Computational Statistics and Data Analysis 210 (2025) 108196

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

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

# A goodness-of-fit test for geometric Brownian motion

Daniel Gaigall [a,b,*], Philipp Wübbolding [a]

[a] Faculty 09 - Medical Engineering and Technomathematics, FH Aachen - University of Applied Sciences, Heinrich-Mußmann Straße 1, 52428 Jülich, Germany

[b] House of Insurance, Leibniz University Hannover, Welfengarten 1, 30167 Hannover, Germany

ARTICLE INFO

*MSC:*
62G10

*Keywords:*
Black-Scholes model
Brownian motion
Functional data
Geometric Brownian motion
Goodness-of-fit test

ABSTRACT

A new goodness-of-fit test for the composite null hypothesis that data originate from a geometric Brownian motion is studied in the functional data setting. This is equivalent to testing if the data are from a scaled Brownian motion with linear drift. Critical values for the test are obtained, ensuring that the specified significance level is achieved in finite samples. The asymptotic behavior of the test statistic under the null distribution and alternatives is studied, and it is also demonstrated that the test is consistent. Furthermore, the proposed approach offers advantages in terms of fast and simple implementation. A comprehensive simulation study shows that the power of the new test compares favorably to that of existing methods. A key application is the assessment of financial time series for the suitability of the Black-Scholes model. Examples relating to various stock and interest rate time series are presented in order to illustrate the proposed test.

## 1. Introduction

The geometric Brownian motion is a stochastic process used in the Black-Scholes model (Black and Scholes, 1973; Merton, 1973). Under this model, the stock price is given by

$$S(t) = S_0 \exp\left(\left(a - \frac{\sigma^2}{2}\right)t + \sigma B(t)\right),\ t \in [0,1],$$

with initial value $S_0 \in (0,\infty)$, drift $a \in \mathbb{R}$, volatility $\sigma \in (0,\infty)$ and with $B(t)$ a standard Brownian motion. In applications, the initial value $S_0$ is typically fixed, while the drift $a$ and volatility $\sigma$ are unknown parameters. Set $\mu = a - \frac{\sigma^2}{2}$, $S_0 = 1$ and take the logarithm, the stochastic process

$$X(t) = \mu t + \sigma B(t),\ t \in [0,1],$$

is obtained. This process is a standard Brownian motion $B(t)$, $t \in [0,1]$, scaled by the volatility $\sigma \in (0,\infty)$ and with linear drift with slope $\mu \in \mathbb{R}$. The exponential (or logarithm) function is bijective (on the positive half-line). As a result, an observation of the stochastic process $X(t)$, $t \in [0,1]$, contains all information about an observation of the stochastic process $S(t)$, $t \in [0,1]$, and vice versa. For this reason, it is referred to the stochastic process $X(t)$, $t \in [0,1]$, as a scaled Brownian motion with linear drift. Methods for testing

* Corresponding author at: FH Aachen - University of Applied Sciences, Faculty 09 - Medical Engineering and Technomathematics, Campus Jülich, Heinrich-Mußmann-Straße 1, 52428 Jülich, Germany.
*E-mail address:* gaigall@fh-aachen.de (D. Gaigall).

https://doi.org/10.1016/j.csda.2025.108196
Received 3 October 2024; Received in revised form 17 April 2025; Accepted 17 April 2025

Available online 23 April 2025
0167-9473/© 2025 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

the assumptions that underlie the Black-Scholes model are of practical importance in finance. One possible approach to test these assumptions in the functional data setting is via goodness-of-fit testing.

Advances in the area of functional data are described in Cuesta-Albertos et al. (2006) and Cuesta-Albertos et al. (2007), where random projections are applied to the data to obtain real-valued observations and finally Kolmogorov-Smirnov type tests are used. Cuevas and Fraiman (2009) extends this procedure to the more general spaces. A similar approach is considered in Cuesta-Albertos and Febrero-Bande (2010) but in the context of an ANOVA test for functional data. Bugni et al. (2009) proposed a bootstrap based goodness-of-fit test for functional data, where the null hypothesis may consist of a parametric family and Bugni and Horowitz (2018) treats an extension to the multi-sample case. Ditzhaus and Gaigall (2018) and Ditzhaus and Gaigall (2022) use a general Hilbert space approach including in particular the functional data case. They apply Cramér-von-Mises type statistics to projected observations, where an integration with respect to a suitable probability measure on the space of projections is used to obtain the test statistics. Gonzãlez-Manteiga (2022) provides a review on goodness-of-fit tests, together with testing in regression models based on distance correlation for functional data. An overview of goodness-of-fit tests for functional data is presented in Gonzãlez-Manteiga et al. (2021).

Henze and Jiménez Gamero (2020) generalized the BHEP-test, with a focus on Gaussian random variables, to the separable Hilbert space case. The classical BHEP test, see Baringhaus and Henze (1988), extends the idea of Epps and Pulley (1983) to the multivariate case and serves as a powerful goodness-of-fit test for multivariate normality. Note that Epps and Pulley (1983) suggest testing normality in the univariate case on the basis of an integral distance between the empirical characteristic function of the data and the characteristic function of the standard normal distribution. The generalization of the classical BHEP test by Henze and Jiménez Gamero (2020) is done by replacing the usual normality in finite dimension with the more general concept of Gaussianity in the infinite dimensional case. The characteristic function is replaced by the characteristic functional. This approach is justified since the characteristic functional uniquely determines the distribution even in the general Hilbert space setting, see Laha and Rohatgi (1979). Coupek et al. (2023) provides an extension in the form of a test for subfamilies of Gaussian distributions. Due to empirical standardization, the distribution of the classical BHEP test statistic in the finite dimensional case is independent of the underlying unknown parameters, i.e., of the mean vector and the covariance matrix, if the null hypothesis of normality holds. For this reason, critical values for the classical BHEP test are readily available. In contrast, in the infinite dimensional case, this is no longer the case. In particular, a counterpart of the standard normal distribution does not exist in general. As a consequence, the tests proposed in Henze and Jiménez Gamero (2020) and Coupek et al. (2023) require computationally expensive resampling techniques to obtain critical values.

Keeping the recent generalization of the BHEP test in mind, a new goodness-of-fit test for geometric Brownian motion (or equivalently for the scaled Brownian motion with linear drift) is proposed in a functional data framework. The test statistic is based on an integral distance between the empirical characteristic functional of the data and the characteristic functional of the standard Brownian motion, which is the counterpart of the standard normal distribution in this case. A BHEP type test is considered for two reasons. First, BHEP type tests have been shown to be powerful, compare Baringhaus and Henze (1988), Henze and Jiménez Gamero (2020) and Coupek et al. (2023). Second, this allows to find a representation of the test statistic useful both for theoretical analysis as well as implementation. The advantages of the approach are summarized as follows: In the particular situation under consideration, an empirical standardization of the data can be used to obtain a test statistic with the useful property that the related distribution is independent of the underlying unknown parameters. That is, if the null hypothesis of a scaled Brownian motion with linear drift holds, then this distribution is independent of the slope of the linear drift and the volatility. For this reason, critical values can readily be computed. This ensures a reduction in the required computational effort, allows quick application of the test in practice and facilitates comprehensive simulation studies. Moreover, it will be demonstrated that the test statistic is simple to implement and that different choices of the integrator in the test statistic enable flexibility when applying the test. Although empirical standardization results in data that are no longer independent, it can be proved that the test has the desired theoretical properties. That is, the test attains the specified nominal significance level in finite sample settings and is consistent as the sample size approaches infinity. In addition, the test exhibits exceptional power, even when small and moderate sample sizes are used.

The remainder of this paper is structured as follows. Section 2 presents the proposed testing procedure. The testing problem is introduced in section 2.1, a maximum likelihood estimation approach for the unknown slope and volatility parameter is proposed in section 2.2 and the test statistic as well as the test is suggested in section 2.3. An alternative representation of the test statistic, useful for theoretical analyses and for implementation purposes, is available. The distribution of the test statistic is independent of the unknown underlying parameters. Furthermore, critical values are readily available and the resulting test maintains the significance level in finite samples.

Section 3 investigates the theoretical properties of the test statistic and the test. In section 3.1, the limit null distribution of the test statistic is established and discussed. Here a fairly general structure of the parameter estimators is assumed. A technical tool of central importance for obtaining the statement is a generalization of the central limit theorem to random variables taking values in a separable Hilbert space. Applying a strong law of large numbers in Hilbert spaces, an almost sure limit result of the test statistic is proved in section 3.2. In this setting, slightly stronger assumptions on the parameter estimators and regularity of the integrator in the test statistic are required. This result holds under the null hypothesis as well as under alternatives and is used to show the test's consistency.

The implementation of the test from a practical perspective is discussed in section 4. A feasible method to obtain estimates of the parameters is provided in section 4.1. In section 4.2, simple formulas for the calculation of the scalar products occurring in

the characteristic functionals of the test statistic are derived. These formulas make it possible to circumvent the use of numerical integration routines and ultimately reduce the required computational effort. A Monte-Carlo approach for the calculation of the integral distance in the test statistic is suggested, see section 4.3.

In Section 5 the results of simulation studies as well as real data examples are presented. The performance of the test, with a special focus on the power under alternatives, is investigated. Gaussian and non-Gaussian alternatives are considered in settings obtained from literature (section 5.1) and in new settings (section 5.2). The empirical findings indicate that the test has power, including in small and moderate samples. In most of the situations under consideration, the test compares favorably to other procedures or provides comparable results. The application of the test is demonstrated on the basis of stock market and interest rate time series data in section 5.3. For the stock prices under consideration, the calculated $p$-values are in line with the model assumption of a geometric Brownian motion. In contrast, when turning the attention to the observed interest rate data, the calculated $p$-values indicate that the geometric Brownian motion is not a suitable model.

## 2. The testing procedure

In this section, the testing problem is introduced, estimators for the unknown slope and volatility parameters are proposed and justified, the test statistic and the test are suggested and different choices of the integrator in the test statistic are discussed.

### 2.1. The testing problem

Let $\mathrm{L}^2$ be the Hilbert space of equivalence classes of $(\mathfrak{B}([0,1]),\mathfrak{B})$-measurable functions $f:[0,1]\to\mathbb{R}$ with $\int_0^1 f(x)^2\mathrm{d}x<\infty$. Here, $\mathfrak{B}$ or $\mathfrak{B}([0,1])$ denotes the Borel $\sigma$-field on $\mathbb{R}$ or $[0,1]$. The Hilbert space $\mathrm{L}^2$ is separable and can be equipped with the scalar product $\langle f,g\rangle=\int_0^1 f(x)g(x)\mathrm{d}x$ and with the norm $||f||=\sqrt{\langle f,f\rangle}$ for $f,g\in\mathrm{L}^2$. The norm induces a distance and the concept of open sets applies. $\mathrm{L}^2$ can be equipped with the Borel $\sigma$-field $\mathfrak{B}(\mathrm{L}^2)$ that is generated by all open sets. Let $(\Omega,\mathfrak{A},P)$ be a probability space and let $X:\Omega\to\mathrm{L}^2$ be a $(\mathfrak{A},\mathfrak{B}(\mathrm{L}^2))$-measurable random variable with distribution $F$; write $X\sim F$. It is supposed that $E(||X||^2)<\infty$. Then the random variable $X$ has a mean function $\nu\in\mathrm{L}^2$ given by

$$\nu(t)=E\big(X(t)\big),\ t\in[0,1],$$

and a covariance operator $C$ given by

$$Cf(s)=\int_0^1 f(t)c(s,t)\mathrm{d}t,\ f\in\mathrm{L}^2,\ s\in[0,1],$$

where $c(s,t)=\mathrm{cov}(X(s),X(t))$, $s,t\in[0,1]$, is the covariance function. The distribution of the Hilbert space valued random variable $X$ is uniquely determined by its characteristic functional

$$\phi_X(f)=E\big(\exp(i\langle X,f\rangle)\big),\ f\in\mathrm{L}^2,$$

see Proposition 7.6.3 in Laha and Rohatgi (1979). The random variable $X$ is called Gaussian, if its characteristic functional is of the form

$$\phi_X(f)=\phi_{\nu,C}(f)=\exp\left(i\langle f,\nu\rangle-\frac{1}{2}\langle Cf,f\rangle\right),\ f\in\mathrm{L}^2.$$

If the stochastic process is a scaled Brownian motion with linear drift, where $\mu\in\mathbb{R}$ is the slope and $\sigma\in(0,\infty)$ is the volatility, write $X\sim\mathrm{BM}(\mu,\sigma)$; in this case, the mean function $\nu=\nu_\mu$ is given by

$$\nu_\mu(t)=\mu t,\ t\in[0,1],$$

the covariance function $c=c_\sigma$ has the expression

$$c_\sigma(s,t)=\sigma\min(s,t),\ s,t\in[0,1],$$

and the covariance operator $C=C_\sigma$ satisfies

$$C_\sigma f(s)=\int_0^1 f(t)c_\sigma(s,t)\mathrm{d}t,\ f\in\mathrm{L}^2,\ s\in[0,1].$$

Consider the null hypothesis that a scaled Brownian motion with linear drift applies and the general alternative that this is not the case. That is, the testing problem is

$$H_0:F\in\mathcal{F},\ H_1:F\notin\mathcal{F},$$

where $\mathcal{F} = \{\mathrm{BM}(\mu,\sigma); \mu \in \mathbb{R}, \sigma \in (0,\infty)\}$ denotes the family of distributions of scaled Brownian motions with linear drift. Given that $n \in \mathbb{N}$ is the sample size, the testing problem is treated on the basis of the data $X_1,\dots,X_n$, which are copies of $X$. In particular, $X, X_1,\dots,X_n$ are independent and identically distributed random variables on $(\Omega,\mathfrak{A},P)$.

### 2.2. Parameter estimation

The testing approach requires the estimation of the unknown parameters $\mu \in \mathbb{R}$ and $\sigma \in (0,\infty)$ under the null hypothesis of a scaled Brownian motion with linear drift. For this purpose, let $\hat{\mu}_n$ and $\hat{\sigma}_n$ be estimators, that are real-valued random variables on $(\Omega,\mathfrak{A},P)$. Although the usage of other estimators is possible and covered by the theoretical results stated later, a special focus is on a maximum likelihood estimation approach. The estimation of the slope $\mu$ is motivated by Subsection 2.1 in Mishura et al. (2017) and the idea can also be adopted for estimating the volatility $\sigma$. Theoretically, the volatility can be directly obtained from a single observation of a scaled Brownian motion with linear drift, precisely from the so-called quadratic variation. For details on quadratic variation, see Revuz and Yor (2010). A method is needed that provides stable values for the estimator of $\sigma$ simultaneously under alternatives on the one hand and that is applicable in practice on the other hand. The following estimation approach takes both considerations into account.

The derivation of the estimators requests some $N \in \mathbb{N}$ and a partition $0 = t_0 < \dots < t_N = 1$ of $[0,1]$. Note that the final estimators do not depend on the choice of $N$ or the partition. However, the following derivation gives a hint as to how to obtain estimates in practice, where the complete functions are typically not available in the observed data; details are given later. Evaluating the data along the partition results in the increments $\Delta X_{i,j} = X_j(t_i) - X_j(t_{i-1})$, $i = 1,\dots,N$, $j = 1,\dots,n$, which are in all $nN$ real-valued random variables. If the null hypothesis of a scaled Brownian motion with linear drift applies, these random variables are independent with distribution $\Delta X_{i,j} \sim N((t_i - t_{i-1})\mu, (t_i - t_{i-1})\sigma^2)$, $i = 1,\dots,N$, $j = 1,\dots,n$. For that reason, the maximum likelihood estimator of $\mu$, on the basis of $\Delta X_{i,j}$, $i = 1,\dots,N$, $j = 1,\dots,n$, is

$$\hat{\mu}_{n,N} = \frac{1}{n}\sum_{j=1}^{n}\sum_{i=1}^{N}\Delta X_{i,j}.$$

Using $X_j(0) = 0$ and the telescopic sum $\sum_{i=1}^{N}\Delta X_{i,j} = X_j(1)$, $j = 1,\dots,n$, for scaled Brownian motions with linear drift, it is

$$\hat{\mu}_{n,N} = \hat{\mu}_n = \frac{1}{n}\sum_{j=1}^{n} X_j(1), \tag{1}$$

independent of $N$. The strong law of large numbers implies the almost sure convergence $\hat{\mu}_n \xrightarrow{a.s.} E(X(1))$ as $n \to \infty$ under the null hypothesis as well as under alternatives. Furthermore, $E(X(1)) = \mu$ if the null hypothesis of a scaled Brownian motion with linear drift applies. This motivates the use of $\hat{\mu}_n$ as an estimator of $\mu$.

Similarly, the maximum likelihood estimator of $\sigma$ on the basis of $\Delta X_{i,j}$, $i = 1,\dots,N$, $j = 1,\dots,n$, in the situation above, is given by

$$\hat{\sigma}_{n,N} = \sqrt{\frac{1}{nN}\sum_{j=1}^{n}\sum_{i=1}^{N}\frac{1}{t_i - t_{i-1}}\left(\Delta X_{i,j} - (t_i - t_{i-1})\hat{\mu}_n\right)^2}. \tag{2}$$

Supposing that the stochastic process $X(t)$, $t \in [0,1]$, admits the quadratic variation at least at $t = 1$ and that the partition satisfies $\lim_{N\to\infty}\max_{i=1,\dots,N}(t_i - t_{i-1}) = 0$, it holds the convergence in probability $\sum_{i=1}^{N}(\Delta X_{i,j})^2 \xrightarrow{P} [X_j](1)$ as $N \to \infty$, where the real-valued random variable $[X_j](1)$ denotes the quadratic variation of $X_j$ at $t = 1$. Considering in particular the equidistant partition, i.e., $\Delta_N = t_i - t_{i-1} = \frac{1}{N}$, $i = 1,\dots,N$, the expression

$$\hat{\sigma}_{n,N} = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(\sum_{i=1}^{N}(\Delta X_{i,j})^2 - 2\frac{1}{N}\hat{\mu}_n X_j(1) + \frac{1}{N}\hat{\mu}_n^2\right)}$$

shows that

$$\hat{\sigma}_{n,N} \xrightarrow{P} \hat{\sigma}_n = \sqrt{\frac{1}{n}\sum_{j=1}^{n}[X_j](1)} \text{ as } N \to \infty.$$

It follows from the strong law of large numbers that $\hat{\sigma}_n \xrightarrow{a.s.} \sqrt{E([X](1))}$ as $n \to \infty$ under the null hypothesis as well as under alternatives if $E([X](1)) < \infty$. That is, if the expectation of the quadratic variation at $t = 1$ is finite, and from the fact that the standard Brownian motion has the quadratic variation equal to 1 at $t = 1$ it follows that $[X](1) = \sigma^2$ and so $\hat{\sigma}_n = \sigma$ if the null hypothesis of a scaled Brownian motion with linear drift applies. This motivates the use of $\hat{\sigma}_n$ as an estimator of $\sigma$.

### 2.3. Test statistic and test

The test statistic uses estimators $\hat{\mu}_n$ and $\hat{\sigma}_n$ of $\mu \in \mathbb{R}$ and $\sigma \in (0, \infty)$ if the null hypothesis of a scaled Brownian motion with linear drift applies, e.g., the estimators obtained by the maximum likelihood approach in section 2.2. Given that $\hat{\sigma}_n > 0$ is satisfied, empirically standardized data $Y_1, \ldots, Y_n$ are obtained by

$$Y_j(t) = \frac{X_j(t) - \hat{\mu}_n t}{\hat{\sigma}_n}, \ t \in [0,1], \ j = 1, \ldots, n.$$

Note that these random variables each have the same distribution but they are no longer independent. If the estimators obtained by the maximum likelihood approach in section 2.2 are used, their (joint) distribution does not depend on the true underlying parameters $\mu$ and $\sigma$ if the null hypothesis $H_0$ holds. The related empirical characteristic functional is

$$\phi_n(f) = \frac{1}{n} \sum_{j=1}^{n} \exp\left(i \left\langle f, Y_j \right\rangle\right), \ f \in \mathrm{L}^2.$$

The characteristic functional of a standard Brownian motion is

$$\phi_{0,C_1}(f) = \exp\left( -\frac{1}{2} \langle C_1 f, f \rangle \right), \ f \in \mathrm{L}^2.$$

Note that the map $(\omega, f) \mapsto |\phi_n(f)(\omega) - \phi_{0,C_1}(f)|^2$, $(\omega, f) \in \Omega \times \mathrm{L}^2$, is $(\mathfrak{A} \otimes \mathfrak{B}(\mathrm{L}^2), \mathfrak{B})$-measurable and moreover bounded by 4. A BHEP type test statistic similar as in Henze and Jiménez Gamero (2020) is given by

$$T_n = n \int_{\mathrm{L}^2} |\phi_n(f) - \phi_{0,C_1}(f)|^2 Q(\mathrm{d}f),$$

where $Q$ is a probability measure defined on the measurable space $(\mathrm{L}^2, \mathfrak{B}(\mathrm{L}^2))$. In applications, different choices of $Q$ are possible; details are given later. From Fubini's theorem, $T_n$ is a real-valued random variable.

Large values of the test statistic indicate that the null hypothesis is to be rejected in favor of the alternative $H_1$. Given that $\alpha \in (0,1)$ is the significance level, a quantile of order $1-\alpha$ of the distribution of $T_n$ under the null hypothesis $H_0$ is denoted by $c_{n,1-\alpha}$. The test that rejects the null hypothesis $H_0$ if and only if $T_n > c_{n,1-\alpha}$ maintains the significance level, i.e., it is

$$P(T_n > c_{n,1-\alpha}) \leq \alpha$$

if the null hypothesis $H_0$ is valid. If the estimators obtained by the maximum likelihood approach in section 2.2 are used, the distribution of $T_n$ does not depend on the true underlying parameters $\mu$ and $\sigma$ if the null hypothesis $H_0$ holds. This is due to the fact that $T_n$ depends only on the empirically standardized data $Y_1, \ldots, Y_n$. In this case, the quantiles $c_{n,1-\alpha}$ are readily available as critical values for the test in applications.

Now, consider the case where $Q$ is symmetric in the sense that it is invariant under the map $f \mapsto -f$, $f \in \mathrm{L}^2$. Subsequently, suitable choices of $Q$ are presented such that this property is satisfied. Then, analogously to Henze and Jiménez Gamero (2020), an alternative expression of the test statistic is obtained by

$$T_n = \int_{\mathrm{L}^2} V_n^2(f) Q(\mathrm{d}f), \tag{3}$$

where

$$V_n(f) = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \left( \cos\left\langle f, Y_j \right\rangle + \sin\left\langle f, Y_j \right\rangle - \exp\left( -\frac{1}{2} \langle C_1 f, f \rangle \right) \right), \ f \in \mathrm{L}^2. \tag{4}$$

This formula is useful for theoretical analysis and for implementation purposes.

### 2.4. The choice of the integrator in the test statistic

Two suitable specifications of $Q$ are presented. The first specification is that $Q$ is the Wiener measure associated with the standard Brownian motion analogously to Henze and Jiménez Gamero (2020). Another possible choice of $Q$ is associated with random linear combinations of the Legendre polynomials $P_j$, $j \in \mathbb{N}$, on $[0,1]$ similar to the approach developed in Ditzhaus and Gaigall (2022) and Ditzhaus and Gaigall (2018). Both of these measures have been shown to provide powerful tests in the respective literature. Note that the Legendre polynomials can be used to build an orthogonal basis on the separable Hilbert space $\mathrm{L}^2$. For this purpose, choose a discrete distribution $\zeta$ on $(\mathbb{N}, \mathfrak{P}(\mathbb{N}))$ with a finite fourth moment, another discrete distribution $\eta$ on $(\mathbb{N}, \mathfrak{P}(\mathbb{N}))$ and a distribution $\xi$ on $(\mathbb{R}, \mathfrak{B})$ with a finite fourth moment that is symmetric in the sense that $\xi$ is invariant under the map $x \mapsto -x$, $x \in \mathbb{R}$. Here, $\mathfrak{P}(\mathbb{N})$ denotes the power set of $\mathbb{N}$. For example, choose with $\zeta$ or $\eta$ a Poisson distribution $\mathrm{Poi}(\lambda) + 1$ for some expectation $\lambda \in (0, \infty)$ and shifted by one and with $\xi$ the normal distribution $N(0, \tau^2)$ for some variance $\tau^2 \in (0, \infty)$. The procedure to generate a realization $f$ of $Q$ is given as follows:

1. Generate a realization $k \in \mathbb{N}$ of the distribution $\zeta$.
2. Independently of step 1., generate $i_1, \dots, i_k \in \mathbb{N}$ by sampling without replacement according to the distribution $\eta$.
3. Independently of steps 1. and 2., generate $a_1, \dots, a_k \in \mathbb{R}$ as independent realizations of the distribution $\xi$.
4. Set $f = \sum_{j=1}^{k} a_j P_{i_j}$.

## 3. Theoretical properties

Below, the limit distribution of the test statistic under the null hypothesis is established and discussed, an almost sure limit result of the test statistic is proved and the test's consistency is deduced.

### 3.1. The limit null distribution

Considering general estimators $\hat{\mu}_n$ and $\hat{\sigma}_n$ of $\mu \in \mathbb{R}$ and $\sigma \in (0, \infty)$ in the case that the null hypothesis of a scaled Brownian motion with linear drift applies. The following assumption is required to obtain a distributional limit of the test statistic under the null hypothesis. It is easy to see that this assumption is satisfied for the estimators obtained by the maximum likelihood approach in section 2.2.

**Assumption 1.** If the null hypothesis $H_0$ that a scaled Brownian motion with linear drift with slope $\mu \in \mathbb{R}$ and volatility parameter $\sigma \in (0, \infty)$ applies there exists a $(\mathfrak{B}(\mathrm{L}^2), \mathfrak{B}^2)$-measurable map $l : \mathrm{L}^2 \to \mathbb{R}^2$ such that

$$\sqrt{n}\left((\hat{\mu}_n, \hat{\sigma}_n) - (\mu, \sigma)\right)^\top = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} l(X_i) + o_P(1) \text{ as } n \to \infty,$$

and $E(|l(X)|^2) < \infty$ as well as $E(l(X)) = 0$ is satisfied. Here, $|\cdot|$ is the usual norm in $\mathbb{R}^2$. Write $l = (l_\mu, l_\sigma)^\top$ and denote by $\Sigma_l = E(l(X)l(X)^\top)$ the covariance matrix of $l(X)$.

Moreover, some regularity on the probability measure $Q$ is needed, analogously to the situation considered in Henze and Jiménez Gamero (2020). For the two specifications of $Q$ presented in section 2.4, the stated properties are satisfied. In the case where $Q$ is the Wiener measure, this is clear. In the case that $Q$ is associated with the random linear combinations of the Legendre polynomials, this follows from Wald's formula and Jensen's inequality.

**Assumption 2.** The probability measure $Q$ satisfies $\int_{\mathrm{L}^2} ||f||^4 Q(\mathrm{d}f) < \infty$ and is symmetric in the sense that it is invariant under the map $f \mapsto -f$, $f \in \mathrm{L}^2$.

Similar as in Henze and Jiménez Gamero (2020), the limit null distribution is derived from a general version of the central limit theorem in separable Hilbert spaces, see Theorem 7.5.1 in Laha and Rohatgi (1979).

Denote by $\mathrm{L}^2_Q$ the Hilbert space of equivalence classes of $(\mathfrak{B}(\mathrm{L}^2), \mathfrak{B})$-measurable functions $S : \mathrm{L}^2 \to \mathbb{R}$ with $\int_{\mathrm{L}^2} S(f)^2 Q(\mathrm{d}f) < \infty$. The scalar product is denoted by $\langle S, T \rangle_Q = \int_{\mathrm{L}^2} S(f)T(f)Q(\mathrm{d}f)$ and the norm by $||S||_Q = \sqrt{\langle S, S \rangle_Q}$ for $S, T \in \mathrm{L}^2_Q$. Because $\mathrm{L}^2$ is separable $\mathrm{L}^2_Q$, is also separable. This is an easy consequence of Proposition 3.4.5 in Cohn (2013). In what follows, write $\sigma_f^2 = \int_0^1 \int_0^1 \min(x, y) f(x) f(y) \mathrm{d}y \, \mathrm{d}x$, $\sigma_{f,g}^2 = \int_0^1 \int_0^1 \min(x, y) f(y) g(x) \mathrm{d}y \, \mathrm{d}x$, $f, g \in \mathrm{L}^2$ and $\mathrm{id}_{\mathbb{R}} : \mathbb{R} \to \mathbb{R}$ for the identity map for $f, g \in \mathrm{L}^2$. Moreover, remember that under Assumption 2 there is an alternative expression of the test statistic $T_n$ given in eq. (3) and $V_n$ defined in eq. (4). As a preparation, the following result is stated.

**Proposition 1.** *Suppose that Assumption 1 and Assumption 2 hold. In addition, assume the null hypothesis $H_0$ that a scaled Brownian motion with linear drift applies with slope $\mu \in \mathbb{R}$ and volatility parameter $\sigma \in (0, \infty)$. Defining the $(\mathfrak{B}(\mathrm{L}^2) \otimes \mathfrak{B}(\mathrm{L}^2), \mathfrak{B})$-measurable map $\Psi : \mathrm{L}^2 \times \mathrm{L}^2 \to \mathbb{R}$ by*

$$\Psi(f, x) = \cos\left\langle f, \frac{x - \mu \cdot \mathrm{id}_{\mathbb{R}}}{\sigma} \right\rangle + \sin\left\langle f, \frac{x - \mu \cdot \mathrm{id}_{\mathbb{R}}}{\sigma} \right\rangle - \exp(-\frac{1}{2}\sigma_f^2) + \exp\left(-\frac{\sigma_f^2}{2}\right)\left(\left(\frac{C_1 f(1)}{\sigma}, \frac{\sigma_f^2}{\sigma}\right) l(x)\right),$$

*$f, x \in \mathrm{L}^2$, it is $E(\Psi(f, X)) = 0$ for all $f \in \mathrm{L}^2$ as well as $E(||\Psi(\cdot, X)||_Q^2) = \int_{\mathrm{L}^2} E(\Psi(f, X)^2) Q(\mathrm{d}f) < \infty$. Moreover, setting*

$$V_{n,0}(f) = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \Psi(f, X_j), \ f \in \mathrm{L}^2,$$

*it follows that*

$$||V_n - V_{n,0}||_Q = o_P(1) \text{ as } n \to \infty.$$

**Proof.** Set

$$Y_j^*(t) = \frac{X_j(t) - \mu t}{\sigma},\ t \in [0,1],\ j = 1, \dots, n,$$

and note that $Y_j^*$, $j = 1, \dots, n$ are now independent, each with the same distribution, namely with that of $Y^*(t) = \frac{X(t)-\mu t}{\sigma}$, $t \in [0,1]$, that is a standard Brownian motion. Note that Assumption 1 and Assumption 2 ensure that $E(\cos\langle f, Y^*\rangle) = \exp(-\frac{\sigma_f^2}{2})$, $E(\sin\langle f, Y^*\rangle) = 0$, $E(l(X)) = 0$ and so $E(\Psi(f, X)) = 0$ for all $f \in \mathrm{L}^2$ as well as $\int_{\mathrm{L}^2} E(\Psi(f,X)^2) Q(\mathrm{d}f) < \infty$.

Now, elementary calculations show that the empirically standardized random variables can be rewritten as

$$Y_j(t) = Y_j^*(t) + \left( \frac{\sigma - \hat{\sigma}_n}{\hat{\sigma}_n} Y_j^*(t) + \frac{(\mu - \hat{\mu}_n)t}{\hat{\sigma}_n} \right),\ t \in [0,1],\ j = 1, \dots, n. \tag{5}$$

Applying for each $j = 1, \dots, n$ and each $f \in \mathrm{L}^2$ a Taylor expansion of the functions cos and sin about $\langle f, Y_j^* \rangle$, there are (random) numbers $\xi_{1,j}(f)$ and $\xi_{2,j}(f)$ between $\langle f, Y_j \rangle$ and $\langle f, Y_j^* \rangle$ such that

$$\begin{aligned}
\cos\left\langle f, Y_j \right\rangle &= \cos\left\langle f, Y_j^* + \left( \frac{\hat{\sigma}_n - \sigma}{\hat{\sigma}_n} Y_j^* + \frac{(\hat{\mu}_n - \mu)\mathrm{id}_{\mathbb{R}}}{\hat{\sigma}_n} \right) \right\rangle \\
&= \cos\left\langle f, Y_j^* \right\rangle - \left\langle f, \frac{\hat{\sigma}_n - \sigma}{\hat{\sigma}_n} Y_j^* + \frac{(\hat{\mu}_n - \mu)\mathrm{id}_{\mathbb{R}}}{\hat{\sigma}_n} \right\rangle \sin\left\langle f, Y_j^* \right\rangle \\
&\quad + \left\langle f, \frac{\hat{\sigma}_n - \sigma}{\hat{\sigma}_n} Y_j^* + \frac{(\hat{\mu}_n - \mu)\mathrm{id}_{\mathbb{R}}}{\hat{\sigma}_n} \right\rangle^2 \frac{1}{2} \cos(\xi_{1,j}(f))
\end{aligned}$$

and

$$\begin{aligned}
\sin\left\langle f, Y_j \right\rangle &= \sin\left\langle f, Y_j^* + \left( \frac{\hat{\sigma}_n - \sigma}{\hat{\sigma}_n} Y_j^* + \frac{(\hat{\mu}_n - \mu)\mathrm{id}_{\mathbb{R}}}{\hat{\sigma}_n} \right) \right\rangle \\
&= \sin\left\langle f, Y_j^* \right\rangle + \left\langle f, \frac{\hat{\sigma}_n - \sigma}{\hat{\sigma}_n} Y_j^* + \frac{(\hat{\mu}_n - \mu)\mathrm{id}_{\mathbb{R}}}{\hat{\sigma}_n} \right\rangle \cos\left\langle f, Y_j^* \right\rangle \\
&\quad - \left\langle f, \frac{\hat{\sigma}_n - \sigma}{\hat{\sigma}_n} Y_j^* + \frac{(\hat{\mu}_n - \mu)\mathrm{id}_{\mathbb{R}}}{\hat{\sigma}_n} \right\rangle^2 \frac{1}{2} \sin(\xi_{2,j}(f)),
\end{aligned}$$

noticing that the maps $(\omega, f) \mapsto \cos(\xi_{1,j}(f))(\omega)$ and $(\omega, f) \mapsto \sin(\xi_{2,j}(f))(\omega)$, $(\omega, f) \in \Omega \times \mathrm{L}^2$, are $(\mathfrak{A} \otimes \mathfrak{B}(\mathrm{L}^2), \mathfrak{B})$-measurable for each $j = 1, \dots, n$. By substituting the terms into eq. (4), it is

$$\begin{aligned}
V_n(f) &= \frac{1}{\sqrt{n}} \sum_{j=1}^n \left( \cos\left\langle f, Y_j^* \right\rangle + \sin\left\langle f, Y_j^* \right\rangle - \exp\left(-\frac{1}{2}\sigma_f^2\right) \right) \\
&\quad + \frac{\hat{\sigma}_n - \sigma}{\hat{\sigma}_n} \frac{1}{\sqrt{n}} \sum_{j=1}^n \left\langle f, Y_j^* \right\rangle \left( \cos\left\langle f, Y_j^* \right\rangle - \sin\left\langle f, Y_j^* \right\rangle \right) \qquad (6) \\
&\quad + \frac{\hat{\mu}_n - \mu}{\hat{\sigma}_n} \frac{1}{\sqrt{n}} \sum_{j=1}^n \langle f, \mathrm{id}_{\mathbb{R}} \rangle \left( \cos\left\langle f, Y_j^* \right\rangle - \sin\left\langle f, Y_j^* \right\rangle \right) \qquad (7) \\
&\quad + R_{1,n}(f),\ f \in \mathrm{L}^2,
\end{aligned}$$

with the remainder term

$$R_{1,n}(f) = \frac{1}{\sqrt{n}} \sum_{j=1}^n \frac{1}{2} \left( \cos(\xi_{1,j}(f)) - \sin(\xi_{2,j}(f)) \right) \left\langle f, \frac{\hat{\sigma}_n - \sigma}{\hat{\sigma}_n} Y_j^* + \frac{(\hat{\mu}_n - \mu)\mathrm{id}_{\mathbb{R}}}{\hat{\sigma}_n} \right\rangle^2,\ f \in \mathrm{L}^2.$$

Applying triangle inequality, the boundedness of the functions sin and cos as well as the Cauchy-Schwarz inequality, the remainder term is

$$\begin{aligned}
||R_{1,n}||_Q &\leq \frac{\sqrt{n}(\hat{\sigma}_n - \sigma)^2}{\hat{\sigma}_n^2} \left( \int_{\mathrm{L}^2} \|f\|^4 Q(\mathrm{d}f) \right)^{\frac{1}{2}} \frac{1}{n} \sum_{j=1}^n \|Y_j^*\|^2 \\
&\quad + \frac{2\sqrt{n}|\hat{\sigma}_n - \sigma||\hat{\mu}_n - \mu|}{\hat{\sigma}_n^2} \left( \int_{\mathrm{L}^2} \|f\|^4 Q(\mathrm{d}f) \right)^{\frac{1}{2}} \|\mathrm{id}_{\mathbb{R}}\| \frac{1}{n} \sum_{j=1}^n \|Y_j^*\| + \frac{\sqrt{n}(\hat{\mu}_n - \mu)^2}{\hat{\sigma}_n^2} \left( \int_{\mathrm{L}^2} \|f\|^4 Q(\mathrm{d}f) \right)^{\frac{1}{2}} \|\mathrm{id}_{\mathbb{R}}\|^2.
\end{aligned}$$

Using Assumption 1 it follows that $\hat{\sigma}_n \xrightarrow{P} \sigma$ and $\hat{\mu}_n \xrightarrow{P} \mu$ from the law of large numbers combined with Slutsky's theorem and the central limit theorem combined with Slutsky's theorem gives the convergences in distribution of $\sqrt{n}(\hat{\sigma}_n - \sigma)$ as well as of $\sqrt{n}(\hat{\mu}_n - \mu)$

as $n \to \infty$. Moreover, the law of large numbers implies the convergence of $\frac{1}{n}\sum_{j=1}^{n}\|Y_j^*\|$ and $\frac{1}{n}\sum_{j=1}^{n}\|Y_j^*\|^2$ in probability to real numbers as $n \to \infty$. In addition, Assumption 2 ensures that $(\int_{\mathrm{L}^2}\|f\|^4 Q(\mathrm{d}f))^{\frac{1}{2}} < \infty$. Combining these findings with Slutsky's theorem to obtain $||R_{1,n}||_Q \xrightarrow{P} 0$ as $n \to \infty$. Turning the attention to the term in eq. (6) and using Assumption 1, it is

$$\frac{\hat{\sigma}_n - \sigma}{\hat{\sigma}_n}\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\left\langle f, Y_j^*\right\rangle\left(\cos\left\langle f, Y_j^*\right\rangle - \sin\left\langle f, Y_j^*\right\rangle\right) = -\frac{1}{\sigma}\sigma_f^2\exp(-\frac{\sigma_f^2}{2})\frac{1}{\sqrt{n}}\sum_{j=1}^{n} l_\sigma(X_j) + R_{2,n}(f),\ f \in \mathrm{L}^2,$$

with the remainder term

$$R_{2,n}(f) = \frac{1}{\hat{\sigma}_n}\frac{1}{n}\sum_{j=1}^{n}\left\langle f, Y_j^*\right\rangle\left(\cos\left\langle f, Y_j^*\right\rangle - \sin\left\langle f, Y_j^*\right\rangle\right)\sqrt{n}(\hat{\sigma}_n - \sigma) + \frac{1}{\sigma}\sigma_f^2\exp(-\frac{\sigma_f^2}{2})\frac{1}{\sqrt{n}}\sum_{j=1}^{n} l_\sigma(X_j),\ f \in \mathrm{L}^2.$$

From Assumption 1 it is $\hat{\sigma}_n - \sigma \xrightarrow{P} 0$ as a consequence of the law of large numbers combined with Slutsky's theorem as well as $\sqrt{n}(\hat{\sigma}_n - \sigma) - \frac{1}{\sqrt{n}}\sum_{j=1}^{n} l_\sigma(X_j) \xrightarrow{P} 0$ as $n \to \infty$. In addition, from a general version of the law of large numbers in Hilbert spaces, see Theorem 7.2.1 in Laha and Rohatgi (1979), it follows that $\int_{\mathrm{L}^2}(\frac{1}{n}\sum_{j=1}^{n}\langle f, Y_j^*\rangle(\cos\langle f, Y_j^*\rangle - \sin\langle f, Y_j^*\rangle) + \sigma_f^2\exp(-\frac{\sigma_f^2}{2}))^2 Q(\mathrm{d}f) \xrightarrow{P} 0$ as $n \to \infty$, noticing that $E(\langle f, Y^*\rangle(\cos\langle f, Y^*\rangle - \sin\langle f, Y^*\rangle)) = -\sigma_f^2\exp(-\frac{\sigma_f^2}{2})$, $f \in \mathrm{L}^2$. Combining these results with Slutsky's theorem, a basic calculation shows that $||R_{2,n}||_Q \xrightarrow{P} 0$ as $n \to \infty$. Treating the term in eq. (7) analogously it is obtained that

$$\frac{\hat{\mu}_n - \mu}{\hat{\sigma}_n}\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\langle f, \mathrm{id}_{\mathbb{R}}\rangle\left(\cos\left\langle f, Y_j^*\right\rangle - \sin\left\langle f, Y_j^*\right\rangle\right) = \frac{1}{\sigma}C_1 f(1)\exp(-\frac{\sigma_f^2}{2})\frac{1}{\sqrt{n}}\sum_{i=1}^{n} l_\mu(X_i) + R_{3,n}(f),\ f \in \mathrm{L}^2,$$

with the remainder term

$$R_{3,n}(f) = \frac{1}{\hat{\sigma}_n}\frac{1}{n}\sum_{j=1}^{n}\langle f, \mathrm{id}_{\mathbb{R}}\rangle\left(\cos\left\langle f, Y_j^*\right\rangle - \sin\left\langle f, Y_j^*\right\rangle\right)\sqrt{n}(\hat{\mu}_n - \mu) - \frac{1}{\sigma}C_1 f(1)\exp(-\frac{\sigma_f^2}{2})\frac{1}{\sqrt{n}}\sum_{i=1}^{n} l_\mu(X_i),\ f \in \mathrm{L}^2.$$

From Assumption 1 it follows that $\hat{\mu}_n - \mu \xrightarrow{P} 0$ as a consequence of the law of large numbers combined with Slutsky's theorem as well as $\sqrt{n}(\hat{\mu}_n - \mu) - \frac{1}{\sqrt{n}}\sum_{j=1}^{n} l_\mu(X_j) \xrightarrow{P} 0$ as $n \to \infty$. Furthermore, again with the law of large numbers in Hilbert spaces, see Theorem 7.2.1 in Laha and Rohatgi (1979), it follows that $\int_{\mathrm{L}^2}(\frac{1}{n}\sum_{j=1}^{n}\langle f, \mathrm{id}_{\mathbb{R}}\rangle(\cos\langle f, Y_j^*\rangle - \sin\langle f, Y_j^*\rangle) - C_1 f(1)\exp(-\frac{\sigma_f^2}{2}))^2 Q(\mathrm{d}f) \xrightarrow{P} 0$ as $n \to \infty$, noticing that $E(\langle f, \mathrm{id}_{\mathbb{R}}\rangle(\cos\langle f, Y^*\rangle - \sin\langle f, Y^*\rangle)) = C_1 f(1)\exp(-\frac{\sigma_f^2}{2})$, $f \in \mathrm{L}^2$. Combining these finding with Slutsky's theorem, a basic calculation serves $||R_{3,n}||_Q \xrightarrow{P} 0$ as $n \to \infty$. In fact, it is $V_n - V_{n,0} = R_{1,n} + R_{2,n} + R_{3,n}$ and the findings for the remainder terms result in

$$||V_n - V_{n,0}||_Q = ||R_{1,n} + R_{2,n} + R_{3,n}||_Q \le ||R_{1,n}||_Q + ||R_{2,n}||_Q + ||R_{3,n}||_Q \xrightarrow{P} 0 \text{ as } n \to \infty$$

and so $||V_n - V_{n,0}||_Q \xrightarrow{P} 0$. This completes the proof. □

Now, a result about the asymptotic null distribution of the test is stated.

**Theorem 1.** *Suppose that Assumption 1 and Assumption 2 hold. In addition, assume the null hypothesis $H_0$ that a scaled Brownian motion with linear drift applies with slope $\mu \in \mathbb{R}$ and volatility parameter $\sigma \in (0, \infty)$. Then the convergence in distribution*

$$T_n \xrightarrow{D} ||\mathcal{V}||_Q^2 = T_\infty \text{ as } n \to \infty$$

*holds, where $\mathcal{V}$ is a $L_Q^2$-valued and centered Gaussian random variable with covariance operator given by the covariance function*

$$\begin{aligned} E\big(\mathcal{V}(f)\mathcal{V}(g)\big) = \exp\left(-\frac{1}{2}(\sigma_f^2 + \sigma_g^2)\right)\Bigg(\exp(\sigma_{f,g}^2) + \left(\frac{C_1 f(1)}{\sigma}, \frac{\sigma_f^2}{\sigma}\right)\Sigma_l\left(\frac{C_1 g(1)}{\sigma}, \frac{\sigma_g^2}{\sigma}\right)^{\top} - 1\Bigg) \\ + \sqrt{2}E\left(\left(\frac{C_1 f(1)}{\sigma}, \frac{\sigma_f^2}{\sigma}\right) l(X)\left(\exp\left(-\frac{1}{2}\sigma_f^2\right)\sin\left(\left\langle g, \frac{X - \mu\cdot\mathrm{id}_{\mathbb{R}}}{\sigma}\right\rangle + \frac{\pi}{4}\right)\right)\right) \\ + \sqrt{2}E\left(\left(\frac{C_1 g(1)}{\sigma}, \frac{\sigma_g^2}{\sigma}\right) l(X)\left(\exp\left(-\frac{1}{2}\sigma_g^2\right)\sin\left(\left\langle f, \frac{X - \mu\cdot\mathrm{id}_{\mathbb{R}}}{\sigma}\right\rangle + \frac{\pi}{4}\right)\right)\right),\ f, g \in \mathrm{L}^2, \end{aligned}$$

*and $T_\infty$ is a real-valued random variable.*

**Proof.** As direct consequence of Proposition 1, a general version of the central limit theorem in separable Hilbert spaces, see Theorem 7.5.1 in Laha and Rohatgi (1979), applies and it follows that

$$V_{n,0} \xrightarrow{D} \mathcal{V} \text{ as } n \to \infty,$$

where the convergence in distribution takes place in the Hilbert space $L^2_Q$, compare Laha and Rohatgi (1979), and $\mathcal{V}$ is a $L^2_Q$-valued stochastic process, in particular a centered Gaussian random variable. The related covariance operator or the related covariance function $E(\mathcal{V}(f)\mathcal{V}(g)) = E(\Psi(f,X)\Psi(g,X))$, $f,g \in \mathrm{L}^2$, is obtained by standard calculation. Moreover

$$||V_{n,0}||_Q \xrightarrow{D} ||\mathcal{V}||_Q \text{ as } n \to \infty$$

follows immediately. Now, the reverse triangle inequality implies

$$\Big| ||V_n||_Q - ||V_{n,0}||_Q \Big| \leq ||V_n - V_{n,0}||_Q.$$

From Proposition 1 it follows that $||V_n - V_{n,0}||_Q \xrightarrow{P} 0$ as $n \to \infty$ and so $||V_n||_Q - ||V_{n,0}||_Q \xrightarrow{P} 0$. An application of Slutsky's theorem implies that

$$||V_n||_Q = ||V_{n,0}||_Q + (||V_n||_Q - ||V_{n,0}||_Q) \xrightarrow{D} ||\mathcal{V}||_Q \text{ as } n \to \infty.$$

Hence $T_n = ||V_n||^2_Q \xrightarrow{D} ||\mathcal{V}||^2_Q = T_\infty$ as $n \to \infty$. Because $\mathcal{V}$ is a $L^2_Q$-valued stochastic process, $T_\infty$ is a real-valued random variable. □

The following remark deals with the special case where the estimators are obtained by the maximum likelihood approach in section 2.2.

**Remark 1.** Consider the estimators obtained by the maximum likelihood approach in section 2.2. It is easy to see that Assumption 1 is satisfied. Suppose that Assumption 2 holds and that the null hypothesis $H_0$ that a scaled Brownian motion with linear drift applies with slope $\mu \in \mathbb{R}$ and volatility parameter $\sigma \in (0,\infty)$. Then, Theorem 1 implies the convergence in distribution

$$T_n \xrightarrow{D} ||\mathcal{V}||^2_Q = T_\infty \text{ as } n \to \infty,$$

where $\mathcal{V}$ is a $L^2_Q$-valued and centered Gaussian random variable; the expression

$$E\big(\mathcal{V}(f)\mathcal{V}(g)\big) = \exp\Big(-\frac{1}{2}(\sigma_f^2 + \sigma_g^2)\Big)(\exp(\sigma_{f,g}^2) + 3\sigma_f^2\sigma_g^2 - 1),\ f,g \in \mathrm{L}^2,$$

of the related covariance function is obtained.

The following remark considers an alternative expression of the limit null distribution of the test statistic.

**Remark 2.** Under additional regularity conditions on the Gaussian random variable $\mathcal{V}$ in Theorem 1, a Karhunen-Loéve expansion can be used to obtain that $\mathcal{V}$ has the same distribution as $\sum_{i=1}^\infty \sqrt{\lambda_i}\xi_i\varphi_i$, where $\sqrt{\lambda_i}\varphi_i$, $i \in \mathbb{N}$, is an orthonormal basis of $\mathrm{L}^2_Q$ consisting of eigenfunctions $\varphi_i$, $i \in \mathbb{N}$, of the related covariance operator with corresponding non-negative eigenvalues $\lambda_i$, $i \in \mathbb{N}$, and $\xi_i$, $i \in \mathbb{N}$, are independent and standard normal distributed random variables. In this case, the limit null distribution of the test statistic is that of $\sum_{j=1}^\infty \lambda_j \xi_j^2$.

### 3.2. Consistency

In order to prove consistency, a slightly stronger assumption on the parameter estimators $\hat{\mu}_n$ and $\hat{\sigma}_n$ of $\mu \in \mathbb{R}$ and $\sigma \in (0,\infty)$ is needed, stated simultaneously under the null hypothesis and under alternatives. Note that, by the law of large numbers combined with Slutsky's theorem, Assumption 1 implies the convergence in probability $(\hat{\mu}_n, \hat{\sigma}_n) \xrightarrow{P} (\mu,\sigma)$ as $n \to \infty$ if the null hypothesis applies with slope $\mu \in \mathbb{R}$ and volatility $\sigma \in (0,\infty)$. This assumption is satisfied by the estimators obtained by the maximum likelihood approach in section 2.2 as it is discussed there.

**Assumption 3.** There exist $\mu \in \mathbb{R}$, $\sigma \in (0,\infty)$ such that

$$(\hat{\mu}_n, \hat{\sigma}_n) \xrightarrow{a.s.} (\mu,\sigma) \text{ as } n \to \infty,$$

where $\mu$ is the underlying slope and $\sigma$ is the underlying volatility if the null hypothesis of a scaled Brownian motion with linear drift applies.

First, a statement about an almost sure limit of the test statistic is given. The result is valid under the null hypothesis as well as under alternatives. Similar as in Henze and Jiménez Gamero (2020), the fact that the Hilbert space $\mathrm{L}^2$ is separable is used.

**Theorem 2.** *Suppose that Assumption 3 holds with related $\mu \in \mathbb{R}$ and $\sigma \in (0,\infty)$. Denote by $\phi_{Y^*}(f)$, $f \in \mathrm{L}^2$, the characteristic functional of $Y^*$, defined by $Y^*(t) = \frac{X(t)-\mu t}{\sigma}$, $t \in [0,1]$. Then the following almost sure convergence holds:*

$$\frac{1}{n}T_n = \int_{\mathrm{L}^2} |\phi_n(f) - \phi_{0,C_1}(f)|^2 Q(\mathrm{d}f) \xrightarrow{\text{a.s.}} \int_{\mathrm{L}^2} |\phi_{Y^*}(f) - \phi_{0,C_1}(f)|^2 Q(\mathrm{d}f) = \tau_Q \text{ as } n \to \infty.$$

**Proof.** At first, the almost surely convergence of $\phi_n(f)$ to $\phi_{Y^*}(f)$ as $n \to \infty$ is shown, where the related null set does not depend on the choice of $f \in \mathrm{L}^2$. In this regard, note that eq. (5) is valid under the null hypothesis as well as under alternatives and remember on the definition of the independent and identically distributed random variables $Y_j^*$, $j = 1,\dots,n$, there. Applying for each $j = 1,\dots,n$ and each $f \in \mathrm{L}^2$ a Taylor expansion of the functions cos and sin about $\langle f, Y_j^* \rangle$, there are (random) numbers $\xi_{1,j}(f)$ and $\xi_{2,j}(f)$ between $\langle f, Y_j \rangle$ and $\langle f, Y_j^* \rangle$ such that

$$\begin{aligned}\cos\left\langle f, Y_j \right\rangle &= \cos\left\langle f, Y_j^* + \left(\frac{\hat\sigma_n - \sigma}{\hat\sigma_n} Y_j^* + \frac{(\hat\mu_n - \mu)\mathrm{id}_{\mathbb{R}}}{\hat\sigma_n}\right)\right\rangle \\ &= \cos\left\langle f, Y_j^* \right\rangle - \left\langle f, \frac{\hat\sigma_n - \sigma}{\hat\sigma_n} Y_j^* + \frac{(\hat\mu_n - \mu)\mathrm{id}_{\mathbb{R}}}{\hat\sigma_n}\right\rangle \sin(\xi_{1,j}(f))\end{aligned}$$

and

$$\begin{aligned}\sin\left\langle f, Y_j \right\rangle &= \sin\left\langle f, Y_j^* + \left(\frac{\hat\sigma_n - \sigma}{\hat\sigma_n} Y_j^* + \frac{(\hat\mu_n - \mu)\mathrm{id}_{\mathbb{R}}}{\hat\sigma_n}\right)\right\rangle \\ &= \sin\left\langle f, Y_j^* \right\rangle + \left\langle f, \frac{\hat\sigma_n - \sigma}{\hat\sigma_n} Y_j^* + \frac{(\hat\mu_n - \mu)\mathrm{id}_{\mathbb{R}}}{\hat\sigma_n}\right\rangle \cos(\xi_{2,j}(f)),\end{aligned}$$

noticing that the maps $(\omega, f) \mapsto \sin(\xi_{1,j}(f))(\omega)$ and $(\omega, f) \mapsto \cos(\xi_{2,j}(f))(\omega)$, $(\omega, f) \in \Omega \times \mathrm{L}^2$, are $(\mathfrak{A} \otimes \mathfrak{B}(\mathrm{L}^2), \mathfrak{B})$-measurable for each $j = 1,\dots,n$. Substituting this in $\cos\left(\langle f, Y_j \rangle\right) + i \sin\left(\langle f, Y_j \rangle\right)$ for each $j = 1,\dots,n$ and each $f \in \mathrm{L}^2$ gives

$$\phi_n(f) = \frac{1}{n}\sum_{j=1}^n \exp(i\left\langle Y_j, f \right\rangle) = \frac{1}{n}\sum_{j=1}^n \exp\left(i\left\langle Y_j^*, f \right\rangle\right) + R_n(f),\ f \in \mathrm{L}^2, \tag{8}$$

with the remainder term

$$R_n(f) = \frac{1}{n}\sum_{j=1}^n (i\cos(\xi_{j,2}(f)) - \sin(\xi_{j,1}(f)))\left\langle \frac{\sigma - \hat\sigma_n}{\hat\sigma_n} Y_j^* + \frac{(\mu - \hat\mu_n)\mathrm{id}_{\mathbb{R}}}{\hat\sigma_n}, f \right\rangle,\ f \in \mathrm{L}^2.$$

Using the triangle inequality, the boundedness of the functions sin and cos as well as the Cauchy-Schwarz inequality, it holds for the remainder term that

$$|R_n(f)| \le 2||f|| \left|\frac{\sigma - \hat\sigma_n}{\hat\sigma_n}\right| \frac{1}{n}\sum_{j=1}^n ||Y_j^*|| + 2||f|| \left|\frac{\mu - \hat\mu_n}{\hat\sigma_n}\right| ||\mathrm{id}_{\mathbb{R}}||,\ f \in \mathrm{L}^2.$$

By Assumption 3 there exists a set $\Omega' \in \mathfrak{A}$ with $P(\Omega') = 1$ such that $(\hat\mu_n(\omega), \hat\sigma_n^2(\omega)) \mapsto (\mu, \sigma)$ as $n \to \infty$ for all $\omega \in \Omega'$. By the strong law of large numbers there exists another set $\Omega'' \in \mathfrak{A}$ such that $P(\Omega'') = 1$ and $\frac{1}{n}\sum_{j=1}^n ||Y_j^*||(\omega) \to E(||Y^*||)$ as $n \to \infty$ for all $\omega \in \Omega''$. For this reason, it follows for all $\omega \in \Omega' \cap \Omega''$ and all $f \in \mathrm{L}^2$ that

$$R_n(f)(\omega) \mapsto 0 \text{ as } n \to \infty. \tag{9}$$

By the strong law of large numbers it follows that for each $f \in \mathrm{L}^2$ there exists a set $\Omega_f \in \mathfrak{A}$ with $P(\Omega_f) = 1$, such that for all $\omega \in \Omega_f$,

$$\lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^n \exp\left(i\left\langle Y_j^*, f \right\rangle\right)(\omega) = \phi_{Y^*}(f).$$

Since $\mathrm{L}^2$ is separable there is a countable dense subset $D \subset \mathrm{L}^2$. Using the fact that the countable intersection of sets of probability 1 is again a set of probability 1 a set $\Omega''' = \bigcap_{f\in D} \Omega_f$ is obtained such that $P(\Omega''') = 1$ with the property that the latter limit holds for all $\omega \in \Omega'''$ and all $f \in D$. Setting $\Omega_0 := \Omega' \cap \Omega'' \cap \Omega'''$ it is for all $\omega \in \Omega_0$ and all $f \in D$

$$\lim_{n\to\infty} \phi_n(f)(\omega) = \phi_{Y^*}(f). \tag{10}$$

Now, fix some $\omega \in \Omega_0$ and some $f \in \mathrm{L}^2$. In addition, let $g \in D$ be arbitrary. The triangle inequality gives

$$|\phi_n(f)(\omega) - \phi_{Y^*}(f)| \leq |\phi_n(f)(\omega) - \phi_n(g)(\omega)| + |\phi_n(g)(\omega) - \phi_{Y^*}(g)| + |\phi_{Y^*}(f) - \phi_{Y^*}(g)|.$$

Treat all three terms on the right hand side of the latter inequality. Firstly, from eq. (10) it is

$$\lim_{n\to\infty} |\phi_n(g)(\omega) - \phi_{Y^*}(g)| = 0.$$

Secondly, applying the inequality $|\exp(ix) - \exp(iy)| \leq |x - y|$, $x, y \in \mathbb{R}$, the triangle inequality and the Cauchy-Schwarz inequality to obtain

$$|\phi_{Y^*}(f) - \phi_{Y^*}(g)| \leq ||f - g|| E(||Y^*||).$$

Thirdly, using eq. (8), eq. (9) and in addition the same inequalities as before, it follows that

$$|\phi_n(f)(\omega) - \phi_n(g)(\omega)| \leq ||f - g|| \frac{1}{n} \sum_{i=1}^{n} ||Y_i^*(\omega)|| + |R_n(f)| + |R_n(g)| \to ||f - g|| E(||Y^*||) \text{ as } n \to \infty.$$

In all, it is

$$\limsup_{n\to\infty} |\phi_n(f)(\omega) - \phi_{Y^*}(f)| \leq 2||f - g|| E(||Y^*||).$$

Because $g \in D$ is arbitrary and $D$ is dense in $\mathrm{L}^2$, letting $||f - g|| \to 0$ implies $\lim_{n\to\infty} \phi_n(f)(\omega) = \phi_{Y^*}(f)$ and finally

$$\lim_{n\to\infty} |\phi_n(f) - \phi_{0,C_1}(f)|^2 = |\phi_{Y^*}(f)(\omega) - \phi_{0,C_1}(f)|^2$$

for all $\omega \in \Omega_0$ and all $f \in \mathrm{L}^2$. Because the latter term is bounded by four, the result follows from the dominated convergence theorem. □

If the null hypothesis $H_0$ is true, the almost sure limit $\tau_Q$ as $n \to \infty$ in Theorem 2 vanishes. To ensure the reverse of this implication and finally the consistency of the test under alternatives $H_1$, an additional condition on the probability measure $Q$ is required. For the two specifications of $Q$ presented in section 2.4, the stated properties are obviously satisfied. Let $B(f, \epsilon) = \{g \in \mathrm{L}^2; ||f - g|| < \epsilon\}$, $\epsilon \in (0, \infty)$, $f \in \mathrm{L}^2$, be the $\epsilon$-ball centered at $f$ and note that $B(f, \epsilon) \in \mathfrak{B}(\mathrm{L}^2)$ because it is an open subset of $\mathrm{L}^2$.

**Assumption 4.** The measure $Q$ does not vanish on open balls of $\mathrm{L}^2$, i.e., it holds that

$$Q(B(f, \epsilon)) > 0 \text{ for all } f \in \mathrm{L}^2 \text{ and all } \epsilon \in (0, \infty).$$

Finally, the following result on the test's consistency is stated.

**Theorem 3.** *Suppose that Assumption 1, Assumption 2, Assumption 3 and Assumption 4 are valid. Then the test that rejects the null hypothesis if and only if $T_n > c_{n,1-\alpha}$ is consistent, i.e., under alternatives $H_1$ it follows that*

$$\lim_{n\to\infty} P(T_n > c_{n,1-\alpha}) = 1.$$

**Proof.** From Assumption 3, Theorem 2 implies that $\frac{1}{n} T_n \xrightarrow{\text{a.s.}} \tau_Q$ as $n \to \infty$. Due to the condition that the alternative $H_1$ is valid, $Y^*$ is not a standard Brownian motion. For this reason and the fact that the distribution of the Hilbert space valued random variable is uniquely determined by its characteristic functional, see Proposition 7.6.3 in Laha and Rohatgi (1979), it is $|\phi_{Y^*}(f_0) - \phi_{0,C_1}(f_0)| > 0$ for some $f_0 \in \mathrm{L}^2$. Because the map $f \mapsto |\phi_{Y^*}(f) - \phi_{0,C_1}(f_0)|$, $f \in \mathrm{L}^2$, is continuous there exists $\epsilon > 0$ and $\delta > 0$ such that $|\phi_Y(f) - \phi_{0,C_1}(f_0)| > \delta$ for all $f \in B(f_0, \epsilon)$. Due to Assumption 4 it follows that

$$\tau_Q = \int_{\mathrm{L}^2} |\phi_{Y^*}(f) - \phi_{0,C_1}(f_0)|^2 Q(\mathrm{d}f) > \delta Q(B(f_0, \epsilon)) > 0.$$

Therefore it is $T_n \xrightarrow{\text{a.s.}} \infty$ as $n \to \infty$. It is easy to see that this implies

$$\lim_{n\to\infty} P(T_n > M) = 1 \tag{11}$$

for all fixed thresholds $M \in \mathbb{R}$. Now, it is shown that the sequence of critical values $c_{n,1-\alpha}$, $n \in \mathbb{N}$, is bounded. Because this is the sequence of quantiles of order $1 - \alpha$ of the distribution of the test statistic $T_n$ if the null hypothesis $H_0$ holds, consider the case that the null hypothesis $H_0$ is true for a moment. Due to the validity of Assumption 1 and Assumption 2, Theorem 1 implies that $T_n \xrightarrow{D} T_\infty$ as $n \to \infty$ if the null hypothesis $H_0$ is true, where $T_\infty$ is a real-valued random variable. Without loss of generality suppose that $T_n \xrightarrow{a.s.} T_\infty$ as $n \to \infty$. This implies, in particular, that

$$\lim_{n\to\infty} P(|T_n - T_\infty| > M) = 0, \tag{12}$$

for all fixed thresholds $M \in \mathbb{R}$. Assume that the sequence of critical values $c_{n,1-\alpha}$, $n \in \mathbb{N}$, is unbounded. Then there exists a subsequence $c_{n',1-\alpha}$, $n' \in \mathbb{N}$, with $\lim_{n'\to\infty} c_{n',1-\alpha} = \infty$. For this reason, for all given thresholds $M \in \mathbb{R}$, it is $c_{n',1-\alpha} \geq M$ for $n' \in \mathbb{N}$ sufficiently large. Because $c_{n',1-\alpha}$ is still a quantile of order $1-\alpha$, it is $P(T_{n'} \geq c_{n',1-\alpha}) \geq \alpha$ and so

$$\liminf_{n'\to\infty} P(T_{n'} > c_{n',1-\alpha}) \geq \alpha \tag{13}$$

on the one hand. On the other hand, given an arbitrary threshold $M \in \mathbb{R}$, elementary deliberations yield that for $n' \in \mathbb{N}$ sufficiently large

$$P(T_{n'} \geq c_{n',1-\alpha}) \leq P(|T_{n'} - T_\infty| \geq M) + P(T_\infty \geq M).$$

Using this and eq. (12) results in $\limsup_{n'\to\infty} P(T_{n'} \geq c_{n',1-\alpha}) \leq P(T_\infty \geq M)$. Because $M$ is arbitrary, letting $M \to \infty$ implies $P(T_\infty \geq M) \to 0$ and so

$$\lim_{n'\to\infty} P(T_{n'} \geq c_{n',1-\alpha}) = 0,$$

which contradicts eq. (13) and so the sequence of critical values $c_{n,1-\alpha}$, $n \in \mathbb{N}$, is bounded. Therefore, there exists a threshold $M \in \mathbb{R}$ satisfying $c_{n,1-\alpha} \leq M$ for all $n \in \mathbb{N}$. With $P(T_n > c_{n,1-\alpha}) \geq P(T_n > M)$ the statement follows from eq. (11). □

## 4. Implementation

In this section, the implementation of the test is discussed from a practical point of view. A feasible method to obtain estimates of the parameters is provided, simple formulas for the calculation of the scalar products that occur in the characteristic functionals of the test statistic are derived and a Monte-Carlo approach for the calculation of the integral distance in the test statistic is suggested.

### 4.1. Implementation of the parameter estimators

Going back to the maximum likelihood estimation approach in section 2.2, the derivation of the estimators given there provides a parameter estimation procedure that is applicable in practice, where typically only a finite number of points are available as observations and a subsequent interpolation gives the functional data. Suppose $N \in \mathbb{N}$ and a partition $0 = t_0 < \cdots < t_N = 1$ of $[0,1]$, where realizations of the related functional values $X_j(t_i)$, $i = 1,\ldots,N$, $j = 1,\ldots,n$ are observed. Then the increments $\Delta X_{i,j} = X_j(t_i) - X_j(t_{i-1})$, $i = 1,\ldots,N$, $j = 1,\ldots,n$, are obtained. The maximum likelihood estimation approach motivates the use of the random variable in eq. (1), that is

$$\hat{\mu}_{n,N} = \hat{\mu}_n = \frac{1}{n}\sum_{j=1}^{n} X_j(1),$$

as an estimator of $\mu \in \mathbb{R}$ and the random variable in eq. (2), that is

$$\hat{\sigma}_{n,N} = \sqrt{\frac{1}{nN}\sum_{j=1}^{n}\sum_{i=1}^{N}\frac{1}{t_i - t_{i-1}}\left(\Delta X_{i,j} - (t_i - t_{i-1})\hat{\mu}_{n,N}\right)^2},$$

as an approximation of the estimator $\hat{\sigma}_n$ of $\sigma \in (0,\infty)$, where $\hat{\sigma}_n$ is introduced in section 2.2.

### 4.2. Implementation of the scalar products

Evaluations of certain scalar products that occur in the characteristic functionals are required to obtain the value of the test statistic. This results in the calculation of the (double) integrals of type $\langle C_1 f, f\rangle = \int_0^1\int_0^1 \min(s,t)f(t)f(s)\mathrm{d}t\mathrm{d}s$ and $\langle f, Y\rangle = \int_0^1 f(t)Y(t)\mathrm{d}t$, where $f \in \mathrm{L}^2$ is a realization of $Q$ and $Y$ is a standard Brownian motion. The data $X_1,\ldots,X_n$ are used in the implementation. As it is explained later, a sufficiently large number $M \in \mathbb{N}$ of independent realizations $f_1,\ldots,f_M$ of $Q$ is needed for the calculation of the test statistic, and the number of evaluations for both scalar products is $n \times M$. This comes with noticeable computational costs if standard numerical integration routines are used for the implementation. Alternatively, simple formulas for the calculation of these scalar products are provided that can be used in applications to reduce the computational effort. Only a finite number of points are available as observations in practice. For this reason, suppose $N \in \mathbb{N}$ and a partition $0 = t_0 < \cdots < t_N = 1$ of $[0,1]$, where realizations of the related functional values $X_j(t_i)$, $i = 1,\ldots,N$, $j = 1,\ldots,n$ are observed. The same applies to realizations $f$ of $Q$; in practice, corresponding functional values can be obtained by $f_j(t_i)$, $i = 1,\ldots,N$, $j = 1,\ldots,M$. Subsequent interpolation results in the approximative empirically standardized data of the form

$$Y_{j,N}(t) = \sum_{i=1}^{N} \mathrm{I}_{[t_{i-1},t_i)}(t)(m_{i,j}^{(Y)}t + b_{i,j}^{(Y)}),\ t \in [0,1),\ j = 1,\ldots,n,$$

and $Y_{j,N}(1) = Y_j(1)$, where $m_{i,j}^{(Y)} = \frac{X_j(t_i) - X_j(t_{i-1}) + (t_i - t_{i-1})\hat{\mu}_{n,N}}{(t_i - t_{i-1})\hat{\sigma}_{n,N}}$, $i = 1, \dots, n$, are given slopes and $b_{i,j}^{(Y)} = \frac{Y_j(t_{i-1})t_i - Y_j(t_i)t_{i-1}}{(t_i - t_{i-1})\hat{\sigma}_{n,N}}$, $i = 1, \dots, N$, $j = 1, \dots, n$, are given intercepts of the linear interpolation. Note that the (joint) distribution of these slopes and intercepts does not depend on the true underlying parameters $\mu$ and $\sigma$ if the null hypothesis $H_0$ is valid. $\mathrm{I}_A$ denotes the indicator function of the set $A \in \mathfrak{B}$. Similarly, subsequent interpolation provides the approximative realizations of $Q$ of the form

$$f_{j,N}(t) = \sum_{i=1}^{N} \mathrm{I}_{[t_{i-1}, t_i)}(t)(m_{i,j}^{(f)} t + b_{i,j}^{(f)}), \ t \in [0,1), \ j = 1, \dots, M,$$

and $f_{j,N}(1) = f_j(1)$, where $m_{i,j}^{(f)} = \frac{f_j(t_i) - f_j(t_{i-1})}{t_i - t_{i-1}}$, $i = 1, \dots, n$, are given slopes and $b_{i,j}^{(f)} = \frac{f_j(t_{i-1})t_i - f_j(t_i)t_{i-1}}{t_i - t_{i-1}}$, $i = 1, \dots, N$, $j = 1, \dots, M$, are given intercepts of the linear interpolation. Tedious calculations give the formulas

$$\begin{aligned}
S_{j,N}^{(f)} = \left\langle C_1 f_{j,N}, f_{j,N} \right\rangle &= \sum_{i'=1}^{N} \sum_{i=1}^{i'-1} \left( \frac{m_{i,j}^{(f)}}{3}(t_i^3 - t_{i-1}^3) + \frac{b_{i,j}^{(f)}}{2}(t_i^2 - t_{i-1}^2) \right) \left( \frac{m_{i',j}^{(f)}}{2}(t_{i'}^2 - t_{i'-1}^2) + b_{i',j}^{(f)}(t_{i'} - t_{i'-1}) \right) \\
&+ \sum_{i'=1}^{N} \sum_{i=i'+1}^{N} \left( \frac{m_{i,j}^{(f)}}{2}(t_i^2 - t_{i-1}^2) + b_{i,j}^{(f)}(t_i - t_{i-1}) \right) \left( \frac{m_{i',j}^{(f)}}{3}(t_{i'}^3 - t_{i'-1}^3) + \frac{b_{i',j}^{(f)}}{2}(t_{i'}^2 - t_{i'-1}^2) \right) \\
&- \sum_{i'=1}^{N} \left( \frac{(m_{i',j}^{(f)})^2}{30}(t_{i'}^5 - t_{i'-1}^5) - \frac{2 b_{i',j}^{(f)} m_{i',j}^{(f)}}{3}(t_{i'}^4 - t_{i'-1}^4) \right) \\
&+ \sum_{i'=1}^{N} \left( \frac{m_{i',j}^{(f)} t_{i'}^2}{6} + \frac{b_{i',j}^{(f)} m_{i',j}^{(f)} t_{i'}}{3} - \frac{(b_{i',j}^{(f)})^2}{6} \right)(t_{i'}^3 - t_{i'-1}^3) \\
&+ \sum_{i'=1}^{N} \left( -\frac{(m_{i',j}^{(f)})^2 t_{i'-1}^3}{6} + \frac{b_{i',j}^{(f)} m_{i',j}^{(f)}}{4}(t_{i'}^2 - t_{i'-1}^2) + \frac{(b_{i',j}^{(f)})^2 t_{i'}}{2} \right)(t_{i'}^2 - t_{i'-1}^2) \\
&- \sum_{i'=1}^{N} \left( \frac{b_{i',j}^{(f)} m_{i',j}^{(f)} t_{i'-1}^3}{3} + \frac{(b_{i',j}^{(f)})^2 t_{i'-1}^2}{2} \right)(t_{i'} - t_{i'-1}), \ j = 1, \dots, M,
\end{aligned}$$

and

$$S_{i,j,N}^{(f,Y)} = \left\langle f_{i,N}, Y_{j,N} \right\rangle = \sum_{k=1}^{N} \left( \frac{m_{k,i}^{(f)} m_{k,j}^{(Y)}}{3}(t_k^3 - t_{k-1}^3) + \frac{m_{k,i}^{(f)} b_{k,j}^{(Y)} + m_{k,j}^{(Y)} b_{k,i}^{(f)}}{2}(t_k^2 - t_{k-1}^2) + b_{k,i}^{(f)} b_{k,j}^{(Y)}(t_k - t_{k-1}) \right),$$

$i = 1, \dots, M$, $j = 1, \dots, n$.

### 4.3. Implementation of the integral distance and the test

Remember that under Assumption 2 the alternative expression in eq. (3) of the test statistic $T_n$ and the definition of $V_n$ in eq. (4) are valid. This motivates the approximation of the integral by a mean and to use a Monte-Carlo approach on the basis of the $M$ approximative realizations $f_{i,N}$, $i = 1, \dots, n$, of $Q$ and the $n$ approximative empirically standardized data $Y_{j,N}$, $j = 1, \dots, n$, given in section 4.2 to obtain an approximation of the test statistic in applications. Given the scalar products $S_{i,N}^{(f)}$, $i = 1, \dots, n$, and $S_{i,j,N}^{(f,Y)}$, $i = 1, \dots, n$, $j = 1, \dots, n$, obtained in section 4.2, it is proposed to use the Monte-Carlo approximation

$$T_{n,N,M} = \frac{1}{M} \sum_{i=1}^{M} V_{i,n,N}^2,$$

where

$$V_{i,n,N} = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \left( \cos S_{i,j,N}^{(f,Y)} + \sin S_{i,j,N}^{(f,Y)} - \exp\left( -\frac{1}{2} S_{i,N}^{(f)} \right) \right),$$

for the test statistic $T_n$. A Monte-Carlo approximation for the critical value is proposed as well. For this purpose, repeat the procedure to obtain the approximation of the test statistic $S \in \mathbb{N}$ times, where $S$ is sufficiently large and the repetition number $s = 1, \dots, S$ based on $X_{j,s}^{(0)}(t_i)$, $i = 1, \dots, N$, where $X_{j,s}^{(0)}(t)$, $t \in [0,1]$, $j = 1, \dots, n$, $s = 1, \dots, S$, are artificial data under the null hypothesis $H_0$, e.g., independent standard Brownian motions. This results in approximations of the test statistic $T_{n,N,M,s}^{(0)}$, $s = 1, \dots, S$, generated under the null hypothesis $H_0$. Denoting by $T_{n,N,M,1:S}^{(0)} \leq \dots \leq T_{n,N,M,S:S}^{(0)}$ the related order statistic, the critical value is obtained as the

**Table 1**
Description of alternatives according to Henze and Jiménez Gamero (2020).

| Alternative | Alternative distribution | Non-normal distributed | Normal distributed |
|---|---|---|---|
| A1 | | $\alpha_0, \beta_1, \dots, \beta_5, \gamma_1, \dots, \gamma_5$ | |
| A2 | Half-normal | $\alpha_0, \beta_1, \dots, \beta_3, \gamma_1, \dots, \gamma_3$ | $\beta_4, \beta_5, \gamma_4, \gamma_5$ |
| A3 | | $\alpha_0, \beta_1, \gamma_1$ | $\beta_2, \dots, \beta_5, \gamma_2, \dots, \gamma_5$ |
| A1' | | $\alpha_0, \beta_1, \dots, \beta_5, \gamma_1, \dots, \gamma_5$ | |
| A2' | Laplace | $\alpha_0, \beta_1, \dots, \beta_3, \gamma_1, \dots, \gamma_3$ | $\beta_4, \beta_5, \gamma_4, \gamma_5$ |
| A3' | | $\alpha_0, \beta_1, \gamma_1$ | $\beta_2, \dots, \beta_5, \gamma_2, \dots, \gamma_5$ |
| A1" | | $\alpha_0, \beta_1, \dots, \beta_5, \gamma_1, \dots, \gamma_5$ | |
| A2" | Mixture | $\alpha_0, \beta_1, \dots, \beta_3, \gamma_1, \dots, \gamma_3$ | $\beta_4, \beta_5, \gamma_4, \gamma_5$ |
| A3" | | $\alpha_0, \beta_1, \gamma_1$ | $\beta_2, \dots, \beta_5, \gamma_2, \dots, \gamma_5$ |

empirical quantile $c_{n,N,M,S,1-\alpha} = T^{(0)}_{n,N,M,\lceil S(1-\alpha)\rceil:S}$ of order $1-\alpha$, where $\lceil\cdot\rceil$ denotes the ceiling function. Finally, the null hypothesis $H_0$ is rejected if and only if $T_{n,N,M} > c_{n,N,M,S,1-\alpha}$. Note that the critical value $c_{n,N,M,S,1-\alpha}$ does not depend on the data.

## 5. Simulation studies and real data examples

Finally, the performance of the test is investigated by simulation studies. Small and moderate samples are considered, with a special focus on the power under alternatives. The application of the test is demonstrated by considering stock market and interest rate time series data.

### 5.1. Simulation study I

For the empirical power comparison, tests based on the characteristic functional are considered for comparison; they have demonstrated a good power performance. For the implementation of the test, the maximum likelihood estimation approach is used. The two choices for $Q$ presented above are considered, where the test using the Wiener measure is denoted by BM(W) and the test using the random linear combinations of Legendre polynomials is denoted by BM(L). For BM(L), the distribution $\zeta$ is chosen as $\mathrm{Poi}(5)+1$, the distribution $\nu$ is chosen as $\mathrm{Poi}(1)+1$ and the distribution $\xi$ is chosen as $\mathrm{N}(0,1)$. Critical values are obtained on the basis of a Monte-Carlo simulation using 100,000 replications. The sample sizes, $n = 10, 20, 50, 100, 200$ and the significance levels $\alpha = 0.05, 0.1$ are considered. The Monte-Carlo approach to obtain the integral distance in the test statistic is based on 100 replications.

In the first simulation study, alternatives from simulations studies conducted in Henze and Jiménez Gamero (2020) and Coupek et al. (2023) are treated. In this regard, note that Henze and Jiménez Gamero (2020) considers the general testing problem of Gaussianity, that is in fact a different testing problem. In contrast, the test of Coupek et al. (2023) is developed for the testing of subclasses of Gaussian processes and can be also used for testing whether a scaled Brownian motion with linear drift applies. Although the test of Coupek et al. (2023) can be used with estimates of both, the slope and the volatility parameter, the simulation study in that paper considers only situations where either the slope or the volatility parameter is known (see Table 1 therein). It is important to note that both papers, Henze and Jiménez Gamero (2020) and Coupek et al. (2023), include simulation studies for different testing problems. In particular, the simulation results obtained are not directly comparable. However, the empirical results given in Henze and Jiménez Gamero (2020) and Coupek et al. (2023) will be considered as benchmark that gives insights into the performance of the new test. The tests in Henze and Jiménez Gamero (2020) and Coupek et al. (2023) are denoted by fBHEP and CDHH, respectively.

At first, consider different non-Gaussian alternatives in Henze and Jiménez Gamero (2020) given by linear combinations of sine and cosine functions in the form of

$$Z(t) = \alpha_0 + \sqrt{2}\sum_{j=1}^{5}\left(\beta_j \cos(2\pi j t) + \gamma_j \sin(2\pi j t)\right),\ t \in [0,1],$$

where $\alpha_0, \beta_1, \dots, \beta_5, \gamma_1, \dots, \gamma_5$, are independent random variables, with distributions according to Table 1. The alternatives A1-A3 use a standard half-normal distribution, the alternatives A1'-A3' use a standard Laplace distribution and the alternatives A1"-A3" use an equal mixture of a standard normal and a standard half-normal distribution for the non-normal distribution of the random factors.

For details, see Henze and Jiménez Gamero (2020). That work considers only sample sizes $n = 50, 100$; here $n = 10, 20$ are considered in addition. Based on 10,000 replications, empirical rejection rates (in %) of the new test BM(W) or BM(L) are reported in Table 2. For the corresponding results of fBHEP, see Henze and Jiménez Gamero (2020). Both BM(W) and BM(L) clearly start detecting all alternatives already for a sample size of $n = 10$ and have a rejection rate of nearly or exactly 100% for a sample size of $n = 20$ and higher. It is seen that the test BM(L) reports equal or higher powers than BM(W) in all cases considered. Furthermore, it is observed that BM(W) and BM(L) have higher rejection rates than is the case for fBHEP in all cases, meaning that the new test performs favorably.

Now, consider Gaussian alternatives from simulation studies in Coupek et al. (2023), that are the location and scale alternatives $W^{\mathrm{loc}}(\delta)$ and $W^{\mathrm{scl}}(\gamma)$ with parameters $\delta = 0.5$ and $\gamma = -0.5, 0.5$. For details, see Coupek et al. (2023). Based on 10,000 replications,

**Table 2**
Empirical rejection rates (in %) of the new test for non-Gaussian alternatives considered in Henze and Jiménez Gamero (2020) for different significance levels $\alpha$ and sample sizes $n$. BM(W) or BM(L) denotes the specification using the Wiener measure or random linear combinations of Legendre polynomials.

| | | BM(W) | | BM(L) | |
|---|---|---|---|---|---|
| Alternative | $n$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ |
| | 10 | 100 | 100 | 100 | 100 |
| A1 | 20 | 100 | 100 | 100 | 100 |
| | 50 | 100 | 100 | 100 | 100 |
| | 100 | 100 | 100 | 100 | 100 |
| | 10 | 99.97 | 100 | 100 | 100 |
| A2 | 20 | 100 | 100 | 100 | 100 |
| | 50 | 100 | 100 | 100 | 100 |
| | 100 | 100 | 100 | 100 | 100 |
| | 10 | 84.81 | 97.16 | 100 | 100 |
| A3 | 20 | 99.92 | 100 | 100 | 100 |
| | 50 | 100 | 100 | 100 | 100 |
| | 100 | 100 | 100 | 100 | 100 |
| | 10 | 60.26 | 87.57 | 89.10 | 97.96 |
| A1' | 20 | 99.24 | 99.94 | 99.97 | 100 |
| | 50 | 100 | 100 | 100 | 100 |
| | 100 | 100 | 100 | 100 | 100 |
| | 10 | 56.87 | 83.97 | 82.53 | 95.53 |
| A2' | 20 | 98.31 | 99.78 | 99.83 | 99.95 |
| | 50 | 100 | 100 | 100 | 100 |
| | 100 | 100 | 100 | 100 | 100 |
| | 10 | 52.48 | 79.91 | 78.78 | 93.23 |
| A3' | 20 | 97.59 | 99.73 | 99.49 | 99.93 |
| | 50 | 100 | 100 | 100 | 100 |
| | 100 | 100 | 100 | 100 | 100 |
| | 10 | 96.97 | 99.34 | 99.8 | 99.99 |
| A1" | 20 | 100 | 100 | 100 | 100 |
| | 50 | 100 | 100 | 100 | 100 |
| | 100 | 100 | 100 | 100 | 100 |
| | 10 | 88.62 | 97.51 | 99.09 | 99.94 |
| A2" | 20 | 99.94 | 100 | 100 | 100 |
| | 50 | 100 | 100 | 100 | 100 |
| | 100 | 100 | 100 | 100 | 100 |
| | 10 | 68.88 | 91.6 | 97.12 | 99.72 |
| A3" | 20 | 99.52 | 99.96 | 100 | 100 |
| | 50 | 100 | 100 | 100 | 100 |
| | 100 | 100 | 100 | 100 | 100 |

empirical rejection rates (in %) of the new test BM(W) or BM(L) are presented in Table 3. For the corresponding results of CDHH, see Coupek et al. (2023). It is seen that BM(W) performs well against location and scale alternatives with parameter $\gamma = 0.5$, where for $\gamma = -0.5$ and smaller sample sizes rejection rates are clearly lower. This observation is very similar to the results of CDHH. Furthermore, BM(L) achieves very high rejection rates in all situations under consideration, also with respect to the comparison with CDHH, where no overall excellent performance can be achieved.

### 5.2. Simulation study II

In an additional simulation study, new settings for the comparison of the new test with that in Henze and Jiménez Gamero (2020) and Coupek et al. (2023) are treated. In detail, consider alternatives of type $A(t) = W(t) + Z$, $t \in [0,1]$, referred to as random start, alternatives of type $A(t) = Z \cdot B(t)$, $t \in [0,1]$, referred to as random factor, and alternatives of type $A(t) = B(t) + Z \cdot t$, $t \in [0,1]$, referred to as random drift, where $B(t)$, $t \in [0,1]$, is a standard Brownian motion and $Z \sim \mathrm{t}_1$ is a t-distributed random variable with one degree of freedom independent of $B$. In particular, these are non-Gaussian alternatives. The simulations are conducted on the basis of contamination alternatives of type $X = (1-I)B + IA$, where the random variable $I$ follows the binomial distributions $I \sim \mathrm{Bin}(1,\mathrm{a})$ for a given probability $a \in [0,1]$ and the random variables $I$ and $(B,A)$ are independent. The value of $a$ reflects the grade of contamination of the null hypothesis $B$ by the alternative $A$; consider $a = 0, 0.2, 0.4, 0.6, 0.8, 1$. Moreover, the sample sizes $n = 10, 20, 50$ and the significance levels $\alpha = 0.05, 0.1$ are investigated.

**Table 3**
Empirical rejection rates (in %) of the new test for Gaussian alternatives considered in Coupek et al. (2023) for different significance levels $\alpha$ and sample sizes $n$. BM(W) or BM(L) denotes the specification using the Wiener measure or random linear combinations of Legendre polynomials.

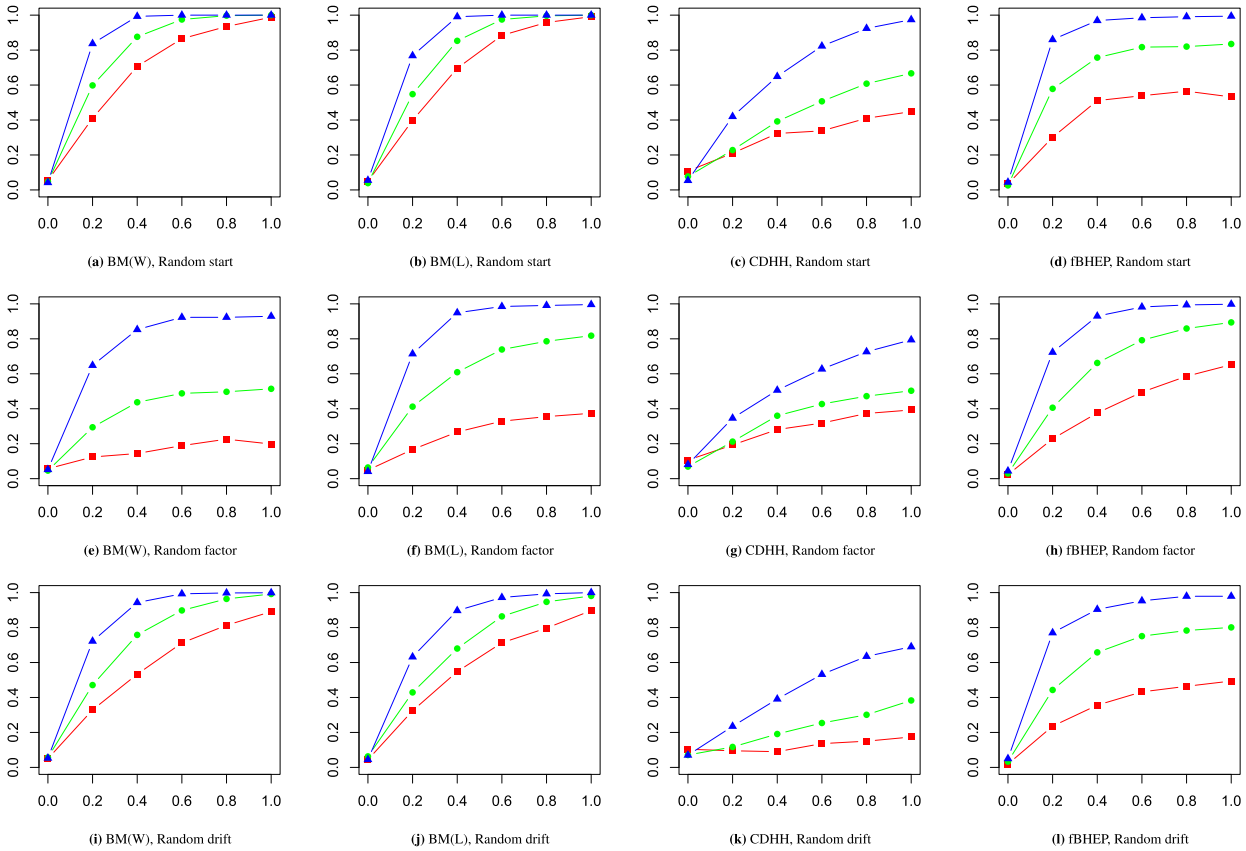
| | | BM(W) | | | BM(L) | | |
|---|---|---|---|---|---|---|---|
| | | $W^{\mathrm{loc}(\delta)}$ | $W^{\mathrm{scl}(\gamma)}$ | | $W^{\mathrm{loc}(\delta)}$ | $W^{\mathrm{scl}(\gamma)}$ | |
| $\alpha$ | $n$ | $\delta = 0.5$ | $\gamma = 0.5$ | $\gamma = -0.5$ | $\delta = 0.5$ | $\gamma = 0.5$ | $\gamma = -0.5$ |
| 0.05 | 10 | 84.71 | 100 | 8.62 | 80.58 | 100 | 100 |
| | 20 | 99.25 | 100 | 15.18 | 98.56 | 100 | 100 |
| | 50 | 100 | 100 | 53.37 | 100 | 100 | 100 |
| | 100 | 100 | 100 | 99.04 | 100 | 100 | 100 |
| | 200 | 100 | 100 | 100 | 100 | 100 | 100 |
| 0.1 | 10 | 91.08 | 100 | 17.12 | 87.78 | 100 | 100 |
| | 20 | 99.64 | 100 | 28.61 | 99.4 | 100 | 100 |
| | 50 | 100 | 100 | 79.93 | 100 | 100 | 100 |
| | 100 | 100 | 100 | 99.96 | 100 | 100 | 100 |
| | 200 | 100 | 100 | 100 | 100 | 100 | 100 |

**Table 4**
Empirical rejection rates (in %) in the setting random start for different significance levels $\alpha$, sample sizes $n$ and grades of contamination $a$. BM(W) or BM(L) denotes the specification of the new test using the Wiener measure or random linear combinations of Legendre polynomials, fBHEP is the test in Henze and Jiménez Gamero (2020) and CDHH is the test in Coupek et al. (2023).

| | | BM(W) | | BM(L) | | fBHEP | | CDHH | |
|---|---|---|---|---|---|---|---|---|---|
| $n$ | $a$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ |
| 10 | 0 | 5.4 | 9.8 | 4.9 | 9.2 | 3.4 | 7.9 | 10.8 | 16.9 |
| | 0.2 | 41 | 47.2 | 39.8 | 46.7 | 30.1 | 36.7 | 21 | 28.6 |
| | 0.4 | 70.7 | 74.9 | 69.6 | 77.1 | 51.1 | 57.5 | 32.3 | 43.1 |
| | 0.6 | 86.5 | 88.7 | 88.4 | 91.8 | 53.9 | 60.9 | 33.8 | 46.8 |
| | 0.8 | 93.5 | 94.8 | 95.8 | 97.7 | 56.5 | 64.3 | 41 | 51 |
| | 1 | 98.8 | 99 | 99.1 | 99.6 | 53.2 | 60.6 | 44.8 | 56.5 |
| 20 | 0 | 4.7 | 9.4 | 3.9 | 8.5 | 2.6 | 8 | 7.7 | 14.1 |
| | 0.2 | 59.8 | 65.5 | 54.8 | 63.1 | 57.8 | 61.6 | 22.8 | 33 |
| | 0.4 | 87.6 | 89.8 | 85.3 | 88.6 | 75.7 | 80.3 | 39.2 | 50.6 |
| | 0.6 | 97.5 | 98.2 | 97.5 | 98.1 | 81.7 | 85.6 | 50.7 | 62.2 |
| | 0.8 | 99.7 | 99.8 | 99.6 | 99.9 | 82 | 87 | 60.8 | 72.9 |
| | 1 | 100 | 100 | 100 | 100 | 83.5 | 87.2 | 66.7 | 81 |
| 50 | 0 | 4.2 | 8.8 | 5.5 | 10.1 | 4.4 | 9.8 | 5.4 | 10.6 |
| | 0.2 | 83.7 | 87 | 76.8 | 82.4 | 86 | 89.5 | 42 | 53.8 |
| | 0.4 | 99.3 | 99.5 | 99.1 | 99.8 | 96.9 | 98.2 | 64.9 | 76.5 |
| | 0.6 | 100 | 100 | 100 | 100 | 98.5 | 98.9 | 82.3 | 92.6 |
| | 0.8 | 100 | 100 | 100 | 100 | 99.1 | 99.3 | 92.4 | 96.9 |
| | 1 | 100 | 100 | 100 | 100 | 99.4 | 99.6 | 97.4 | 99.4 |

Consider the specifications BM(W) and BM(L) of the test documented in section 5.1 as well as the tests fBHEP in Henze and Jiménez Gamero (2020) and CDHH in Coupek et al. (2023) for comparison. For fBHEP, 200 replications for the bootstrap procedure are used to obtain critical values and 100 replications are used for the Monte-Carlo procedure to obtain the integral distance in the test statistic. For CDHH, also 200 replications for the bootstrap procedure are used to generate critical values, 1 is chosen as shape parameter and the identity matrix is chosen as covariance matrix of the marginal distribution of the integrator.

Based on 1,000 replications, empirical rejection rates (in %) of all tests are reported in Table 4, Table 5 and Table 6, where in addition the results are graphically illustrated in Fig. 1 and for $\alpha = 0.05$. Both specifications of the new test BM(W) and BM(L) perform very similarly and seem to have the best power properties for two of the three types of alternatives under consideration, namely random start and random drift, compared with the competitors. Already for small samples, a good performance is observed. For the remaining type of alternative, that is random factor, the fBHEP test performs best and rejection rates of CDHH are higher than of BM(W) or BM(L) for small samples. However, this out-performance disappears as the sample size increases. In what follows, a result of a measurement of the computing time at the local machine in the simulation setting is presented. For a single replication at sample size $n = 10$, the evaluation of the tests took 0.02s and 0.027s for the tests BM(W) and BM(L), respectively, and 0.299s for the test CDHH as well as 2.922s for the test fBHEP. At sample size $n = 20$, evaluation time for BM(W) and BM(L) are still at 0.02s and 0.027s, respectively, while the evaluation time for CDHH increases to 0.506s and for fBHEP to 3.715s. For the sample size $n = 50$, the evaluation time for BM(W) and BM(L) is 0.023s and 0.031s, respectively, and 1.862s for CDHH and 4.178s for fBHEP.



**(a)** BM(W), Random start **(b)** BM(L), Random start **(c)** CDHH, Random start **(d)** fBHEP, Random start

**(e)** BM(W), Random factor **(f)** BM(L), Random factor **(g)** CDHH, Random factor **(h)** fBHEP, Random factor

**(i)** BM(W), Random drift **(j)** BM(L), Random drift **(k)** CDHH, Random drift **(l)** fBHEP, Random drift

**Fig. 1.** Empirical rejection rates (ordinate) for different grades of contamination (abscissa) in the settings random start, random factor and random drift at significance level $\alpha = 0.05$ and for the sample sizes $n = 10$ (red), $n = 20$ (green) and $n = 50$ (blue). BM(W) or BM(L) denotes the specification of the new test using the Wiener measure or random linear combinations of Legendre polynomials, fBHEP is the test in Henze and Jiménez Gamero (2020) and CDHH is the test in Coupek et al. (2023). (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

**Table 5**
Empirical rejection rates (in %) in the setting random factor for different significance levels $\alpha$, sample sizes $n$ and grades of contamination $a$. BM(W) or BM(L) denotes the specification of the new test using the Wiener measure or random linear combinations of Legendre polynomials, fBHEP is the test in Henze and Jiménez Gamero (2020) and CDHH is the test in Coupek et al. (2023).

| | | BM(W) | | BM(L) | | fBHEP | | CDHH | |
|---|---|---|---|---|---|---|---|---|---|
| $n$ | $a$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ |
| 10 | 0 | 5.7 | 10.3 | 4.9 | 9.4 | 2.3 | 6.6 | 10.6 | 15.3 |
| | 0.2 | 12.4 | 19.3 | 16.7 | 25.4 | 22.8 | 28.1 | 19.3 | 25.2 |
| | 0.4 | 14.4 | 25.5 | 26.7 | 39.4 | 37.6 | 44.4 | 28.1 | 33.2 |
| | 0.6 | 18.9 | 30.4 | 32.9 | 44.7 | 49.5 | 57.7 | 31.8 | 39.8 |
| | 0.8 | 22.6 | 35.2 | 35.5 | 52.4 | 58.6 | 67 | 37.3 | 45.2 |
| | 1 | 19.8 | 32.3 | 37.4 | 52.6 | 65.2 | 74 | 39.3 | 46.3 |
| 20 | 0 | 4.6 | 9.5 | 6.4 | 10.5 | 3 | 7.2 | 6.9 | 12.3 |
| | 0.2 | 29.4 | 38 | 41.2 | 48.3 | 40.6 | 46.5 | 21.2 | 28.6 |
| | 0.4 | 43.7 | 53.4 | 60.9 | 68.9 | 66.2 | 71 | 36 | 42.8 |
| | 0.6 | 48.8 | 61.9 | 73.9 | 81.4 | 79.2 | 84 | 42.7 | 49.3 |
| | 0.8 | 49.7 | 63.6 | 78.6 | 85.9 | 85.9 | 89.8 | 47.2 | 55.7 |
| | 1 | 51.4 | 63.4 | 81.8 | 88.8 | 89.4 | 93 | 50.3 | 56.9 |
| 50 | 0 | 5.2 | 11.3 | 4.1 | 9.6 | 4.4 | 9.8 | 8.1 | 12.2 |
| | 0.2 | 64.8 | 69.6 | 71.4 | 76.2 | 72.3 | 75.4 | 34.6 | 41.6 |
| | 0.4 | 85.3 | 89.6 | 94.9 | 96.5 | 93 | 94.8 | 50.6 | 60.6 |
| | 0.6 | 92.3 | 95.3 | 98.5 | 99.1 | 98.2 | 98.9 | 62.7 | 70.1 |
| | 0.8 | 92.3 | 95.4 | 99.1 | 99.5 | 99.4 | 99.8 | 72.6 | 78.5 |
| | 1 | 92.9 | 96 | 99.6 | 99.9 | 99.8 | 99.9 | 79.4 | 83.9 |

**Table 6**
Empirical rejection rates (in %) in the setting random drift for different significance levels $\alpha$, sample sizes $n$ and grades of contamination $a$. BM(W) or BM(L) denotes the specification of the new test using the Wiener measure or random linear combinations of Legendre polynomials, fBHEP is the test in Henze and Jiménez Gamero (2020) and CDHH is the test in Coupek et al. (2023).

| | | BM(W) | | BM(L) | | fBHEP | | CDHH | |
|---|---|---|---|---|---|---|---|---|---|
| $n$ | $a$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ |
| 10 | 0 | 4.8 | 10.8 | 4.4 | 9.5 | 1.9 | 6.7 | 10.4 | 16.2 |
| | 0.2 | 32.9 | 37.5 | 32.4 | 37.8 | 23.6 | 29.5 | 9.5 | 11.9 |
| | 0.4 | 53.3 | 58.4 | 54.8 | 60 | 35.6 | 42.9 | 9 | 11.9 |
| | 0.6 | 71.2 | 74 | 71.3 | 75.1 | 43.2 | 49.7 | 13.6 | 14.8 |
| | 0.8 | 81.3 | 84.2 | 79.7 | 84.2 | 46.4 | 53.5 | 15 | 16.8 |
| | 1 | 89.1 | 92.2 | 89.7 | 91.5 | 49.4 | 57.4 | 17.4 | 19.5 |
| 20 | 0 | 5.7 | 9.9 | 6.3 | 12.2 | 3.3 | 7.9 | 7.1 | 13.3 |
| | 0.2 | 47.1 | 53.2 | 42.9 | 49.1 | 44.3 | 51.2 | 11.7 | 14.6 |
| | 0.4 | 75.8 | 79.6 | 68 | 73.8 | 65.8 | 70.8 | 19.1 | 21.1 |
| | 0.6 | 89.8 | 91.5 | 86.4 | 90.8 | 75.1 | 79.4 | 25.4 | 28.2 |
| | 0.8 | 96.4 | 97.3 | 94.7 | 96.7 | 78.3 | 83.2 | 30.1 | 31.9 |
| | 1 | 99.2 | 99.6 | 98.1 | 98.8 | 80.1 | 84.7 | 38.3 | 40.9 |
| 50 | 0 | 5.3 | 10.3 | 4.5 | 9.8 | 5 | 10.6 | 7 | 13 |
| | 0.2 | 72.2 | 75.7 | 63.2 | 70.1 | 77 | 80.4 | 23.5 | 26.2 |
| | 0.4 | 94.3 | 95.8 | 89.7 | 93.4 | 90.4 | 92.6 | 39.1 | 42.9 |
| | 0.6 | 99.3 | 99.6 | 97.2 | 98.1 | 95.3 | 96.5 | 53.2 | 56.5 |
| | 0.8 | 99.8 | 99.9 | 99.3 | 99.7 | 97.9 | 98.7 | 63.5 | 68.2 |
| | 1 | 99.9 | 99.9 | 100 | 100 | 97.9 | 98 | 69 | 71.8 |

### 5.3. Application to real data

Below the application of the new test to real data is demonstrated. Consider different financial time series, obtained by daily values from the year 2022 to 2023 for three US stock indices, namely the Dow Jones, Standard & Poors and NASDAQ, see Marketwatch (2024), Yahoo (2024) and Investing (2024) respectively, and for three US Federal Reserve government (interest rate) bonds with maturities of one month, six months and one year, see Reserve (2024). In detail, the first 17 values in each month are used for the analysis, that is the minimal number of days recorded in every month of all time series. A visualization of the stock or bond times series is given in Fig. 2 or Fig. 3.

It is investigated whether a geometric Brownian motion, in other words the Black-Scholes model, is suitable for modeling the data. For this purpose, the approach in Ditzhaus and Gaigall (2022) is used. Treating in all $K = 6$ time series and a time horizon $T_0 = 24$ (months), the stochastic processes $S^{(k)}(t)$, $t \in [0, T_0]$, $k = 1, \dots, K$, are obtained as data. Note that with the daily values only a finite number of points are available as observations and a subsequent interpolation gives the functional data exactly as it is explained in section 4. The samples are obtained by introducing the new time horizon $T = 1$ and subsequently splitting the time horizon in sections of equal length $T$. Precisely, the sample size is obtained as $n = T_0/T = 24$ and the log-returns

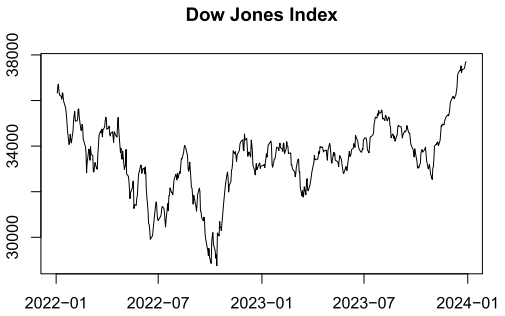
$$X_j^{(k)}(t) = \log \frac{S^{(k)}(t + (j-1)T)}{S^{(k)}((j-1)T)}, \ t \in [0, T], \ j = 1, \dots, n, \ k = 1, \dots, K,$$

are the samples from the different time series. In fact, if the model of a geometric Brownian motion applies for $k \in \{1, \dots, K\}$, $X_j^{(k)}(t)$, $t \in [0, T]$, $j = 1, \dots, n$ is a sample of size $n$ consisting of independent and identically distributed random variables with values in $L^2$, where the underlying distribution is that of a scaled Brownian motion with linear drift.

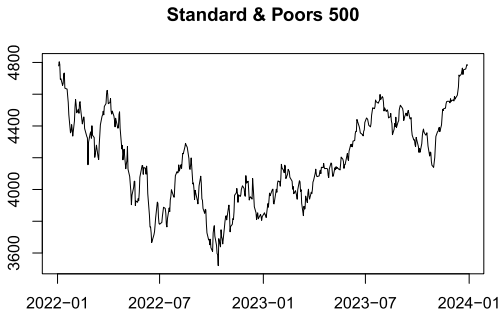
Consider the specifications BM(W) and BM(L) of the test documented in section 5.1. The resulting $p$-values are given in Table 7. For all three stock indices, it is seen that both specifications of the new test BM(W) and BM(L) report very high $p$-values that are, in particular, clearly above the typical significance level of 0.05. This indicates that a geometric Brownian motion is suitable for modeling the stock indices under consideration. In contrast, $p$-values obtained for the interest rate data are considerably lower and in particular below the typical significance level of 0.05, except for the specification BM(W) and the 1 year maturity, where the $p$-value is a bit larger on the one hand but clearly lower than for all stock indices. The overall result for the interest rate data indicates that the hypothesis of a geometric Brownian motion is not very suitable in this situation.

## 6. Conclusion

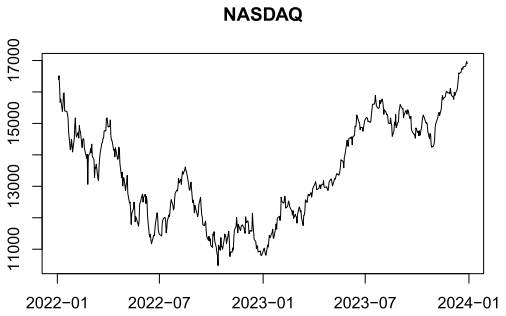
A test for the null hypothesis that observed data are realized from a geometric Brownian motion is introduced. In the functional data setting, a maximum likelihood estimation approach can be used to obtain estimators for the underlying parameters (the drift and volatility) under the null hypothesis. The test can be applied to standardized data, which means that the resulting test statistic is independent of the unknown parameters if the null hypothesis holds. Critical values are readily available and the test maintains the




**(a)** Daily values Dow Jones Index

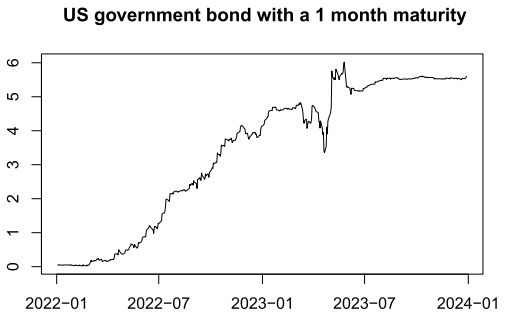



**(b)** Daily values Standard & Poors 500 index

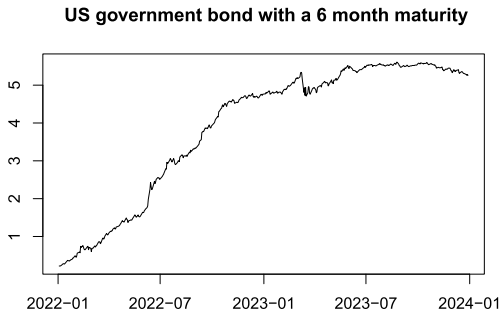



**(c)** Daily values NASDAQ index

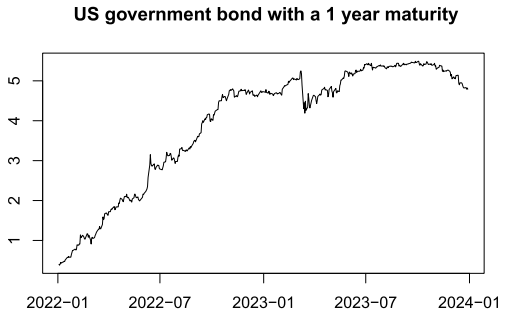
**Fig. 2.** Daily values of three US stock indices for the years 2022 – 2023.




**(a)** Daily values 1 month maturity




**(b)** Daily values 6 month maturity




**(c)** Daily values 1 year maturity

**Fig. 3.** Daily values of three US government (interest rate) bonds for the years 2022 – 2023.

**Table 7**
*p*-values for testing a geometric Brownian motion based on three US stock indices and three US government (interest rate) bond data. BM(W) or BM(L) denotes the specification of the new test using the Wiener measure or random linear combinations of Legendre polynomials.

| | | US bond | | | US stock index | | |
|---|---|---|---|---|---|---|---|
| | Dataset | 1 month | 6 month | 1 year | Dow Jones | S&P 500 | NASDAQ |
| BM(W) | *p*-value | 0.04405 | 0.01739 | 0.1209 | 0.74238 | 0.86867 | 0.99757 |
| BM(L) | *p*-value | 0.00754 | 0.00395 | 0.04352 | 0.81952 | 0.84595 | 0.97158 |

significance level in finite samples. The asymptotic properties of the test are investigated using Hilbert space methods. The asymptotic null distribution of the test statistic is obtained as the distribution of an integral of a Gaussian process, where the related covariance function is explicitly given. The test's consistency is proven using an almost sure limit which is valid under the null hypothesis as well as alternatives. Formulas and methods for a fast and simple implementation of the test are provided. It is seen that the test is powerful in a range of settings and it is demonstrated that the distributional assumptions of the Black-Scholes model can be checked using the test.

An important question to be addressed in practice is whether or not a geometric Brownian motion is an appropriate model for observed data. In this setting, the underlying drift and volatility parameters could be time-dependent or could even depend on different stochastic processes. It is likely that a generalization of the testing approach in this respect would require additional assumptions regarding the parameters. Furthermore, an alternative approach is to replace the integral distance in the test statistic with another suitable distance, e.g., a supremum type distance. The study of these possible generalizations or alternative approaches would require independent and detailed investigations. For this reason, these ideas are deferred to future research.

## Acknowledgements

The authors are grateful to the Editor, the Associate Editor and the two Referees for handling the manuscript and for valuable comments and suggestions that substantially improved the quality of the paper. The authors thank María Dolores Jiménez Gamero and Zdeněk Hlávka for providing the implementations of their tests. A special thanks goes to James Allison.

## References

Baringhaus, L., Henze, N., 1988. A consistent test for multivariate normality based on the empirical characteristic function. Metrika 35 (1), 339–348.
Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. J. Polit. Econ. 81 (3), 637–654.
Bugni, F.A., Hall, P., Horowitz, J.L., Neumann, G.R., 2009. Goodness-of-fit tests for functional data. Econom. J. 12 (s1), 1–18.
Bugni, F., Horowitz, J., 2018. Permutation tests for equality of distributions of functional data. J. Appl. Econom. 36.
Cohn, D.L., 2013. Measure Theory. Birkhäuser, New York, NY.
Coupek, P., Dolník, V., Hlávka, Z., Hlubinka, D., 2023. Fourier approach to goodness-of-fit tests for Gaussian random processes. Stat. Pap.
Cuesta-Albertos, J., del Barrio, E., Fraiman, R., Matrán, C., 2007. The random projection method in goodness of fit for functional data. Comput. Stat. Data Anal. 51 (10), 4814–4831.
Cuesta-Albertos, J., Febrero-Bande, M., 2010. A simple multiway ANOVA for functional data. Test 19, 537–557.
Cuesta-Albertos, J., Fraiman, R., Ransford, T., 2006. Random projections and goodness-of-fit tests in infinite-dimensional spaces. Bull. Braz. Math. Soc. 37, 477–501.
Cuevas, A., Fraiman, R., 2009. On depth measures and dual statistics. A methodology for dealing with general data. J. Multivar. Anal. 100, 753–766.
Ditzhaus, M., Gaigall, D., 2018. A consistent goodness-of-fit test for huge dimensional and functional data. J. Nonparametr. Stat. 30 (4), 834–859.
Ditzhaus, M., Gaigall, D., 2022. Testing marginal homogeneity in Hilbert spaces with applications to stock market returns. Test 31, 749–770.
Epps, T.W., Pulley, L.B., 1983. A test for normality based on the empirical characteristic function. Biometrika 70 (3), 723–726.
Gonzãlez-Manteiga, W., 2022. A review on specification tests for models with functional data. Span. J. Stat. 4, 9–40.
Gonzãlez-Manteiga, W., Crujeiras, R., García-Portugués, E., 2021. A review of goodness-of-fit tests for models involving functional data.
Henze, N., Jiménez Gamero, M.D., 2020. A test for Gaussianity in Hilbert spaces via the empirical characteristic functional. Scand. J. Stat. 48, 406–428.
Investing, 2024. Nasdaq stock index. https://www.investing.com/indices/nq-100.
Laha, R.G., Rohatgi, V.K., 1979. Probability Theory. Wiley Series in Probability and Mathematical Statistics. Wiley, New York.
Marketwatch, 2024. Dow Jones index. https://www.marketwatch.com/investing/index/djia/download-data.
Merton, R.C., 1973. Theory of rational option pricing. Bell J. Econ. Manag. Sci. 4 (1), 141–183.
Mishura, Y., Ralchenko, K., Shklyar, S., 2017. Maximum likelihood drift estimation for Gaussian process with stationary increments. Austrian J. Stat. 46 (3–4), 67–78.
Federal Reserve, 2024. Us government bond time series. https://www.federalreserve.gov/releases/h15/.
Revuz, D., Yor, M., 2010. Continuous Martingales and Brownian Motion. Springer, Berlin, Heidelberg.
Yahoo, 2024. Standard & poors stock index. https://de.finance.yahoo.com/quote/%5ESPX?tsrc=fin-srch.