's inequality and the fact that

$$\mathbb{E}\|P_N^{(1)}(xh)\|^2 \leq \frac{1}{N} \sum_{i=1}^{\lfloor hN \rfloor} \mathbb{E}\|\sqrt{K}[\widehat{F}_i - \mathbb{E}(\widehat{F}_i)]\|^2 \leq ch,$$

where we have used the moment bound from Lemma B.2 in the last step.

We now proceed to the proof of the above condition (K.2). For this purpose we apply a central limit theorem for triangular arrays in Hilbert spaces, Theorem 1.1 in [29]. This result has three conditions, where i)-ii) are implied by the fact that our random functions $\sqrt{K}[\widehat{F}_i - \mathbb{E}(\widehat{F}_i)]$ have a covariance that converges w.r.t. to the supremum norm to the limiting covariance and thus in particular w.r.t. the trace norm (see Theorem B.1, together with Lemma B.1). The third condition is a Linderberg-type condition, which is implied by the fact that for any $f \in L^2[0,1]$ and $\delta > 0$ it holds that

$$\mathbb{E}\Big[\langle f, \sqrt{K}[\widehat{F}_i - \mathbb{E}(\widehat{F}_i)] \rangle^2 \cdot \mathbb{I}\{ \|\sqrt{K/N}[\widehat{F}_i - \mathbb{E}(\widehat{F}_i)]\| > \delta \} \Big] \to 0.$$

This fact follows directly using the Cauchy-Schwarz inequality (essentially the Lyapunov argument) and recalling the moment bound from Lemma B.2, since

$$\begin{aligned} &\mathbb{E}\Big[\langle f, \sqrt{K}[\widehat{F}_i - \mathbb{E}(\widehat{F}_i)] \rangle^2 \mathbb{I}\{ \|\sqrt{K/N}[\widehat{F}_i - \mathbb{E}(\widehat{F}_i)]\| > \delta \} \Big] \\ \leq &\Big\{ \mathbb{E}\Big[\langle f, \sqrt{K}[\widehat{F}_i - \mathbb{E}(\widehat{F}_i)] \rangle^2 \Big] \Big\}^{1/2} \mathbb{P}(\|\sqrt{K/N}[\widehat{F}_i - \mathbb{E}(\widehat{F}_i)]\| > \delta)^{1/2}. \end{aligned}$$

Again using the Markov inequality and the second moment bound for the variable $\sqrt{K}[\widehat{F}_i - \mathbb{E}(\widehat{F}_i)]$ (see Lemma B.2) we observe that

$$\mathbb{P}(|\sqrt{K/N}[\widehat{F}_i - \mathbb{E}(\widehat{F}_i)]| > \delta) \leq c/N = o(1).$$

Thus, by Theorem 1.1 in [29] weak convergence of $P_N^{(1)}(1)$ to a centered Gaussian variable in $L^2[0,1]$ follows, and then by Theorem 1 in [28] the weak invariance principle for $P_N^{(1)}$ of this theorem. ∎

PROOF OF THEOREM 4.1: The proof is based on Theorem B.2. To demonstrate the weak convergence of $\widehat{S}^{(1)}$ (under $H_0^{(1)}$), we rewrite the statistic as

$$\begin{aligned}
\widehat{S}^{(1)} =& \frac{K}{N^2}\sum_{n=1}^{N}\int_0^1 \Big(\sum_{i=1}^{n}\widehat{F}_i(u) - \frac{n}{N}\sum_{i=1}^{N}\widehat{F}_i(u)\Big)^2 du \\
=& \frac{1}{N}\sum_{n=1}^{N}\int_0^1 \Big(\frac{1}{\sqrt{N}}\sum_{i=1}^{n}\sqrt{K}[\widehat{F}_i(u) - \mathbb{E}[\widehat{F}_i(u)]] - \frac{n}{N\sqrt{N}}\sum_{i=1}^{N}\sqrt{K}[\widehat{F}_i(u) - \mathbb{E}[\widehat{F}_i(u)]]\Big)^2 du \\
=& \int_0^1 \|P_N^{(1)}(x) - (\lfloor xN\rfloor/N)\cdot P_N^{(1)}(1)\|^2 dx \\
\xrightarrow{d}& \int_0^1 \|\mathbb{G}(x) - x\mathbb{G}(1)\|^2 dx.
\end{aligned}$$

Here, $\mathbb{G}$ denotes a functional Brownian motion (in the space $L^2[0,1]$) that characterizes the limit of $P_N^{(1)}$ and is defined in Theorem B.2. It is known that

$$\int_0^1 \|\mathbb{G}(x) - x\mathbb{G}(1)\|^2 dx \stackrel{d}{=} \sum_{\ell=1}^{\infty}\lambda_\ell \int_0^1 B_\ell(u)^2 du,$$

see, e.g. Theorem 1 in [17]. The consistency under a fixed alternative follows from Theorem 4.2 that will be proven in Section B.2. ∎

PROOF OF THEOREM 4.4: Recall the definition of the random variable w_i in (4.11). We can now define the process

$$\begin{aligned}
P_N^{(2)}(x) :=& \frac{1}{\sqrt{N}}\sum_{i=1}^{\lfloor xN\rfloor}\log(\widehat{Q}_i(1)) - \mathbb{E}[\log(\widehat{Q}_i(1))] = \frac{1}{\sqrt{N}}\sum_{i=1}^{\lfloor xN\rfloor} 2g_i + \frac{1}{\sqrt{N}}\sum_{i=1}^{\lfloor xN\rfloor} w_i - \mathbb{E}[w_i] \\
=:& \widetilde{P}_N^{(2)}(x) + R(x),
\end{aligned}$$

where we have used decomposition (4.10) in the second equality. Both terms on the right are defined in the obvious way. We first show that

$$\sup_{x\in[0,1]}|R(x)| = o_P(1), \quad \text{as } N\to\infty. \tag{B.18}$$

Corollary A.1 in [26] implies that

$$\mathbb{E}|w_i - \mathbb{E}w_i|^2 = O(1/K).$$

Since the w_i's are independent, by Doob's martingale inequality, for any $\delta > 0$,

$$\mathbb{P}\Big(\sup_x |R(x)| > \delta\Big) \le \frac{\mathbb{E}|R(1)|^2}{\delta^2} \le \frac{c}{K}.$$

As a consequence, it follows as $N, K \to \infty$ that $P_N^{(2)} = \widetilde{P}_N^{(2)} + o_P(1)$ and hence that

$$\widehat{S}^{(2)} = \frac{1}{N}\sum_{n=1}^{N}\Bigg(\sum_{i=1}^{n}\big(2g_i + \mathbb{E}[w_i]\big) - \frac{n}{N}\sum_{i=1}^{N}\big(2g_i + \mathbb{E}[w_i]\big)\Bigg)^2 + o_P(1) =: \widetilde{S}^{(2)} + o_P(1). \tag{B.19}$$

Here, the o-term is the same on both sides of the equality, which defines $\tilde{S}^{(2)}$.

It remains to show that $\widetilde{S}^{(2)} \to S^{(2)}$. This follows from Condition 4 of Assumption 4.1 and a continuous mapping argument.

The claim that $\widehat{S}^{(2)} \xrightarrow{P} \infty$ if $H_0^{(2)}$ is violated follows from Theorem 4.5 proven in Section B.2.

∎

Proofs of Propositions 4.1 and 4.2: By (B.19),

$$(\widehat{S}^{(1)}, \widehat{S}^{(2)}) = (\widehat{S}^{(1)}, \tilde{S}^{(2)}) + o_P(1).$$

Recall that $\widehat{S}^{(1)}, \tilde{S}^{(2)}$ are independent of each other, since $\widehat{S}^{(1)}$ is only a functions of the Brownian motions $W_1, W_2, ...$ (see eq. (4.2) and (4.4)) and $\tilde{S}^{(2)}$ only of the random factors $g_1, g_2, ...$ (see (B.19)). As a consequence, their limits $S^{(1)}, S^{(2)}$ are independent as well. Recalling that $\Lambda^{(i)}$ is the cdf of $S^{(i)}$, which is continuous, it follows that

$$(p^{(1)}, p^{(2)}) = (1 - \Lambda^{(1)}(\widehat{S}^{(1)}), 1 - \Lambda^{(2)}(\widehat{S}^{(2)})) \xrightarrow{d} (U_1, U_2),$$

where U_1, U_2 are independent uniformly distributed random variables on the unit interval. It is then standard to show that

$$2(\log(p^{(1)}) + \log(p^{(2)})) \xrightarrow{d} 2(\log(U_1) + \log(U_2)) \stackrel{d}{=} \chi^2_4.$$

If H_0 is violated $H_0^{(i)}$ is violated for some $i \in \{1, 2\}$. In this case $\widehat{S}^{(i)} \xrightarrow{P} \infty$, hence $\Lambda^{(i)}(\widehat{S}^{(i)}) \xrightarrow{P} 1$ and

$$\log(p^{(i)}) = -\log(1 - \Lambda^{(i)}(\widehat{S}^{(i)})) \xrightarrow{P} \infty$$

implying that $\widehat{S} \xrightarrow{P} \infty$ leading to rejection with asymptotic probability 1. ∎

B.2 Proofs of Theorems 4.2, 4.3, 4.5, 4.6 and B.3 (behavior under the alternative hypotheses)

Proof of Theorem 4.2: We begin rewriting $\widehat{S}^{(1)}$ as

$$\widehat{S}^{(1)} = \frac{K}{N^2}\sum_{n=1}^{N}\Big\|\sum_{i=1}^{n}\{\widehat{F}_i - \mathbb{E}[\widehat{F}_i]\} - \frac{n}{N}\sum_{i=1}^{N}\{\widehat{F}_i - \mathbb{E}[\widehat{F}_i]\} \tag{B.20}$$

$$- \Big(\frac{n}{N}\sum_{i=1}^{N}\mathbb{E}[\widehat{F}_i] - \sum_{i=1}^{n}\mathbb{E}[\widehat{F}_i]\Big)\Big\|^2 \tag{B.21}$$

$$= \int_0^1 \|P_N^{(1)}(x) - (\lfloor xN \rfloor / N) \cdot P_N^{(1)}(1) + h_N(x)\|^2 dx.$$

Here, $P_N^{(1)}$ denotes the partial sum process defined in (B.1) and

$$h_N(x) := -\sqrt{\frac{K}{N}}\Big(\frac{\lfloor Nx \rfloor}{N}\sum_{i=1}^{N}\mathbb{E}[\widehat{F}_i] - \sum_{i=1}^{\lfloor Nx \rfloor}\mathbb{E}[\widehat{F}_i]\Big).$$

We decompose $P_N^{(1)}$ as

$$P_N^{(1)}(x) = P_N^{(1)}(x \wedge \theta) + [P_N^{(1)}(x) - P_N^{(1)}(\theta)]\mathbb{I}\{x > \theta\}. \tag{B.22}$$

The two processes on the right-hand side involve only functions $\widehat{F}_i$ before (left term) or after the change point (right term). Moreover, they are stochastically independent of each other, since the $\widehat{F}_i$ are independent along i. Consequently, using exactly the same arguments as in the proof of Theorem B.2, we conclude that

$$\begin{aligned}\{P_N^{(1)}(x)\}_{x\in[0,\theta]} &\xrightarrow{d} \{\mathbb{G}_1(x)\}_{x\in[0,\theta]}\\ \{[P_N^{(1)}(x) - P_N^{(1)}(\theta)]\}_{x\in[\theta,1]} &\xrightarrow{d} \{\mathbb{G}_2(x-\theta)\}_{x\in[\theta,1]},\end{aligned}$$

where $\mathbb{G}_1, \mathbb{G}_2$ are independent functional Brownian motions characterized by the covariance functions

$$c_{F,1}(u,v) := \lim_{K\to\infty} K \cdot \mathbb{E}\left[\{\widehat{F}_1(u) - \mathbb{E}[\widehat{F}_1(u)]\}\{\widehat{F}_1(v) - \mathbb{E}[\widehat{F}_1(v)]\}\right],$$

$$c_{F,2}(u,v) := \lim_{K\to\infty} K \cdot \mathbb{E}\left[\{\widehat{F}_N(u) - \mathbb{E}[\widehat{F}_N(u)]\}\{\widehat{F}_N(v) - \mathbb{E}[\widehat{F}_N(v)]\}\right],$$

respectively (covariance before and after the change). These limits exist, as demonstrated in Theorem B.1. Decomposition (B.22), together with the continuous mapping theorem, implies

$$\{P_N^{(1)}(x)\}_{x\in[0,1]} \xrightarrow{d} \mathbb{G}^{(1)}(x \wedge \theta) + [\mathbb{G}^{(2)}(x-\theta)]\mathbb{I}\{x > \theta\}. \tag{B.23}$$

Next, we turn to the analysis of the deterministic function $h_N(x)$ that can be written as

$$h_N(x) = \sqrt{NK}\begin{cases} x(1-\theta)(\mathbb{E}[\widehat{F}_1] - \mathbb{E}[\widehat{F}_N]) & \text{if} \quad x \le \theta,\\ \theta(1-x)(\mathbb{E}[\widehat{F}_1] - \mathbb{E}[\widehat{F}_N]) & \text{if} \quad x > \theta.\end{cases}$$

According to Lemma B.1, $\|\mathbb{E}[\widehat{F}_i] - F_i\|_\infty = O(1/K)$, where the O-term is independent of the volatility functions $\sigma_{(j)}, \sigma_{(j)}^{-1}$. Consequently, setting

$$\tilde{h}_N(x) := \begin{cases} x(1-\theta)(F_1 - F_N)/a_N & \text{if} \quad x \le \theta,\\ \theta(1-x)(F_1 - F_N)/a_N & \text{if} \quad x > \theta,\end{cases}$$

we get the identity

$$h_N(x) = O\Big(\sqrt{\frac{N}{K}}\Big) + a_N\sqrt{NK}\cdot\tilde{h}_N. \tag{B.24}$$

Now, a simple calculation shows that

$$\begin{aligned}
\frac{F_1 - F_N}{a_N} &= \frac{\int_0^t \sigma_{(1)}^2(u)du}{a_N \int_0^1 \sigma_{(1)}^2(u)du} - \frac{\int_0^t [\sigma_{(1)}(u) + a_N\tilde{\sigma}(u)]^2 du}{a_N \int_0^1 [\sigma_{(1)}(u) + a_N\tilde{\sigma}(u)]^2 du} \\
=&\frac{\int_0^t \sigma_{(1)}^2(u)du \int_0^1 [\sigma_{(1)}(u) + a_N\tilde{\sigma}(u)]^2 du - \int_0^t [\sigma_{(1)}(u) + a_N\tilde{\sigma}(u)]^2 du \int_0^1 \sigma_{(1)}^2(u)du}{a_N \int_0^1 \sigma_{(1)}^2(u)du \int_0^1 [\sigma_{(1)}(u) + a_N\tilde{\sigma}(u)]^2 du} \\
=&\left(2\int_0^t \sigma_{(1)}^2(u)du \int_0^1 \sigma_{(1)}(u)\tilde{\sigma}(u)du + a_N \int_0^t \sigma_{(1)}^2(u)du \int_0^1 \tilde{\sigma}^2(u)du \right. \\
&\left. - 2\int_0^t \sigma_{(1)}(u)\tilde{\sigma}(u)du \int_0^1 \sigma_{(1)}^2 du - a_N \int_0^t \tilde{\sigma}^2(u)du \int_0^1 \sigma_{(1)}^2(u)du \right) \\
&\left(\int_0^1 \sigma_{(1)}^2(u)du \int_0^1 [\sigma_{(1)}(u) + a_N\tilde{\sigma}(u)]^2 du \right)^{-1} \\
\to& 2\frac{\int_0^t \sigma_{(1)}^2(u)du \int_0^1 \sigma_{(1)}(u)\tilde{\sigma}(u)du - \int_0^t \sigma_{(1)}(u)\tilde{\sigma}(u)du \int_0^1 \sigma_{(1)}^2(u)du}{\left(\int_0^1 \sigma_{(1)}^2(u)du \right)^2} =: \bar{h}(t).
\end{aligned}$$

The function in the numerator of $\bar{h}$ is not identically equal to 0. To see this, let us define

$$f(t) := \int_0^t \sigma_{(1)}^2(u)du, \qquad g(t) = \int_0^t \sigma_{(1)}(u)\tilde{\sigma}(u)du.$$

Notice that by assumption $\tilde{\sigma}/\sigma_{(1)}$ is not constant. Now, the numerator being 0 would mean that $f(t)g(1) = g(t)f(1)$, which is equivalent to g/f being constant. But this means, that $g'/f' = \tilde{\sigma}/\sigma_{(1)}$ must be constant too, which contradicts our assumption.
This means, that

$$\tilde{h}_N \to h := \begin{cases} x(1-\theta)\bar{h} & \text{if} \quad x \le \theta, \\ \theta(1-x)\bar{h} & \text{if} \quad x > \theta, \end{cases}$$

where h is a nonzero function. Notice that by assumption $a_N K \to \infty$, which implies $\sqrt{N/K} = o(a_N\sqrt{NK})$. This fact, together with the convergence $\tilde{h}_N \to h$ and identity (B.24), implies that

$$h_N = a_N\sqrt{NK}h + o(a_N\sqrt{NK}).$$

Now, recalling the representation of $\widehat{S}^{(1)}$ in (B.20), the weak convergence (B.23) and the fact that $a_N\sqrt{NK} \to \infty$, we obtain

$$\frac{\widehat{S}^{(1)}}{a_N^2 NK} \xrightarrow{d} \int_0^1 \|h(x)\|^2 dx > 0, \tag{B.25}$$

concluding our proof. ∎

PROOF OF THEOREM 4.3: Recall the definition of the change point estimator $\hat{\theta}^{(1)}$ in (4.7). Setting

$$M(n) := \Big\| \sum_{i=1}^{n} \frac{1}{N} \widehat{F}_i - \frac{n}{N^2} \sum_{i=1}^{N} \widehat{F}_i \Big\|^2$$

and $\hat{n} := N\hat{\theta}^{(1)}$, we see that

$$\hat{n} = \text{argmax}_{n=1,\dots,N}[M(n) - M(n^*)].$$

Here $n^* = \lfloor N\theta \rfloor$ and the above equality holds because $M(n^*)$ is a constant. Let us define for ease of notation the terms

$$\begin{aligned}
E(n) :=& \frac{1}{N} \sum_{i=1}^{n} \{\widehat{F}_i - \mathbb{E}[\widehat{F}_i]\} - \frac{n}{N^2} \sum_{i=1}^{N} \{\widehat{F}_i - \mathbb{E}[\widehat{F}_i]\}, \\
D(n) :=& \frac{n}{N^2} \sum_{i=1}^{N} \mathbb{E}[\widehat{F}_i] - \frac{1}{N} \sum_{i=1}^{n} \mathbb{E}[\widehat{F}_i], \\
\tilde{D}(n) :=& \frac{n(n^* - N)}{N^2} F_1 + \frac{n(N - n^*)}{N^2} F_N.
\end{aligned}$$

We now show that

$$\begin{aligned}
&\lim_{M\to\infty} \limsup_{N} \mathbb{P}\Big(\max_{n \le n^* - b_N,} M(n) - M(n^*) \ge 0 \Big) = 0, \\
&\lim_{M\to\infty} \limsup_{N} \mathbb{P}\Big(\max_{n > n^* + b_N,} M(n) - M(n^*) \ge 0 \Big) = 0
\end{aligned} \tag{B.26}$$

where $b_N := \max(M, Ma_N^{-2}/K)$ and $M > 0$. For simplicity, we focus on the first identity. Notice that we can rewrite the difference inside the maximum as

$$\begin{aligned}
&M(n) - M(n^*) = \langle D(n) - D(n^*) - E(n) + E(n^*), D(n) + D(n^*) + E(n) + E(n^*) \rangle \\
=&\{\|D(n)\|^2 - \|D(n^*)\|^2\} + \{-\|E(n)\|^2 + \|E(n^*)\|^2\} + \langle D(n) - D(n^*), E(n) + E(n^*) \rangle \\
&- \langle E(n) - E(n^*), D(n) + D(n^*) \rangle =: A_1(n) + A_2(n) + A_3(n) - A_4(n).
\end{aligned} \tag{B.27}$$

It is now enough to prove that

$$\lim_{M\to\infty} \limsup_{N} \mathbb{P}\Big(\max_{n \le n^* - b_N} A_1(n)/3 + A_i(n) \ge 0 \Big) = 0, \qquad \text{for } i = 2, 3, 4.$$

To obtain these bounds, we use the estimates from Lemma B.3. For $i = 2, 3$, we can simply show that $A_1(n)$ asymptotically dominates $A_2(n), A_3(n)$. More precisely, using Lemma B.3 parts a), b) shows that

$$A_1(n)/3 + A_2(n) \le -ca_N^2(n^* - n)/N + O_P(1/(NK)).$$

Now, $|n-n^*| \geq b_N \geq Ma_N^{-2}/K$, implies that

$$ca_N^2(n^*-n)/N \geq c\frac{M}{NK},$$

which goes to ∞ as M does, whereas the part $O_P(1/(NK))$ is independent of M. Similarly, we can use a) and c) of Lemma B.3, to see that

$$\begin{aligned}&A_1(n)/3 + A_2(n) \leq -ca_N^2(n^*-n)/N + ca_N(n-n^*)/NO_P(1/\sqrt{NK}) \\ =&\frac{a_N(n^*-n)}{N}\big(-ca_N + O_P(1/\sqrt{NK})\big).\end{aligned}$$

Since a_N is of larger order than $1/\sqrt{NK}$ by assumption, the part inside the brackets is negative, with probability converging to 1, as $N \to \infty$.
Finally, we consider the sum $A_1(n)/3 + A_4(n)$. Here a slightly more subtle argument is needed. We decompose the indices $n = 1, ..., n^* - b_N$ in blocks of size $b_N 2^\ell$ for $\ell = 1, 2...$, with the the first block consisting of the indices $n^* - b_N, n^* - b_N - 1, ..., n^* - 3b_N$, the second consisting of the indices $n^* - 4b_N - 1, ..., n^* - 7b_N$ and so on. We call these blocks $B_1, B_2, ...$ Then, we observe that

$$\max_{n \leq n^* - b_N} A_1(n)/3 + A_4(n) \geq \max_\ell \max_{n \in B_\ell} A_1(n)/3 + A_4(n).$$

Using estimate a) from Lemma B.3, we have

$$\max_{n \in B_\ell} A_1(n)/3 \geq -ca_N^2(n^* - b_N - \ell b_N)/N.$$

We can then use the bound d) from Lemma B.3 to see that

$$\begin{aligned}&\mathbb{P}(\max_{n \leq n^* - b_N} A_1(n)/3 + A_4(n) \geq 0) \\ \leq &\sum_{\ell \geq 1} \mathbb{P}(\max_{n \in B_\ell} A_1(n)/3 + A_4(n) \geq 0) \\ \leq &\sum_{\ell \geq 1} \mathbb{P}(\max_{n \in B_\ell} A_4(n) \geq ca_N^2 b_N 2^\ell/N) \\ \leq &c\sum_{\ell \geq 1} \frac{a_N^2 b_N 2^\ell/(N^2K)}{a_N^4 b_N^2 2^{2\ell-2}/N^2} \leq \sum_{\ell \geq 1} \frac{c}{M2^{\ell-2}}\end{aligned}$$

The right side does not depend on N anymore and converges to 0, as $M \to \infty$. This completes the proof. ∎

The following lemma was used in the proof of Theorem 4.3.

LEMMA B.3 *Under the assumptions of Theorem 4.2 (and so of Theorem 4.3),*

a) $A_1(n) < -ca_N^2(n^* - n)/N$.

b) $\max_{n\leq n^*}|A_2(n)| = O_P(1/(NK))$.

c) $|A_3(n)| \leq ca_N(n-n^*)/NO_P(1/\sqrt{NK})$.

d) For any $\kappa = 2, ..., N$ *it holds that* $\mathbb{E}\max_{|n-n^*|\leq\kappa}|A_4(n)|^2 \leq ca_N O_P(\sqrt{\kappa}/(N\sqrt{K}))$.

Here all Landau symbols and constants are independent of M.

PROOF. We begin by considering $A_1(n)$. More precisely, we investigate first $\tilde{A}_1(n) := \{\|\tilde{D}(n)\|^2 - \|\tilde{D}(n^*)\|^2\}$. By the definition of $\tilde{D}$,

$$\tilde{A}_1(n) = \|F_N - F_1\|^2\Big(\frac{n^2(N-n^*)^2 - (n^*)^2(N-n^*)^2}{N^4}\Big). \tag{B.28}$$

In the proof of Theorem 4.2, we have demonstrated that for a sufficiently small constant $c > 0$ it holds that $\|F_N - F_1\|^2 > ca_N^2$. Next, we consider the second factor in (B.28). Recall that $N-n^* \geq cN$ for some small enough $c > 0$ and that $n^2-(n^*)^2 = (n+n^*)(n-n^*)$, where again $n + n^* \geq n^* \geq cN$. Putting these results together yields

$$\tilde{A}_1(n) < -\frac{c(n^*-n)a_N^2}{N}.$$

Next, we consider the difference $|A_1(n) - \tilde{A}_1(n)|$, which can be expressed as

$$\Big|\|F_N - F_1\|^2 - \|\mathbb{E}\widehat{F}_N - \mathbb{E}\widehat{F}_1\|^2\Big|\Big|\frac{n^2(N-n^*)^2 - (n^*)^2(N-n^*)^2}{N^4}\Big|.$$

By analogous arguments as before, the second factor is upper bounded by $c|n-n^*|/N$. The first factor can be expressed as

$$|2\langle[\mathbb{E}\widehat{F}_N - \mathbb{E}\widehat{F}_1] - [F_N - F_1], F_N - F_1\rangle + \|[\mathbb{E}\widehat{F}_N - \mathbb{E}\widehat{F}_1] - [F_N - F_1]\|^2|,$$

which follows by a version of the third binomial formula for Hilbert spaces. Recalling Lemma B.1, we observe that

$$\|[\mathbb{E}\widehat{F}_N - \mathbb{E}\widehat{F}_1] - [F_N - F_1]\|^2 \leq \frac{c}{K^2}$$

and using additionally that $\|F_1 - F_N\| \leq c/a_N$

$$|\langle[\mathbb{E}\widehat{F}_N - \mathbb{E}\widehat{F}_1] - [F_N - F_1], F_N - F_1\rangle| \leq \frac{ca_N}{K}.$$

Employing the assumption $a_N K \to \infty$, it follows that the term $(ca_N)/K$ dominates. Hence, we conclude, that

$$|A_1(n) - \tilde{A}_1(n)| \leq \frac{c(n^*-n)a_N}{NK}$$

Since a_N is asymptotically dominated by $1/K$, it follows that $\tilde{A}_1(n)$ dominates the remainder $|A_1(n) - \tilde{A}_1(n)|$, proving the rate a).

To demonstrate b), it suffices to show the desired rate for $A_{2,1}(n) := \|E(n)\|^2$, since $\max_n \|E(n)\|^2 \geq \|E(n^*)\|^2$. Next, we recall that by definition of $E(n)$ we can decompose

$$\|E(n)\| \leq \frac{1}{N}\Big\| \sum_{i=1}^{n} \{\widehat{F}_i - \mathbb{E}[\widehat{F}_i]\}\Big\| + \Big\| \frac{n}{N^2} \sum_{i=1}^{N} \{\widehat{F}_i - \mathbb{E}[\widehat{F}_i]\}\Big\| =: A_{2,1,1}(n)(n) + A_{2,1,2}(n).$$

It suffices to show the rate $A_{2,1,1}(n), A_{2,1,2}(n) = O_P(1/\sqrt{NK})$ (uniformly in n) to get the desired result and we focus on the more difficult term $A_{2,1,1}(n)$ only. Notice that for all i it holds that

$$\mathbb{E}\|\widehat{F}_i - \mathbb{E}[\widehat{F}_i]\|^2 \leq c/K$$

(see Lemma B.1). Now, given the independence of the random variables across i it holds for any two indices $1 \leq n_1 < n_2 \leq N$, that

$$\mathbb{E}\Big\| \sum_{i=n_1}^{n_2} \widehat{F}_i - \mathbb{E}[\widehat{F}_i]\Big\|^2 \leq \frac{c(n_2 - n_1)}{K}. \tag{B.29}$$

So, using Theorem 3.1 in [32] implies that

$$\mathbb{E}\big[\| \max_n A_{2,1,1}(n)\|^2\big] \leq \frac{c}{KN},$$

with a possibly larger constant c. Notice that the cited theorem was originally formulated for real valued random variables, but the proof carries over directly to random variables on a Hilbert space. This concludes the proof of b)

c) follows by similar techniques as before. We observe that

$$A_3(n) \leq \|D(n) - D(n^*)\|\|E(n) + E(n^*)\|.$$

By the same bounds as in the last step, we conclude that $max_n\|E(n) + E(n^*)\| = O_P(1/\sqrt{NK})$. Turning to the first factor, we see that

$$\begin{aligned} \|D(n) - D(n^*)\| &\leq \|D(n) - D(n^*) - [\tilde{D}(n) - \tilde{D}(n^*)]\| + \|\tilde{D}(n) - \tilde{D}(n^*)\| \\ &=: A_{3,1}(n) + A_{3,2}(n). \end{aligned} \tag{B.30}$$

Using the analogous calculations as in a) shows that

$$\begin{aligned} A_{3,1}(n) &= \Big(\frac{(n^* - n)(N - n^*)}{N^2}\Big)\|F_1 - F_N - [\mathbb{E}\widehat{F}_1 - \mathbb{E}\widehat{F}_N]\| \\ &\leq c\frac{n^* - n}{NK}, \end{aligned}$$

where we have used Lemma B.1, to bound the second factor in the first line. Turning to $A_{3,2}(n)$, we observe (again with the same techniques as in a)), that

$$A_{3,2}(n) \leq \frac{ca_N(n^* - n)}{N}.$$

Putting these rates together and noticing that a_N dominates $1/K$, we observe that

$$A_3(n) \leq \frac{ca_N(n^* - n)}{N} O_P\Big(\frac{1}{\sqrt{NK}}\Big),$$

proving c).

Finally, we turn to d). We can decompose $A_4(n)$ into

$$A_4(n) \leq \|E(n) - E(n^*)\|\|D(n) + D(n^*)\| =: A_{4,1}(n) \cdot A_{4,2}(n).$$

$A_{4,2}(n)$ is bounded by $\|D(n)\| + \|D(n^*)\|$. Again, using Lemma B.1 we observe that

$$\|D(n)\| + \|D(n^*)\| \leq \|\tilde{D}(n)\| + \|\tilde{D}(n^*)\| + \frac{c}{K}.$$

Furthermore, by definition we have $\|\tilde{D}(n)\| \leq \|F_1 - F_N\| \leq a_N$. Since a_N dominates $1/K$ we have

$$A_{4,2}(n) \leq ca_N.$$

Next, we turn to $A_{4,1}(n)$, which using the definition of E can be upper bounded by

$$\Big\|\frac{1}{N}\sum_{i=n}^{n^*}\{\widehat{F}_i - \mathbb{E}[\widehat{F}_i]\}\Big\| + \Big\|\frac{n - n^*}{N^2}\sum_{i=1}^{N}\{\widehat{F}_i - \mathbb{E}[\widehat{F}_i]\}\Big\| =: A_{4,1,1}(n) + A_{4,1,2}(n).$$

We confine our proof to the more difficult term $A_{4,1,1}(n)$. We employ Theorem 3.1 in [32] (where the condition of the named theorem is satisfied due to (B.29)), which entails

$$\mathbb{E}[\max_{\kappa \leq n \leq n^*} \|A_{4,1,1}(n)\|^2] \leq \frac{\kappa}{N^2 K}$$

and hence that $\max_{\kappa \leq n \leq n^*} A_{4,1,1}(n) = O_P(\sqrt{\kappa}/(N\sqrt{K}))$. Combining the rates of $A_{4,1}$ and $A_{4,2}$ now yields the desired result d). ∎

Proof of Theorems 4.5 and 4.6 The proofs of these theorems follow similar steps as those for Theorems 4.2 and 4.3. The details are, however, easier given that in this case we consider real-valued time series. We have therefore decided to omit a proof of these results to avoid redundancy. ∎

Proof of Proposition 4.3 We confine the proof to the more difficult case of the mixed alternative $\sigma_{(2)} = (1 + 1/\sqrt{K})\sigma_{(1)} + \tilde{\sigma}/\sqrt{K}$, that violates both $H^{(1)}$ and $H^{(2)}$. We notice, we can rewrite

$$\sigma_{(2)} = \sigma_{(1)} + \big(\tilde{\sigma} + \sigma_{(1)}\big)/\sqrt{K}$$

and by the proof of Theorem 4.2 (see the convergence (B.25) in particular), we observe that

$$\frac{\widehat{S}^{(1)}}{N} \xrightarrow{P} c_1 > 0,$$

where $c_1 > 0$ is a constant depending on $\sigma_{(1)}$ and $\tilde{\sigma}$. By a similar line of argumentation, we can show that

$$\frac{K\widehat{S}^{(2)}}{N} \xrightarrow{P} c_2 > 0,$$

where $c_2 > 0$ is a constant that depends on $\sigma_{(1)}$ and $\tilde{\sigma}$ as well. Now, notice that we can rewrite the following fraction of p-values in terms of the distribution functions

$$\frac{p^{(1)}}{p^{(1)} + p^{(2)}} = \frac{1 - \Lambda^{(1)}[\widehat{S}^{(1)}]}{(1 - \Lambda^{(1)}[\widehat{S}^{(1)}]) + (1 - \Lambda^{(2)}[\widehat{S}^{(2)}])}. \tag{B.31}$$

If we can prove for an arbitrarily small, but fixed $c_3 > 0$ that

$$\frac{1 - \Lambda^{(1)}[\widehat{S}^{(1)}]}{1 - \Lambda^{(2)}[\widehat{S}^{(2)}]} = \mathcal{O}(\exp(-c_3 N)), \tag{B.32}$$

it follows that the ratio in (B.31) is exponentially decaying and hence, by definition of $\hat{\theta}$ in (4.20) that

$$\hat{\theta} = \hat{\theta}^{(1)} + \mathcal{O}_P\left(\frac{1}{N}\right).$$

This, together with the fact that

$$\hat{\theta}^{(1)} = \theta + \mathcal{O}_P\left(\frac{1}{N}\right),$$

(Theorem 4.3) would already imply the desired result. Now, to prove (B.32), we notice that by the continuous mapping theorem

$$\frac{1 - \Lambda^{(1)}[\widehat{S}^{(1)}]}{1 - \Lambda^{(2)}[\widehat{S}^{(2)}]} \frac{1 - \Lambda^{(2)}[(N/K)c_2]}{1 - \Lambda^{(1)}[Nc_1]} = \frac{1 - \Lambda^{(1)}[N(\widehat{S}^{(1)}/N)]}{1 - \Lambda^{(2)}[(N/K)(K\widehat{S}^{(2)}/N)]} \frac{1 - \Lambda^{(2)}[(N/K)c_2]}{1 - \Lambda^{(1)}[Nc_1]} \xrightarrow{P} 1.$$

Consequently, it suffices to analyze the deterministic scaling ratio and show that

$$\frac{1 - \Lambda^{(1)}[Nc_1]}{1 - \Lambda^{(2)}[(N/K)c_2]} = \mathcal{O}(\exp(-c_3 N)),$$

to prove (B.32). Since $K \to \infty$ as $N \to \infty$, this implies that for any arbitrarily small constant $c_L > 0$ there exists an $N_0 > 0$ such that for all $N > N_0$ we have $c_2/K < c_1 c_L$. It follows by monotonicity of the denominator and Theorem B.3 that as $N \to \infty$

$$\frac{1 - \Lambda^{(1)}[Nc_1]}{1 - \Lambda^{(2)}[(N/K)c_2]} \leq \frac{1 - \Lambda^{(1)}[Nc_1]}{1 - \Lambda^{(2)}[Nc_1 c_L]} \leq \exp(-c_3 N)$$

for an adequate choice of c_L and a sufficiently small constant $c_3 > 0$, which concludes the proof. ∎

THEOREM B.3 *Under Assumption 4.1, there exist constants* $0 < c_L < c_U < \infty$, *and* $c'_L c'_U > 0$ *such that for sufficiently large* x

$$\frac{1-\Lambda^{(1)}(x)}{1-\Lambda^{(2)}(c_L x)} \leq \exp(-c'_L\, x), \qquad \frac{1-\Lambda^{(1)}(x)}{1-\Lambda^{(2)}(c_U x)} \geq \exp(c'_U\, x).$$

The constants in this Theorem depend on the eigenvalues λ_i *in Theorem 4.1 and the long–run variance* λ *in Theorem 4.4.*

PROOF. We confine the proof to the second equation (the first one follows by the same reasoning). Let us define $B := \int_0^1 [\mathbb{B}(x)]^2 dx$ and let $B_1, B_2, ...$ be i.i.d. copies of B. Then

$$S^{(1)} \stackrel{d}{=} \sum_{i=1}^{\infty} \lambda_i B_i, \qquad S^{(2)} \stackrel{d}{=} (4\lambda)B,$$

where $\lambda_1 \geq \lambda_2 \geq ...$ are defined in Theorem 4.1 and λ in Theorem 4.4. For simplicity, we set $\lambda = 1$. According to eq. (49) in [42] we can write for large x

$$1-\Lambda^{(2)}(x) = \frac{c_1 \exp(-c_2 x)(1+O(1/x))}{\sqrt{x}}, \tag{B.33}$$

for absolute constants c_1, c_2. Now, let us investigate $\Lambda^{(1)}$. Therefore, let us assume that there exists an $\eta > 0$ such that $c_3 := \sum_{i\geq 1} \lambda_i i^\eta < \infty$ and set $c_L := 3c_3$. Now, it follows that

$$\begin{aligned}
1-\Lambda^{(1)}(c_L x) &= \mathbb{P}\Big(S^{(1)} > c_L x\Big) = \mathbb{P}\Big(\sum_{i=1}^{\infty} \lambda_i B_i > (c_L x/c_3) \sum_{i=1}^{\infty} \lambda_i i^\eta\Big) \\
&= \mathbb{P}\Big(\sum_{i=1}^{\infty} \lambda_i (B_i - 3x i^\eta) > 0\Big) \leq \sum_{i=1}^{\infty} \mathbb{P}\Big(\lambda_i B_i > 3x i^\eta \lambda_i\Big) \\
&= \sum_{i=1}^{\infty} \frac{c_1 \exp(-3c_2 i^\eta x)(1+O(1/x))}{\sqrt{[3x i^\eta]}} \\
&= \frac{\exp(-2c_2 x)}{\sqrt{x}} \sum_{i=1}^{\infty} \frac{c_1 \exp(-c_2 x[3i^\eta - 2])(1+O(1/x))}{\sqrt{3i^\eta}}.
\end{aligned}$$

The series on the right is of size $O(1)$ and the first factor decays at a rate $\exp(-2c_2 x)$, which comparing it to (B.33) yields the desired result.

Finally, we have to show that the eigenvalues $(\lambda_i)_{i\in\mathbb{N}}$ have indeed the property that for some $\eta > 0$ it holds that $\sum_{i\geq 1} \lambda_i i^\eta < \infty$. This follows directly from the fact that the asymptotic covariance kernel of $\widehat{F}_1$ is Lipschitz continuous (defined in Theorem B.1; the verification of the Lipschitz property is easy). It thus satisfies the conditions of Theorem 3.2 in [16] and hence there exists a sufficiently large constant $c > 0$ that $|\lambda_i| \leq c i^{-2}$ ∎

C Implementation

In this section, we provide details on the implementation of our approach. In the following, for a matrix $A \in \mathbb{R}^{L\times K}$, we refer to the ℓ–th row as $A[\ell,]$, to the k–th column as $A[,k]$ and to the entry $A_{\ell,k}$ as $A[\ell,k]$. When applying a function $f : \mathbb{R} \to \mathbb{R}$ to a vector or matrix, we apply it entry-wise. The Euclidean norm and inner product are denoted by $\|\cdot\|$ and $\langle\cdot,\cdot\rangle$, respectively.

We begin by transforming our raw data into estimates of the quadratic variation process, $\widehat{Q}$ (appearing in (3.6)) in Algorithm 1.

Algorithm 1 Empirical Quadratic Variation $\{\widehat{Q}_i(k/K) : i = 1,\dots,N,\ k = 1,\dots,K\}$

input: N (number of curves), K (number of measurements), $R \in \mathbb{R}^{N\times(K+1)}$ (matrix of observations $\{R_i(k/K) : i = 1,\dots,N,\ k = 0,\dots,K\}$)
output: Quadratic variation matrix $Q \in \mathbb{R}^{N\times K}$
function Q-Functions(N, K, R)
  Define $SI \in \mathbb{R}^{N\times K}$ with 0 entries
  for $l = 1,\dots,K$ **do**
    $SI[,l] := (R[,l+1] - R[,l])^2$
  end for
  Define $Q \in \mathbb{R}^{N\times K}$ with $Q[,1] = SI[,1]$, 0 otherwise
  for $l = 2,\dots,K$ **do**
    $Q[,l] := Q[,l-1] + SI[,l]$
  end for
return Q
end function

Algorithm 2 below transforms the estimated quadratic variation into the normalized version $\widehat{F}_i$, appearing in (4.1). This quantity is used as inputs of the test of hypothesis $H_0^{(1)}$.

Algorithms 3, 4 and 5, respectively, calculate the CUSUM statistic $\widehat{S}^{(1)}$, the empirical p-value $\hat{p}^{(1)}$ and the estimator $\hat{\theta}^{(1)}$. In the estimation of $\hat{p}^{(1)}$ there exist two user-determined parameters: r, the number of simulations for the limiting distribution and B, a dimension reduction parameter that is used to truncate the infinite sum in the definition of $\widehat{S}^{(1)}$ (see (4.5)). As defaults, we recommend using $r = 1000$ and choosing B large enough such that at least 95% of the empirical variance is explained. More precisely, we pick B as the smallest value s.t.

$$\frac{\sum_{b=1}^{B}\hat{\lambda}_b}{Tr[\hat{c}_F]} \geq 95\%$$

where Tr refers to the trace of a covariance kernel $\hat{c}_F$, defined as $\int_0^1 \hat{c}_F(u,u)du$. The empirical eigenvalues and the estimator $\hat{c}_F$ are defined in Algorithm 4.

Algorithm 2 $\{\widehat{F}_i(k/K) : i = 1, \ldots, N,\ k = 1, \ldots, K\}$

input: N (number of curves), K (number of measurements), $Q \in \mathbb{R}^{N \times K}$ (output of Algorithm 1)
output: The standardized quadratic variation matrix $F \in \mathbb{R}^{N \times K}$
function F-FUNCTIONS(N, K, Q)
  Define $F \in \mathbb{R}^{N \times K}$ with 0 entries
  for $n = 1, \ldots, N$ **do**
    for $k = 1, \ldots, K$ **do**
      $F[n, k] := Q[n, k]/Q[n, K]$
    end for
  end for
return F
end function

Algorithm 3 Test statistic $\widehat{S}^{(1)}$

input: N (number of curves), K (number of measurements), $F \in \mathbb{R}^{N \times K}$ (output of Algorithm 2)
output: Statistic $\widehat{S}^{(1)}$
function F-CUSUM(N, K, F)
  Define $PS \in \mathbb{R}^{N \times K}$ with $PS[1,] := F[1,]$ and otherwise 0 entries
  for $n = 2, \ldots, N$ **do**
    $PS[n,] := PS[n-1,] + F[n,]$
  end for
  Define $S = 0$
  for $n = 1, \ldots, N$ **do**
    $S = S + \|PS[n,] - (n/N)PS[N,]\|^2$
  end for
  $\widehat{S}^{(1)} = S/N^2$
return $\widehat{S}^{(1)}$
end function

Algorithm 4 Empirical p-value $\hat{p}^{(1)}$

input: N (number of curves), K (number of measurements), $F \in \mathbb{R}^{N\times K}$ (output of Algorithm 2), B (dimension reduction parameter), r (simulation number for p-values)
output: $\hat{p}^{(1)}$, empirical p-value for statistic $\widehat{S}^{(1)}$
function F-Cov(N, K, F, B)
  Define $C \in \mathbb{R}^{K\times K}$ with 0-entries
  for $n = 2, \ldots, N$ **do**
    $C = C + (F[n,] - F[n-1,]) \cdot (F[n,] - F[n-1,])^{\top}$.
  end for
  Collect the B largest eigenvalues of $\hat{c}_F := C/(2NK)$ in the vector $v_\lambda := (\hat{\lambda}_1, \ldots, \hat{\lambda}_B)$
return v_λ
end function
function Int-BB(B, N)
  Define $v_B \in \mathbb{R}^B$ with 0-entries
  for $rep = 1, \ldots, B$ **do**
    Define $v_{rep} = (B_{rep}(0), B_{rep}(1/N), B_{rep}(2/N), \ldots, B_{rep}(1))$ (where $B_1, \ldots, B_B$ are independent Brownian Bridges)
    $v_B[rep] = \|v_{rep}\|^2/N$
  end for
return v_B
end function
function pval-1(N, K, B, r, F)
  Define $v_r \in \mathbb{R}^r$ with 0-entries
  Calculate $v_\lambda =$ F-Cov(N, K, F, B)
  for $rep = 1, \ldots, r$ **do**
    Generate fresh $v_B =$Int-BB(B, N)
    $v_r[rep] = \langle v_\lambda, v_B \rangle$
  end for
  Order v_r with entries in decreasing order
  Calculate $c =$F-CUSUM(N, K, F)
  Determine $i^* = \text{argmin}\{|c - v_r[i]| : i = 1, \ldots, r\}$
  $\hat{p}^{(1)} := i^*/r$
return $\hat{p}^{(1)}$
end function

Algorithm 5 Change point estimator $\hat{\theta}^{(1)}$

input: N (number of curves), K (number of measurements), $F \in \mathbb{R}^{N\times K}$ (output of Algorithm 2)
output: Change point estimator $\hat{\theta}^{(1)}$
function F-CP(N, K, F)
  Define $PS \in \mathbb{R}^{N\times K}$ with $PS[1,] := F[1,]$ and otherwise 0 entries
  for $n = 2, \ldots, N$ **do**
    $PS[n,] := PS[n-1,] + F[n,]$
  end for
  Define $v_N \in \mathbb{R}^N$ with 0 entries
  for $n = 1, \ldots, N$ **do**
    Set $v_N[n] = \|PS[n,] - (n/N)PS[N,]\|^2$
  end for
  $n^* = \text{argmax}\{v_N[n] : n = 1, \ldots, N\}$
  $\hat{\theta}^{(1)} = n^*/N$
return $\hat{\theta}^{(1)}$
end function

Next, we turn to the test of $H_0^{(2)}$, beginning with the calculation of the logarithmized total variation $\log(\widehat{Q}_i(1))$, appearing in (4.10), then calculating the CUSUM statistic $\widehat{S}^{(2)}$ in (4.12), and subsequently approximating the p-value. These are implemented in Algorithms 6, 7 and 8. Algorithm 9 entails the (potential) time change $\hat{\theta}^{(2)}$. When approximating the p-value, we call the LAMBDA-function, that gives an estimate of the long-run variance (4λ), where λ is defined in (4.14). There exist many preimplemented methods in statistical softwares to approximate the long–run variance and hence we do not select any specific method.

Algorithm 6 $\{\log \widehat{Q}_i(1) : i = 1, \ldots, N\}$

input: N (number of curves), $Q \in \mathbb{R}^{N\times K}$ (output of Algorithm 1)
output: Log total quadratic variation, $LTQ \in \mathbb{R}^N$
function LOG–TQV(N, Q)
  Define $LTQ \in \mathbb{R}^N$ with $LTQ = \log Q[, K]$
return LTQ
end function

Algorithm 7 Test statistic $\widehat{S}^{(2)}$

input: N (number of curves), LTQ (output of Algorithm 6)
output: Statistic $\widehat{S}^{(2)}$
function LOGQ-CUSUM(N, LTQ)
  Define $PS \in \mathbb{R}^N$ with $PS[1] = LTQ[1]$ and otherwise 0 entries
  for $n = 2, \ldots, N$ **do**
    $PS[n] = PS[n-1] + LTQ[n]$
  end for
  Define $S = 0$
  for $n = 1, \ldots, N$ **do**
    $S = S + |PS[n] - (n/N)PS[N]|^2$
  end for
  $\widehat{S}^{(2)} = S/N^2$
return $\widehat{S}^{(2)}$
end function

Algorithm 8 Empirical p-value $\hat{p}^{(2)}$

input: N (number of curves), LTQ (output of Algorithm 6), r (simulation number for p-values)
output: $\hat{p}^{(2)}$, p-value for statistic $\widehat{S}^{(2)}$
function PVAL-2(N, B, r, LTQ)
  Calculate $\tilde{\lambda}$ = LAMBDA(LTQ)
  Define $v_r \in \mathbb{R}^r$ with 0-entries
  for $rep = 1, \ldots, r$ **do**
    Generate fresh v_B :=Int-BB$(1, N)$
    $v_r[rep] = \tilde{\lambda} \cdot v_B$
  end for
  Order v_r with entries in decreasing order
  Calculate c =LOGQ-CUSUM(N, K, F)
  Determine $i^* = \text{argmin}\{|c - v_r[i]| : i = 1, \ldots, r\}$
  $\hat{p}^{(2)} = i^*/r$
return $\hat{p}^{(2)}$
end function

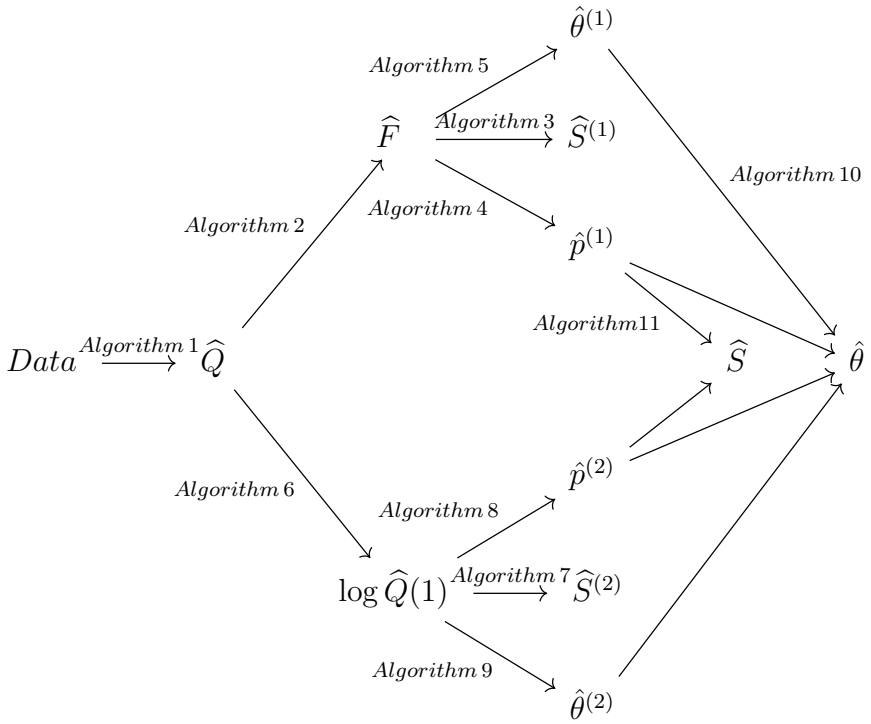

Figure 4: *Flowchart of our inference procedure, with arrows indicating input relations.*

Algorithm 9 Change point estimator $\hat{\theta}^{(2)}$

---

input: N (number of curves), LTQ (output of Algorithm 6)
output: Change point estimator $\hat{\theta}^{(2)}$
function LOGQ-CP(N, LTQ)
  Define $PS \in \mathbb{R}^{N \times K}$ with $PS[1] = LTQ[1]$ and otherwise 0 entries
  for $n = 2, \ldots, N$ **do**
    $PS[n] = PS[n-1] + LTQ[n]$
  end for
  Define $v_N \in \mathbb{R}^N$ with 0 entries
  for $n = 1, \ldots, N$ **do**
    Set $v_N[n] = |PS[n,] - (n/N)PS[N,]|^2$
  end for
  $n^* = \text{argmax}\{v_N[n] : n = 1, \ldots, N\}$
  $\hat{\theta}^{(2)} = n^*/N$
return $\hat{\theta}^{(2)}$
end function

Finally, in Algorithms 10 and 11, we combine the p-values and change point estimators to the pooled change point estimator and the pooled test of H_0. As before, we call $q_{1-\alpha}$ the upper α-quantile of the chi-squared distribution with four degrees of freedom.

Algorithm 10 Change point estimator $\hat{\theta}$

input: $\hat{\theta}^{(1)}$ (output of Algorithm 5), $\hat{\theta}^{(2)}$ (output of Algorithm 9), $\hat{p}^{(1)}$ (output of Algorithm 4), $\hat{p}^{(2)}$ (output of Algorithm 8)
output: Change point estimator $\hat{\theta}$
function CP($\hat{\theta}^{(1)}$,$\hat{\theta}^{(2)}$,$\hat{p}^{(1)}$,$\hat{p}^{(2)}$)
  Define $\hat{\theta} = \frac{\hat{p}^{(1)}}{\hat{p}^{(1)}+\hat{p}^{(2)}}\hat{\theta}^{(2)} + \frac{\hat{p}^{(2)}}{\hat{p}^{(1)}+\hat{p}^{(2)}}\hat{\theta}^{(1)}$
  return $\hat{\theta}$
end function

Algorithm 11 The Statistic $\hat{S}$ and test decision

input: $\hat{p}^{(1)}$ (output of Algorithm 4), $\hat{p}^{(2)}$ (output of Algorithm 8), α (nominal level)
output: Statistic $\hat{S}$, and test decision
function S($\hat{p}^{(1)}$,$\hat{p}^{(2)}$, α)
  $\widehat{S} = -2\{\log(\hat{p}^{(1)}) + \log(\hat{p}^{(2)})\}$
  if $\widehat{S} > q_{1-\alpha}$ **then**
    Define $decision = 1$
  else
    Define $decision = 0$
  end if
return $(\widehat{S}, decision)$
end function

D Details of the computation of test statistics and critical values in Section 5

D.1 Additional information related to Section 5.1

Computation of $\int_0^t \sigma(u)dW(u)$ We can avoid the numerical integral method (e.g. Euler-Maruyama) to calculate the integral $\int_0^t \sigma(u)dW(u)$. According to Dambis-Dubins-Schwarz theorem, see e.g. Section 5.3.2 in [30], any continuous local martingale M can be written as a "time-changed" Brownian motion. In particular, if W is a Brownian motion and σ is W-integrable then the result can be applied to $X(t) = \int_0^t \sigma(u)dW_i(u)$. This gives

$$\forall t, \quad \int_0^t \sigma(u)dW(u) = W\left(\int_0^t \sigma^2(u)du\right) \quad a.s.$$

Thus, generating from $X(t) = \int_0^t \sigma(u)dW(u)$ reduces to generating from Brownian motion. Additionally, Brownian motion admits independent and normally distributed increments. It is enough to generate independent normal variables (precise) and then sum them up (fast) to obtain discrete observations of Brownian paths. Specifically, the increments are in the form

$$\int_0^t \sigma(u)dW(u) - \int_0^s \sigma(u)dW(u) = W\left(\int_0^t \sigma^2(u)du\right) - W\left(\int_0^s \sigma^2(u)du\right), \quad 0 \le s < t < 1,$$

that are normally distributed with mean zero and variance $\int_s^t \sigma^2(u)du$, and non-overlapping increments are independent.

Based on our discretization on t, the increment can be further spelled out as

$$\begin{aligned} X(t_k) - X(t_{k-1}) &= W\left(\int_0^{t_k} \sigma^2(u)du\right) - W\left(\int_0^{t_{k-1}} \sigma^2(u)du\right) \\ &= W(G(t_k)) - W(G(t_{k-1})), \quad k = 1, ..., K. \end{aligned}$$

Denote such increment as $d(t_k) = X(t_k) - X(t_{k-1})$. We can simulate $d(t_k)$ by independent normal variables, i.e. $d(t_k) \sim \mathcal{N}(0, G(t_k) - G(t_{k-1}))$. Therefore, the trajectory of $X(t_k)$ is the summation of those independent normal variables

$$X(t_k) = \sum_{s=1}^{k} d(t_k), \quad k = 1, ..., K.$$

Computation of $\int_0^1 B^2(u)du$ Following [20], we use the expansion discussed in [40], pp 210–211, to approximate the squared integral of Brownian bridge,

$$\int_0^1 B^2(u)du \approx \sum_{j=1}^{J} \frac{Z_j^2}{j^2\pi^2},$$

where $\{Z_j\}_{j=1}^{\infty}$ are i.i.d. standard normal random variables. There is thus no need to simulate the trajectories and to perform numerical integration. In our work, we used $J = 500$.

Simulation procedure for testing a shape change $H_0^{(1)}$

1. Simulate the data $\{R_i(t_k), i = 1, ..., N,\ k = 0, ..., K\}$ based on the DGP.

2. Calculate the realized quadratic variation processes

$$\widehat{Q}_i(t) = \sum_{k=1}^{K} |R_i(t_k) - R_i(t_{k-1})|^2 \mathbb{I}\{t_k \le t\}, \qquad t \in [0, 1], \qquad i = 1, 2, ..., N.$$

3. Calculate the empirical standardized quadratic variation

$$\widehat{F}_i(t) = \frac{\widehat{Q}_i(t)}{\widehat{Q}_i(1)}, \qquad t \in [0,1], \qquad i = 1, 2, ..., N.$$

4. Calculate the test statistic $\widehat{S}^{(1)}$

$$\widehat{S}^{(1)} := \frac{K}{N^2} \sum_{n=1}^{N} \int_0^1 \Big(\sum_{i=1}^{n} \widehat{F}_i(u) - \frac{n}{N} \sum_{i=1}^{N} \widehat{F}_i(u) \Big)^2 du.$$

5. We use the following first-order difference estimator(FDE)[1]:

$$c_F(u,v) = \frac{1}{2(N-1)} \sum_{i=2}^{N} \{F_i(u) - F_{i-1}(u)\} \{F_i(v) - F_{i-1}(v)\}.$$

6. Collect the first largest B eigenvalues $\{\lambda_1, ..., \lambda_B\}$ of $c_F(u,v)$ so that

$$\frac{\sum_{\ell=1}^{B} \lambda_\ell}{\sum_{\ell=1}^{K} \lambda_\ell} \geq 95\%.$$

7. Approximate the limit distribution of $\widehat{S}^{(1)}$ by simulating the following quantity by $r = 5000$ times

$$\sum_{\ell=1}^{B} \lambda_\ell \sum_{j=1}^{J} \frac{Z_j^2}{j^2 \pi^2},$$

   and use its empirical distribution to obtain the $\hat{p}^{(1)}$.

8. If reject $H_0^{(1)}$, perform the change point estimator

$$\hat{\theta}^{(1)} = \frac{1}{N} \operatorname*{argmax}_{n \in \{1,...,N\}} \int_0^1 \Big(\sum_{i=1}^{n} \widehat{F}_i(u) - \frac{n}{N} \sum_{i=1}^{N} \widehat{F}_i(u) \Big)^2 du.$$

9. Repeat Steps (1)–(8) for $M = 5000$ times to obtain the empirical rejection rate.

[1] I think here we can divide by $2(N-1)$, rather than $2N$. This is because there are only $N-1$ of first-order difference observations. Based on a small-scale simulation (not reported), I notice that dividing by $2(N-1)$ gives slightly better size than dividing by $2N$.

Simulation procedure for testing a change in total volatility $H_0^{(2)}$

1. Use $\left\{\widehat{Q}_i(t)\right\}_{i=1}^N$ calculated above to obtain $\left\{\log \widehat{Q}_i(1)\right\}_{i=1}^N$.

2. Calculate the test statistic $\widehat{S}^{(2)}$

$$\widehat{S}^{(2)} := \frac{1}{N^2}\sum_{n=1}^{N}\Big(\sum_{i=1}^{n}\log(\widehat{Q}_i(1)) - \frac{n}{N}\sum_{i=1}^{N}\log(\widehat{Q}_i(1))\Big)^2.$$

3. Calculate the long-run variance of $\left\{\log \widehat{Q}_i(1)\right\}_{i=1}^N$ by

$$\hat{\lambda} = \hat{\gamma}(0) + 2\sum_{h=1}^{N-1} K\left(\frac{h}{H}\right)\hat{\gamma}(h),$$

   where

$$\hat{\gamma}(h) = \frac{1}{N}\sum_{i=1}^{N-h}\left(\log \widehat{Q}_i(1) - \frac{1}{N}\sum_{i=1}^{N}\log \widehat{Q}_i(1)\right)\left(\log \widehat{Q}_{i+h}(1) - \frac{1}{N}\sum_{i=1}^{N}\log \widehat{Q}_i(1)\right).$$

4. Approximate the limit distribution of $\widehat{S}^{(2)}$ by simulating the following quantity by $r = 5000$ times

$$\hat{\lambda}\sum_{j=1}^{J}\frac{Z_j^2}{j^2\pi^2},$$

   and use its empirical distribution to obtain the $\hat{p}^{(2)}$.

5. If reject $H_0^{(2)}$, perform the change point estimator

$$\hat{\theta}^{(2)} = \frac{1}{N}\operatorname*{argmax}_{n\in\{1,\ldots,N\}}\Big(\sum_{i=1}^{n}\log(\widehat{Q}_i(1)) - \frac{n}{N}\sum_{i=1}^{N}\log(\widehat{Q}_i(1))\Big)^2,$$

6. Repeat Steps (1)–(5) for $M = 5000$ times to obtain the empirical rejection rate.

Simulation procedure for testing the global null hypothesis H_0

1. Calculate the test statistic $\widehat{S}$ by

$$\widehat{S} = -2\{\log(\hat{p}^{(1)}) + \log(\hat{p}^{(2)})\}.$$

2. Use its limit distribution χ_4^2 to obtain the p-value of the global test.

3. If reject H_0, perform the pooled change point estimator

$$\hat{\theta} = \frac{\hat{p}^{(1)}}{\hat{p}^{(1)}+\hat{p}^{(2)}}\hat{\theta}^{(2)} + \frac{\hat{p}^{(2)}}{\hat{p}^{(1)}+\hat{p}^{(2)}}\hat{\theta}^{(1)}.$$

4. Repeat Steps (1)–(3) for $M = 5000$ times to obtain the empirical rejection rate.

D.2 Distribution of change point estimators

Now we validate the convergence of the change point estimators, $\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}$, under different alternative hypotheses. Figure 5 provides the violin plot (with included boxplot) for $\hat{\theta}_1$ under $H_{A,1}$ with $\theta = 0.5$. We can observe that $\hat{\theta}_1$ converges to the true $\theta = 0.5$ as N and K increases. Figure 6 provides the violin plot (with included boxplot) for $\hat{\theta}_2$ under $H_{A,2}$ with $\theta = 0.5$. As K is not relevant for $H_{A,2}$, we only present $\hat{\theta}_2$ in terms of different N, and there is a converge in $\hat{\theta}_2$ to the true $\theta = 0.5$. Figure 7 provides the violin plot (with included boxplot) for $\hat{\theta}$ under $H_{A,3}$ with $\theta = 0.5$. We can observe that $\hat{\theta}$ is already very close to $\theta = 0.5$, even under small N and small K. The density is very concentrated around $\theta = 0.5$, and get more concentrated for larger N small K.

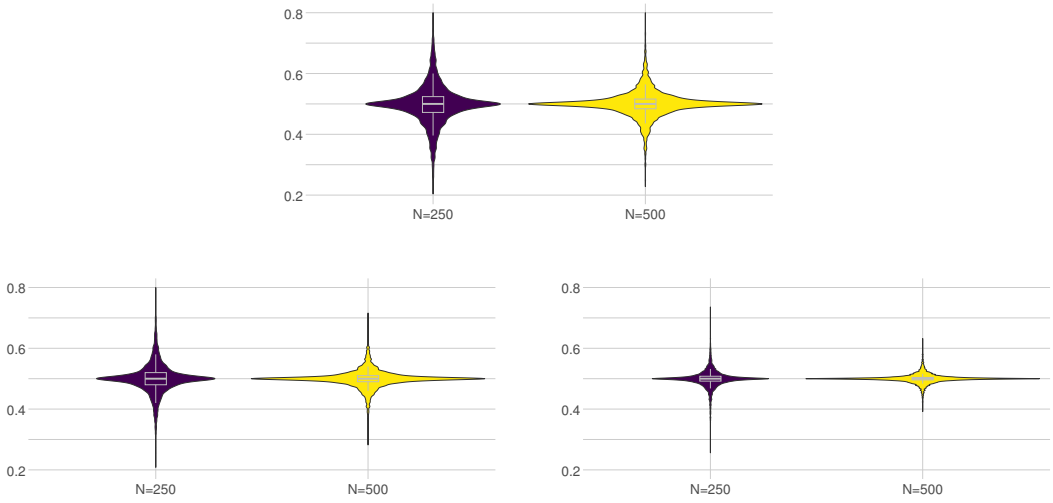

Figure 5: Distribution of $\hat{\theta}_1$ under $H_{A,1}$ with $\theta = 0.5$. Top panel: $K = 26$; Bottom left panel: $K = 39$; Bottom right panel: $K = 78$.

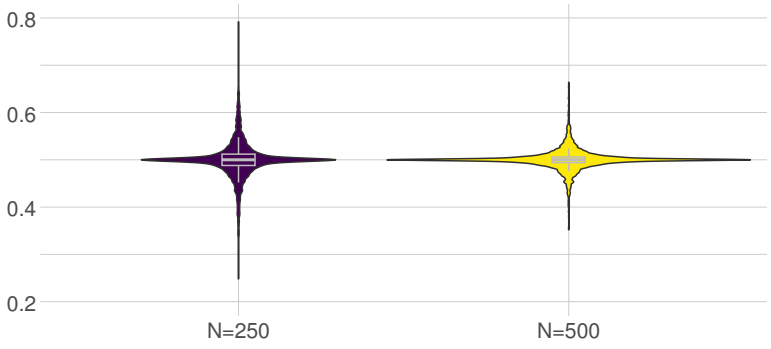

Figure 6: Distribution of $\hat{\theta}_2$ under $H_{A,2}$ with $\theta = 0.5$.

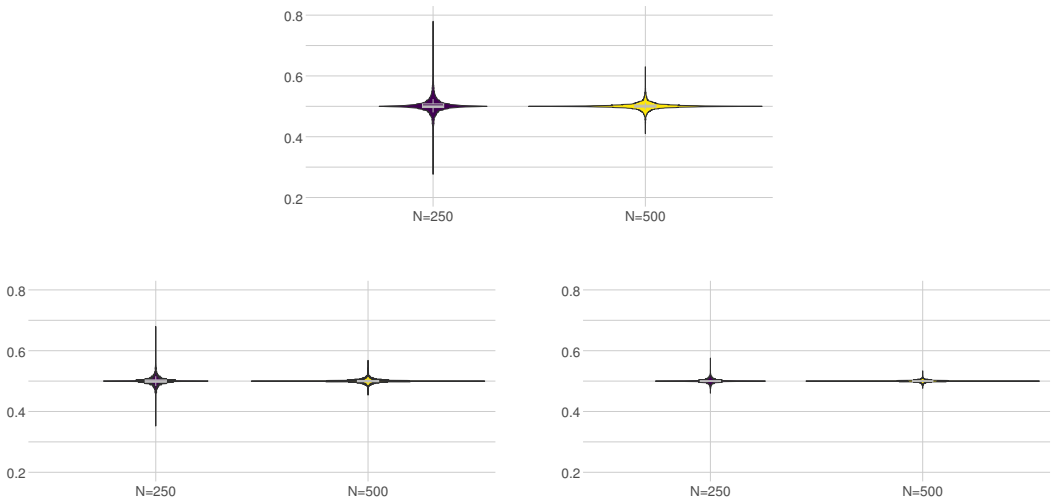

Figure 7: Distribution of $\hat{\theta}$ under $H_{A,3}$ with $\theta = 0.5$. Top panel: $K = 26$; Bottom left panel: $K = 39$; Bottom right panel: $K = 78$.