AN ASYMPTOTIC TEST FOR CONSTANCY OF THE VARIANCE UNDER SHORT-RANGE DEPENDENCE

BY SARA K. SCHMIDT[1,*], MAX WORNOWIZKI[2,‡], ROLAND FRIED[2,§] AND
HEROLD DEHLING[1,†]

[1]*Department of Mathematics, Ruhr-Universität Bochum,* *sara.schmidt@ruhr-uni-bochum.de;*
†*herold.dehling@ruhr-uni-bochum.de*

[2]*Department of Statistics, TU Dortmund University,* ‡*wornowizki@statistik.tu-dortmund.de;* §*fried@statistik.tu-dortmund.de*

We present a novel approach to test for heteroscedasticity of a nonstationary time series that is based on Gini's mean difference of logarithmic local sample variances. In order to analyse the large sample behaviour of our test statistic, we establish new limit theorems for U-statistics of dependent triangular arrays. We derive the asymptotic distribution of the test statistic under the null hypothesis of a constant variance and show that the test is consistent against a large class of alternatives, including multiple structural breaks in the variance. Our test is applicable even in the case of nonstationary processes, assuming a locally stationary mean function. The performance of the test and its comparatively low computation time are illustrated in an extensive simulation study. As an application, we analyse Google Trends data, monitoring the relative search interest for the topic "global warming."

1. Introduction. Constancy of the variance is a common assumption and several authors have proposed tests for it. Wichern, Miller and Hsu [29], Abraham and Wei [1] and Baufays and Rasson [5] do so in the parametric framework of autoregressive models. Inclán and Tiao [23] and Gombay, Horváth and Hušková [22] propose nonparametric tests based on cumulative sums of squares against the alternative of a single structural break in a sequence of independent data. Lee and Park [24] and Wied et al. [30] extend this work to time series data fulfilling different short-range dependence conditions. Gerstenberger, Vogel and Wendler [21] use Gini's mean difference instead of sums of squares to test against the same family of alternatives. Galeano and Peña [18] and Aue et al. [4] consider the multivariate case. Chen and Gupta [12] combine binary segmentation and the Schwarz information criterion for detection of multiple change-points in independent Gaussian data. Bloomfield, Hurd and Lund [6] propose a test against the alternative of a periodic variance cycle.

In these and other papers, the mean of the data is assumed to be constant. Tests of the stationarity of the variance in the presence of a time-varying mean have been derived only recently by Dette, Wu and Zhou [16] and by Gao et al. [19]. While the latter authors assume independent Gaussian observations, the former ones apply wild bootstrap to derive critical values because the covariance structure of the limiting Gaussian process depends in a complex manner on the dependencies in the data generating process.

We construct a nonparametric asymptotic test for the constancy of the variance against the alternative of one or several change-points, allowing the data to be short-range dependent and the mean to be possibly time-varying. The statistic underlying our test has been proposed by Wornowizki, Fried and Meintanis [31], who use the permutation principle to test the constancy of the variance in a sequence of independent observations. They illustrate the advantages of this test statistic over several competitors in a series of simulation experiments.

For the construction of our asymptotic test, we develop new theory under the assumption that we observe time series data $X_1, \ldots, X_n$ generated by a particular locally stationary process in the sense of Dahlhaus [13]. Formally, we work with a triangular array

$$X_i := X_{i,n} = \sigma(i/n)Y_i + \mu(i/n), \tag{1}$$

where $(Y_i)_{i \in \mathbb{N}}$ is a stationary β-mixing process with mean zero and variance one. The local means and variances are described by the functions $\mu : [0, 1] \to \mathbb{R}$ and $\sigma^2 : [0, 1] \to (0, \infty)$, respectively. We will test the null hypothesis $\sigma(x) \equiv \sigma_H$ against the alternative that σ is a nonconstant càdlàg function. We show that under some mild regularity conditions on μ and σ, this test is consistent against all nonconstant variance functions. In contrast, tests that are specifically designed for the alternative of a single change-point in the variance, as modelled by the local variance function $\sigma^2(x) = \sigma_1^2 \mathbb{1}_{[0,\tau)}(x) + \sigma_2^2 \mathbb{1}_{[\tau,1]}(x)$, will hardly be consistent against arbitrary alternatives as considered here.

A model of the above type (1) is quite common, for instance in the context of nonparametric regression (see, for instance, Wu and Zhao [32]), where the mean function is usually assumed to be Lipschitz-continuous, possibly with jumps, and where the variance is required to be stationary. In this regard, conducting our test can be understood as a preliminary step in determining whether the latter model assumption is met. We develop our test in the first place under the assumption of a Lipschitz-continuous mean function μ, and discuss an extension to piecewise Lipschitz-continuous mean functions using the differenced time series later in Section 4. While our test for constant variance is novel even under the restrictive assumption of a constant mean, we thus allow for more realistic scenarios where the mean function is nearly constant on a small time scale $o(n)$, but possibly nonconstant on larger time scales.

More precisely, our test is based on the statistic $U(n)$, which takes the form of Gini's mean difference $\frac{1}{b_n(b_n-1)} \sum_{1 \le j \ne k \le b_n} |\nu_j - \nu_k|$ of the logarithmic local sample variances $\nu_j = \log(\hat{\sigma}_j^2)$. Specifically, we subdivide the time interval $\{1, \ldots, n\}$ into blocks of length ℓ_n, the jth block being given by $\{(j-1)\ell_n + 1, \ldots, j\ell_n\}$. Let b_n denote the number of full blocks that fit into $\{1, \ldots, n\}$, that is, $b_n = \lfloor n/\ell_n \rfloor$, and $\hat{\sigma}_j^2$ be the local sample variance in the jth block,

$$\hat{\sigma}_{j,n}^2 = \hat{\sigma}_j^2 = \frac{1}{\ell_n} \sum_{i=(j-1)\ell_n+1}^{j\ell_n} \left(X_i - \frac{1}{\ell_n} \sum_{r=(j-1)\ell_n+1}^{j\ell_n} X_r \right)^2.$$

Using this notation, the statistic $U(n)$ reads

$$U(n) = \frac{1}{b_n(b_n-1)} \sum_{1 \le j \ne k \le b_n} |\log \hat{\sigma}_j^2 - \log \hat{\sigma}_k^2|. \tag{2}$$

The use of the log-transformed local variances makes $U(n)$ scale invariant. Note that, on a broader level, $U(n)$ constitutes a U-statistic with kernel $h(x, y) = |x - y|$, whose entries are given by the triangular array $\log \hat{\sigma}_{j,n}^2$, $1 \le j \le b_n$. These entries, after proper centering and scaling, converge in distribution to a normal law. We develop U-statistic theory for this type of triangular arrays in order to show that $U(n)$ is asymptotically normal under the null hypothesis. Our results hold for absolutely regular processes, and thus cover a large class of time series models. Absolute regularity is also known under the term β-mixing and is a slightly stronger assumption than the well-known "strong mixing"-condition. Still, it covers a wide range of examples such as certain classes of Markov chains, stationary nondegenerate Gaussian processes with a particular form of the spectral density (see Bradley [8]) as well as ARMA- and GARCH-models (see Example 2.2 below).

Our test is computationally feasible even in case of huge data sets since we analyse deviations between local statistics and compare them to critical values calculated from the asymptotical distribution derived in this paper. Its low computation time is confirmed in an extensive

simulation study. Considering a wide range of data generating processes and alternatives, we find our test to have good finite sample properties and to be especially well suited in case of multiple structural breaks or nonmonotone variation of the variance function. In particular, we compare our procedure to the approach in Dette, Wu and Zhou [15, 16]. Moreover, we illustrate that our test performs well for both types of mean scenarios, Lipschitz-continuous mean functions and mean functions with jumps. As an application, we use our test to detect periods of different volatility in the annual increments of the relative search interest for the topic "global warming" retrieved from Google Trends.

The rest of the paper is structured as follows: Section 2 covers some central definitions, presents the key asymptotic results for the test statistic $U(n)$ and outlines the estimation of the long run variance in our particular setting. The main ideas of the proof are sketched in Section 3. Section 4 treats several extensions of our theory, among which are possible modifications of the test statistic, an application to data with jumps in the mean and the estimation of the change-point locations. The results of the simulation study and the data example are reported in Sections 5 and 6, respectively. All further proofs and some additional simulation results are deferred to the Supplementary Material [28].

2. Main results.

2.1. *Basic definitions and assumptions.* Throughout the paper, we assume that n observations, $X_1, \ldots, X_n$, generated by the model (1) are given, that is, $X_i = \sigma(i/n)Y_i + \mu(i/n)$, where $(Y_i)_{i \in \mathbb{N}}$ is a short-range dependent or, more precisely, an absolutely regular stationary process. In the following, we assume

$$\mu : [0, 1] \to \mathbb{R}$$

to be a Lipschitz-continuous mean function and

$$\sigma : [0, 1] \to [\sigma_0, \infty)$$

for some $\sigma_0 > 0$ to be a càdlàg-function. We want to test the hypothesis of a constant variance, that is, $\sigma \equiv \sigma_H$, against the alternative of a nonconstant variance function. Note that under the null hypothesis, we have $\mathbb{E}(X_i) = \mu(\frac{i}{n})$ and $\mathrm{Var}(X_i) = \sigma_H^2$.

Lipschitz-continuity of the mean function μ implies that the means $\mathbb{E}(X_i) = \mu(i/n)$ are nearly constant on a small time scale $o(n)$, as for $|i - r| = o(n)$, we have $\mu(i/n) - \mu(r/n) = o(1)$. In contrast, the mean function will generally be nonconstant on the larger time scale $O(n)$, as $\mathbb{E}(X_{c_1 \cdot n}) - \mathbb{E}(X_{c_2 \cdot n}) = \mu(c_1) - \mu(c_2)$ for $c_1, c_2 \in [0, 1]$. In fact, one can even relax the assumption on μ to Hölder-continuity, see Remark 4.3 below, but we will restrict ourselves to the notationally more feasible case of Lipschitz-continuity. In addition, we will later also present an approach that can handle piecewise Lipschitz-continuous mean functions by considering first order differences of the data.

The dependence structure of $(X_i)_{i \in \mathbb{N}}$ is determined by the underlying process $(Y_i)_{i \in \mathbb{N}}$, which is assumed to be absolutely regular. Such processes are also known under the term β-mixing.

DEFINITION 2.1. A sequence of random variables $(Y_i)_{i \in \mathbb{N}}$ is called absolutely regular if

$$\beta_Y(k) := \sup_{m \in \mathbb{N}} \beta\big(\sigma(Y_i, 1 \leq i \leq m), \sigma(Y_i, k + m \leq i \leq \infty)\big) \to 0 \quad \text{as } k \to \infty,$$

where the β-mixing coefficient of two σ-fields $\mathcal{A}$ and $\mathcal{B}$ is given by

$$\beta(\mathcal{A}, \mathcal{B}) := \mathbb{E}\big(\operatorname*{ess\,sup}_{A \in \mathcal{A}} |\mathbb{P}(A|\mathcal{B}) - \mathbb{P}(A)|\big).$$

Regarding the moments and the mixing rate of the process $(Y_i)_{i \in \mathbb{N}}$, we assume that there exists a $\vartheta > 0$ such that the following two conditions are satisfied:

$$(A1) \quad \mathbb{E}(|Y_1|^{4+2\vartheta}) < \infty,$$

$$(A2) \quad \sum_{k=1}^{\infty} \beta_Y(k)^{\vartheta/(2+\vartheta)} < \infty.$$

Given these assumptions, the long run variance

$$\kappa^2 := \operatorname{Var}(Y_1^2) + 2 \sum_{k=1}^{\infty} \operatorname{Cov}(Y_1^2, Y_{k+1}^2)$$

is finite since $\beta_{Y^2}(k) \le \beta_Y(k)$ (see Theorem 10.7 in Bradley [9]). Throughout, we will assume the long run variance to be strictly positive, $\kappa^2 > 0$.

Most of the classical models in time series analysis satisfy the above assumptions. The following example points out some suitable processes that meet all the conditions required to derive the asymptotic results below.

EXAMPLE 2.2.

1. Let $(Y_i)_{i \in \mathbb{N}}$ be a strictly stationary, causal ARMA(p, q)-process following the model:

$$Y_i = \varepsilon_i + \sum_{j=1}^{p} \alpha_j Y_{i-j} + \sum_{m=1}^{q} \beta_m \varepsilon_{i-m}$$

with independent innovations $(\varepsilon_i)_{i \in \mathbb{N}}$ and with all roots of $1 - \sum_{j=1}^{p} \alpha_j z^j$ being larger than one in absolute value. Additionally, assume that the AR-and MA-polynomials have no common roots, and that the innovations $(\varepsilon_i)_{i \in \mathbb{N}}$ have an absolutely continuous distribution with respect to the Lebesgue measure. Then, $(Y_i)_{i \in \mathbb{N}}$ is absolutely regular with a geometric rate, that is, $\beta(k) = O(e^{-\xi k})$ for some $\xi > 0$ (see Theorem 1 in Mokkadem [26]).

2. Strictly stationary GARCH(p, q)-processes are strictly stationary solutions to the equations

$$Y_i = \sigma_i \varepsilon_i \quad \text{and} \quad \sigma_i^2 = \alpha_0 + \sum_{j=1}^{p} \alpha_j Y_{i-j}^2 + \sum_{m=1}^{q} \beta_m \sigma_{i-m}^2.$$

They are likewise absolutely regular with a geometric rate if the i.i.d. noise sequence $(\varepsilon_i)_{i \in \mathbb{N}}$ has finite absolute rth moment for some $r \in (0, \infty)$, that is, $\mathbb{E}(|\varepsilon_1|^r) < \infty$, and if ε_1 has an absolutely continuous distribution with a density that is strictly positive in a neighbourhood of zero (see Lindner [25] and the references therein for this result and an overview on the existence of strictly stationary solutions and the existence of moments for GARCH-models).

The aim of this paper is to test for changes in the variance of the time series $(X_i)_{i \in \mathbb{N}}$ by means of the statistic $U(n) = \frac{1}{b_n(b_n-1)} \sum_{1 \le j \ne k \le b_n} |\log \hat{\sigma}_j^2 - \log \hat{\sigma}_k^2|$ from (2), which compares the local estimates $\hat{\sigma}_j^2$, $j = 1, \ldots, b_n$, derived from splitting the data into b_n blocks of length ℓ_n. Both b_n and ℓ_n are assumed to grow with the sample size and, for simplicity, to be integers. Moreover, we have to impose certain growth restrictions on the block length ℓ_n and therewith the number of blocks b_n in order to ensure the desired convergence of the test statistic. Throughout the paper, we will need the block length ℓ_n to grow faster than the number of blocks. In particular, if we set $\ell_n = n^s$ for some $s \in (0, 1)$ and consequently $b_n = n^{1-s}$, this translates to $s > 0.5$. A change in the variance should result in large differences between the block-estimates $\hat{\sigma}_j^2$ and ultimately in a high value of $U(n)$, which will lead to a rejection of the hypothesis.

2.2. *Asymptotical results.* First, we show the convergence of the statistic $U(n)$ defined in (2) towards a two-dimensional Riemann-integral.

THEOREM 2.3. *Given the assumptions* (A1) *and* (A2) *and if* $\ell_n = n^s$ *with* $s \in (0.5, 0.75)$, *it holds*

$$U(n) \xrightarrow{\mathbb{P}} \int_0^1 \int_0^1 |\log \sigma^2(x) - \log \sigma^2(y)| \, dx \, dy \quad \text{as } n \to \infty.$$

Obviously, $U(n)$ converges to 0 in probability under the null hypothesis, while it converges to a strictly positive value under any alternative for which σ is not almost surely constant with respect to the Lebesgue measure on $[0, 1]$. This will imply that our test is consistent against the very general alternative of a changing variance function.

By standardizing the statistic $U(n)$ via the long run variance κ^2 and by using an appropriate scaling $\sqrt{\ell_n}$, we can now state a law of large numbers for $U(n)$ under the null hypothesis.

THEOREM 2.4. *Let the above assumptions* (A1) *and* (A2) *be fulfilled and let* $\ell_n = n^s$ *with* $s \in (0.5, 0.75)$. *If there exists a sequence* $m_n \to \infty$ *as* $n \to \infty$ *such that* $m_n = o(n^{2s-1})$ *and* $b_n \beta_Y(m_n) \to 0$, *then it holds under the null hypothesis*

$$\frac{\sqrt{\ell_n}}{\kappa} U(n) \xrightarrow{\mathbb{P}} \frac{2}{\sqrt{\pi}} \quad \text{as } n \to \infty.$$

Theorem 2.4 already reveals the double asymptotics governing the statistic $U(n)$. While there holds a central limit theorem for the inner block sums $\hat{\sigma}_j$ for which we need the scaling $\sqrt{\ell_n}$, there holds a law of large numbers for the outer structure of a U-statistic. Under the additional assumption of polynomially decaying mixing coefficients, a central limit theorem for the outer U-statistic holds as well.

THEOREM 2.5. *Assume there exist constants* $0 < \delta \leq 1$ *and* $\rho > 1 \vee \frac{9\delta}{(\delta+1)(\delta+2)}$ *such that* $\mathbb{E}(|Y_1|^{4+2\delta}) < \infty$ *and for all* $k \in \mathbb{N}$ *it holds* $\beta_Y(k) \leq Ck^{-\rho(2+\delta)(1+\delta)/\delta^2}$. *Moreover, choose* $\ell_n = n^s$ *with* $s \in (0.5, 0.75)$ *as well as* $s > 1/(1 + \delta\frac{\rho-1}{\rho+1}) \vee (1 + \frac{\delta^2}{\rho(2+\delta)(1+\delta)})/(2 + \frac{\delta^2}{\rho(2+\delta)(1+\delta)})$. *Then it holds under the null hypothesis*

$$\sqrt{b_n}\left(\frac{\sqrt{\ell_n}}{\kappa} U(n) - \frac{2}{\sqrt{\pi}}\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{4}{3} + \frac{8}{\pi}(\sqrt{3} - 2)\right) \quad \text{as } n \to \infty.$$

EXAMPLE 2.6. Coming back to Example 2.2, consider the ARMA(p, q)- and GARCH(p, q)-models presented there with independent standard normally distributed innovations. Their mixing coefficients decay at a geometric rate, $\beta(k) = O(e^{-\xi k})$ for some $\xi > 0$, which corresponds to the border case $\rho \to \infty$ in Theorem 2.5. The conditions incorporating ρ thus boil down to $s > 1/(1 + \delta)$. Put differently, we may choose the tuning parameter $s \in (0.5, 0.75)$ and have to ensure $\mathbb{E}(|Y_1|^{2+2/s+\varepsilon}) < \infty$ for some $\varepsilon > 0$ for Theorem 2.5 to hold.

Theorem 2.5 requires an additional outer scaling factor $\sqrt{b_n}$ depending on the number of blocks for the U-statistic central limit theorem to hold. Referring to the theory of U-statistics, one can then derive convergence towards a normal distribution whose variance $\psi^2 = 4 \text{Var}(h_1(Z))$ with $h_1(x) := \mathbb{E}(|x - Z'|) - 2/\sqrt{\pi}$ for two independent standard normally distributed random variables Z and Z' equals the limit $\lim_{n\to\infty} \text{Var}(\sqrt{n}\tilde{U}_n)$, where $\tilde{U}_n$

denotes Gini's mean difference computed from a sample of n i.i.d. standard normal observations. The latter limit can be calculated explicitly, see Gerstenberger and Vogel [20], such that $\psi^2 = \frac{4}{3} + \frac{8}{\pi}(\sqrt{3} - 2)$.

In particular, Theorem 2.5 can be used for structural break testing. Based on the data $x_1, \ldots, x_n$, one can compute the value of the properly standardized test statistic and compare it to the asymptotic critical values obtained from the limit distribution. We shall reject the hypothesis of a constant variance if the computed value exceeds the $(1 - \alpha)$-quantile of the $\mathcal{N}(0, \frac{4}{3} + \frac{8}{\pi}(\sqrt{3} - 2))$-distribution. Note that this test is consistent against arbitrary alternatives for which $\sigma(x)$ is not almost surely constant with respect to the Lebesgue measure on $[0, 1]$. A key factor is the reliable estimation of the unknown long run variance κ^2, which is discussed in Section 2.3.

2.3. *Estimation of the long-run variance.*

For a practical implementation of our test, we need to estimate the long-run variance

$$\kappa^2 = \mathrm{Var}(Y_1^2) + 2 \sum_{k=1}^{\infty} \mathrm{Cov}(Y_1^2, Y_{k+1}^2).$$

In the literature, there are various procedures for estimating such long-run variances. We employ the subsampling approach, introduced by Carlstein [11]. We will use the relation $\mathbb{E}(|\frac{1}{\sqrt{n}} \sum_{i=1}^{n}(Y_i^2 - \mathbb{E}(Y_1^2))|) \to \kappa \sqrt{\frac{2}{\pi}}$ for the construction of our estimator. For standard subsampling, we divide the observations into $\tilde{b}_n$ nonoverlapping blocks of length $\tilde{\ell}_n$, and consider the estimator

$$\frac{1}{\tilde{b}_n} \sqrt{\frac{\pi}{2}} \sum_{j=1}^{\tilde{b}_n} \left| \frac{1}{\sqrt{\tilde{\ell}_n}} \sum_{i=(j-1)\tilde{\ell}_n+1}^{j\tilde{\ell}_n} \left(Y_i^2 - \frac{1}{n} \sum_{r=1}^{n} Y_r^2 \right) \right|.$$

Various authors have established consistency of this estimator for a wide class of short-range dependent data.

As we do not observe the Y_i^2 directly, but only $X_i = \sigma(\frac{i}{n})Y_i + \mu(\frac{i}{n})$, we need to modify the standard subsampling procedure. We will use a subsampling that is consistent under the null hypothesis $\sigma(x) \equiv \sigma_H$, that is, when the observations are given by $X_i = \sigma_H Y_i + \mu(\frac{i}{n})$. We first center the observations by their local means, defining

$$\tilde{X}_i = X_i - \frac{1}{\ell_n} \sum_{r=(j-1)\ell_n+1}^{j\ell_n} X_r \quad \text{for } i \in \{(j-1)\ell_n + 1, \ldots, j\ell_n\}.$$

Setting $\hat{\sigma}_H^2 = \frac{1}{n} \sum_{i=1}^{n} \tilde{X}_i^2$, we then define the subsampling long-run variance estimator

$$\hat{\kappa} := \frac{1}{\tilde{b}_n} \sqrt{\frac{\pi}{2}} \frac{1}{\hat{\sigma}_H^2} \sum_{j=1}^{\tilde{b}_n} \left| \frac{1}{\sqrt{\tilde{\ell}_n}} \sum_{i=(j-1)\tilde{\ell}_n+1}^{j\tilde{\ell}_n} (\tilde{X}_i^2 - \hat{\sigma}_H^2) \right|.$$

The next proposition shows that consistency of the above estimator $\hat{\kappa}$ indeed remains valid given the additional scaling factor $\sqrt{b_n}$.

PROPOSITION 2.7. *Assume that there exist constants $0 < \delta \leq 1$ and $\rho > 1$ such that $\mathbb{E}(|Y_1|^{4+2\delta}) < \infty$ and for all $k \in \mathbb{N}$ it holds $\beta_Y(k) \leq Ck^{-\rho(2+\delta)(1+\delta)/\delta^2}$. Moreover, let $\ell_n = n^s$ and $\tilde{\ell}_n = n^q$ such that $s > 0.5$, $1 - s < q\delta(\rho - 1)/(\rho + 1)$, $q < s$ and $q < 3(1 - s)$. Then it holds under the null hypothesis*

$$\sqrt{b_n}|\hat{\kappa} - \kappa| \xrightarrow{\mathbb{P}} 0 \quad \text{as } n \to \infty.$$

By Proposition 2.7, we can replace the long run variance κ^2 in the central limit Theorem 2.5 by its estimator:

COROLLARY 2.8. *Assume there exist constants $0 < \delta \leq 1$ and $\rho > 1 \vee \frac{9\delta}{(\delta+1)(\delta+2)}$ such that $\mathbb{E}(|Y_1|^{4+2\delta}) < \infty$ and for all $k \in \mathbb{N}$ it holds $\beta_Y(k) \leq Ck^{-\rho(2+\delta)(1+\delta)/\delta^2}$. Moreover, let $\ell_n = n^s$ and $\tilde{\ell}_n = n^q$ such that $s \in (0.5, 0.75)$, $s > 1/(1 + \delta\frac{\rho-1}{\rho+1}) \vee (1 + \frac{\delta^2}{\rho(2+\delta)(1+\delta)})/(2 + \frac{\delta^2}{\rho(2+\delta)(1+\delta)})$, $1 - s < q\delta(\rho - 1)/(\rho + 1)$, $q < s$ and $q < 3(1 - s)$. Then it holds under the null hypothesis*

$$\sqrt{b_n}\left(\frac{\sqrt{\tilde{\ell}_n}}{\hat{\kappa}}U(n) - \frac{2}{\sqrt{\pi}}\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{4}{3} + \frac{8}{\pi}(\sqrt{3} - 2)\right) \quad \text{as } n \to \infty.$$

Long run variance estimators often have the drawback of diverging under the alternative. This is unfortunate since the long run variance serves in the denominator as a standardization and its overestimation thus results in a nonnegligible loss of power. Still, the following proposition ensures that the divergence here is not fast enough to cancel out the growth rate associated with the test statistic in the numerator.

PROPOSITION 2.9. *Let the assumptions (A1) and (A2) be fulfilled and assume $q < s$ as well as $q < 4(1 - s)$ for $\ell_n = n^s$ and $\tilde{\ell}_n = n^q$. Then it holds*

$$\frac{1}{\sqrt{\tilde{\ell}_n}}\hat{\kappa} \xrightarrow{\mathbb{P}} \frac{\int_0^1 |\sigma^2(x) - \int_0^1 \sigma^2(y)\,dy|\,dx}{\int_0^1 \sigma^2(z)\,dz} \quad \text{as } n \to \infty.$$

COROLLARY 2.10. *Given the assumptions (A1) and (A2) and if $\ell_n = n^s$ with $s \in (0.5, 0.75)$ as well as $\tilde{\ell}_n = n^q$ with $q < s$, it holds*

$$\sqrt{\tilde{\ell}_n}\frac{U(n)}{\hat{\kappa}} \xrightarrow{\mathbb{P}} \frac{\int_0^1 \int_0^1 |\log \sigma^2(x) - \log \sigma^2(y)|\,dx\,dy \cdot \int_0^1 \sigma^2(z)\,dz}{\int_0^1 |\sigma^2(x) - \int_0^1 \sigma^2(y)\,dy|\,dx} \quad \text{as } n \to \infty.$$

Due to $\tilde{\ell}_n = o(\ell_n)$, the expression within the central limit theorem thus still diverges under the alternative, though at a slower rate.

3. Outline and main ideas of the proofs.
This section outlines the key ideas to prove our main results, whereas the technical details are given in the Supplementary Material ([28], Supplement A). The proofs rely on a series of approximations of the statistic

$$U(n) = \frac{1}{b_n(b_n - 1)} \sum_{1 \leq j \neq k \leq b_n} |\log \hat{\sigma}_j^2 - \log \hat{\sigma}_k^2|$$

by simpler statistics that are easier to analyze and have the same large sample behavior as $U(n)$. In a first step, we replace the block means $\frac{1}{\ell_n}\sum_{i=(j-1)\ell_n+1}^{j\ell_n} X_i$ in the definition of the local sample variance $\hat{\sigma}_j^2$ by the expected values $\mu(i/n)$. For the resulting U-statistic $U_1(n)$ one can then derive an analogue of Theorem 2.3.

To obtain additional limit theory under the null hypothesis in Theorems 2.4 and 2.5, two more approximations are required. In a second approximation step, we linearize the logarithm, that is, we use the approximation $\log(1 + x) \approx x$, which is valid for x close to 0. Finally, employing a coupling technique for absolutely regular observations (see Borovkova, Burton and Dehling [7]), we replace the slightly shortened dependent blocks

$(X_{(j-1)\ell_n+1}, \ldots, X_{j\ell_n-m_n})$ by close-by independent blocks $(X'_{(j-1)\ell_n+1}, \ldots, X'_{j\ell_n-m_n})$ with the same marginal distributions. The resulting U-statistic $U_3(n)$ can be analyzed by an adaptation of U-statistic theory to triangular arrays since its entries stem from a triangular array of row-wise independent random variables.

3.1. *A first approximation.* Centering the observations X_i in each block j by their expected value $\mu(i/n)$ instead of the block mean $\frac{1}{\ell_n} \sum_{r=(j-1)\ell_n+1}^{j\ell_n} X_r$, we obtain the following approximation to the empirical block variances:

$$s_{j,n}^2 := s_j^2 := \frac{1}{\ell_n} \sum_{i=(j-1)\ell_n+1}^{j\ell_n} \left(X_i - \mu\left(\frac{i}{n}\right) \right)^2 = \frac{1}{\ell_n} \sum_{i=(j-1)\ell_n+1}^{j\ell_n} \sigma^2\left(\frac{i}{n}\right) Y_i^2.$$

Taking Gini's mean difference of the $\log s_j^2$, we get the statistic

$$U_1(n) = \frac{1}{b_n(b_n - 1)} \sum_{1 \le j \ne k \le b_n} |\log s_j^2 - \log s_k^2|.$$

Our first approximation theorem shows that $U_1(n)$ is sufficiently close to $U(n)$ for all limit theorems to carry over.

PROPOSITION 3.1. *Assume that conditions* (A1) *and* (A2) *hold and that* $\ell_n = n^s$ *with* $s \in (0.5, 0.75)$. *Then, we have*

$$\sqrt{n}|U(n) - U_1(n)| \xrightarrow{\mathbb{P}} 0,$$

as $n \to \infty$.

Proposition 3.1 holds both under the null hypothesis as well as under the alternative. Thus, we may henceforth restrict our analysis to the U-statistic $U_1(n)$, and to the centered data $X_i - \mu(i/n) = \sigma(i/n)Y_i$. Without loss of generality, we may assume from now on that $\mu(i/n) = 0$, and that the data are given by

$$X_i = \sigma\left(\frac{i}{n}\right) Y_i.$$

3.2. *Outline of the proof of Theorem* 2.3. Recall that the statistic $U_1(n)$ employs the arguments $\log s_j^2$ for $1 \le j \le b_n$, where $s_j^2 = \frac{1}{\ell_n} \sum_{i=(j-1)\ell_n+1}^{j\ell_n} \sigma^2(\frac{i}{n}) Y_i^2$. Their mean can be approximated via

$$\mathbb{E}(s_j^2) = \sum_{i=(j-1)\ell_n+1}^{j\ell_n} \sigma^2\left(\frac{i}{n}\right) \approx \sigma^2\left(\frac{j\ell_n}{n}\right) = \sigma^2\left(\frac{j}{b_n}\right)$$

and one can moreover show that $\text{Var}(s_j^2) \to 0$ as $n \to \infty$. Hence, we set $s_j^2 \approx \sigma^2(j/b_n)$ and this in turn implies

$$U_1(n) \approx \frac{1}{b(b_n - 1)} \sum_{1 \le j \ne k \le b_n} \left| \log \sigma^2\left(\frac{j}{b_n}\right) - \log \sigma^2\left(\frac{k}{b_n}\right) \right|.$$

The latter double sum is a Riemann-type approximation of the desired limit integral $\int_0^1 \int_0^1 |\log \sigma^2(x) - \log \sigma^2(y)| \, dx \, dy$. A rigorous proof is given in the Supplementary Material.

3.3. *Two further approximations.* The second and third approximation are essential ingredients in the analysis of the asymptotic behavior of $U(n)$ under the null hypothesis, that is, when $\sigma(x) \equiv \sigma_H$. In this case, we have $s_j^2 = \sigma_H^2 \frac{1}{\ell_n} \sum_{i=(j-1)\ell_n+1}^{j\ell_n} Y_i^2$ and hence

$$\left| \log s_j^2 - \log s_k^2 \right| = \left| \log\left(\frac{1}{\ell_n} \sum_{i=(j-1)\ell_n+1}^{j\ell_n} Y_i^2 \right) - \log\left(\frac{1}{\ell_n} \sum_{i=(k-1)\ell_n+1}^{k\ell_n} Y_i^2 \right) \right|$$

$$= \left| \log(1 + S_j) - \log(1 + S_k) \right|,$$

where we have defined

$$S_j := S_{j,n} := \frac{1}{\ell_n} \sum_{i=(j-1)\ell_n+1}^{j\ell_n} Y_i^2 - 1 = \frac{s_j^2}{\sigma_H^2} - 1.$$

Since $\mathbb{E}(Y_i) = 0$ and $\mathrm{Var}(Y_i) = 1$, the law of large numbers implies that for a sufficiently large sample size, S_j is close to 0. This motivates a Taylor expansion of $\log(1 + x)$ around $x = 0$. Replacing $\log(1 + S_j)$ by S_j in the definition of $U_1(n)$ yields the statistic

$$U_2(n) := \frac{1}{b_n(b_n - 1)} \sum_{1 \leq j \neq k \leq b_n} |S_j - S_k|.$$

PROPOSITION 3.2. *Assume that conditions* (A1) *and* (A2) *hold, and that* $\ell_n = n^s$ *with* $s > 0.5$. *Then, under the null hypothesis,*

$$\sqrt{n}(U_1(n) - U_2(n)) \xrightarrow{\mathbb{P}} 0,$$

as $n \to \infty$.

In a third and final approximation step, we replace the arguments S_j, $1 \leq j \leq b_n$, of $U_2(n)$ by close-by independent random variables S_j', $1 \leq j \leq b_n$. In order to achieve this, we use a coupling technique for β-mixing processes. First, we divide each of the blocks $B_{j,n} = (Y_{(j-1)\ell_n+1}, \ldots, Y_{j\ell_n})$ into a long block

$$\tilde{B}_j := \tilde{B}_{j,n} := (Y_{(j-1)\ell_n+1}, \ldots, Y_{j\ell_n-m_n})$$

of length $\ell_n - m_n$ and a short block

$$R_j := R_{j,n} := (Y_{j\ell_n-m_n+1}, \ldots, Y_{j\ell_n})$$

of length $m_n = o(\ell_n)$. The latter blocks function as separation between the main blocks whose interdependence decreases when $m_n \to \infty$, as $n \to \infty$. At the same time, the separating blocks R_j need to be sufficiently short to be asymptotically negligible compared to the longer blocks $\tilde{B}_j$, and thus m_n should grow only slowly. By Lemma 2.4 in Borovkova, Burton and Dehling [7], there exists a sequence of i.i.d. random vectors

$$\tilde{B}'_{j,n} := \tilde{B}'_j := (Y'_{(j-1)\ell_n+1}, \ldots, Y'_{j\ell_n-m_n})$$

with the same marginal distribution as $\tilde{B}_j$ such that

$$\mathbb{P}(\tilde{B}'_j = \tilde{B}_j) = 1 - \beta_Y(m_n),$$

for all $1 \leq j \leq b_n$. Define the corresponding block sums

$$\tilde{S}'_{j,n} := \tilde{S}'_j := \frac{1}{\ell_n} \sum_{i=(j-1)\ell_n+1}^{j\ell_n-m_n} ((Y_i')^2 - 1),$$

and note that for any n, the random variables $\tilde{S}'_j$, $1 \leq j \leq b_n$, are independent and identically distributed. Finally, we consider the U-statistic

$$U_3(n) := \frac{1}{b_n(b_n - 1)} \sum_{1 \leq j \neq k \leq b_n} |\tilde{S}'_j - \tilde{S}'_k|.$$

PROPOSITION 3.3. *Assume that $\ell_n = n^s$ with $s > 0.5$, and that m_n is chosen in such a way that $m_n = o(n^{2s-1})$ as well as $b_n \beta(m_n) \to 0$. Then, under the null hypothesis,*

$$\sqrt{n}|U_2(n) - U_3(n)| \xrightarrow{\mathbb{P}} 0,$$

as $n \to \infty$.

REMARK 3.4. The exact choice of m_n is insignificant for the structural break testing. For polynomially decaying mixing coefficients, that is, $\beta_Y(k) \leq Ck^{-\rho(2+\delta)(1+\delta)/\delta^2}$, the conditions $m_n = o(n^{2s-1})$ and $b_n \beta_Y(m_n) \to 0$ are met as long as one chooses $s > (1 + \frac{\delta^2}{\rho(2+\delta)(1+\delta)})/(2 + \frac{\delta^2}{\rho(2+\delta)(1+\delta)})$.

3.4. *Outline of the proof of Theorem* 2.4. Given the above approximations, under the null hypothesis, it suffices to analyze the asymptotic behavior of the U-statistic $U_3(n)$, whose entries $\tilde{S}'_j = \frac{1}{\ell_n} \sum_{i=(j-1)\ell_n+1}^{j\ell_n-m_n}((Y'_i)^2 - 1)$ form a row-wise independent and identically distributed triangular array. Moreover, $\mathbb{E}(\tilde{S}'_1) = 0$ and given the assumptions (A1) and (A2), we have

$$\kappa_n^2 := \mathrm{Var}(\sqrt{\ell_n}\tilde{S}'_1)$$

$$= \frac{\ell_n - m_n}{\ell_n} \mathrm{Var}(Y_1^2) + 2 \sum_{k=1}^{\ell_n-m_n-1} \frac{\ell_n - m_n - k}{\ell_n} \mathrm{Cov}(Y_1^2, Y_{k+1}^2) \longrightarrow \kappa^2,$$

by dominated convergence. By the central limit theorem for partial sums of β-mixing processes, $\sqrt{\ell_n}S'_j$ converges in distribution to a normal law with mean 0 and variance κ^2. In our further analysis, we will essentially show that we may replace S'_j by $\frac{\kappa}{\sqrt{\ell_n}}Z_j$, where Z_j are independent standard normal random variables.

We first establish a law of large numbers for a rescaled version of $U_3(n)$.

PROPOSITION 3.5. *Assume that* (A1) *and* (A2) *hold and that $m_n = o(\ell_n)$. Then, under the hypothesis,*

$$\frac{\sqrt{\ell_n}}{\kappa} U_3(n) \xrightarrow{\mathbb{P}} \frac{2}{\sqrt{\pi}},$$

as $n \to \infty$.

PROOF. In order to prove Proposition 3.5, we first show that $\mathbb{E}(\frac{\sqrt{\ell_n}}{\kappa}U_3(n)) \to \frac{2}{\sqrt{\pi}}$. Note that

$$\mathbb{E}\left(\frac{\sqrt{\ell_n}}{\kappa} U_3(n)\right) = \mathbb{E}\left(\left|\frac{\sqrt{\ell_n}}{\kappa} \tilde{S}'_1 - \frac{\sqrt{\ell_n}}{\kappa} \tilde{S}'_2\right|\right).$$

An application of the central limit theorem to the stationary β-mixing process $((Y'_i)^2 - 1)_{i \in \mathbb{N}}$ yields

$$\mathbb{E}\left(\left|\frac{\sqrt{\ell_n}}{\kappa} \tilde{S}'_1 - \frac{\sqrt{\ell_n}}{\kappa} \tilde{S}'_2\right|\right) \to \mathbb{E}(|Z - Z'|) = \frac{2}{\sqrt{\pi}}$$

and it suffices to additionally verify that $\mathrm{Var}(\sqrt{\ell_n}U_3(n)) \to 0$. By the definition of $U_3(n)$ and by the independence of the S'_j, we obtain

$$
\begin{aligned}
\mathrm{Var}(\sqrt{\ell_n}U_3(n)) ={}& \frac{1}{b_n(b_n-1)}\,\mathrm{Var}(|\sqrt{\ell_n}\tilde{S}'_1 - \sqrt{\ell_n}\tilde{S}'_2|) \\
&+ \frac{2}{b_n}\,\mathrm{Cov}(|\sqrt{\ell_n}\tilde{S}'_1 - \sqrt{\ell_n}\tilde{S}'_2|, |\sqrt{\ell_n}\tilde{S}'_1 - \sqrt{\ell_n}\tilde{S}'_3|).
\end{aligned}
$$

Now, one can check that the right-hand side converges to zero. Details of the proof are given in the Supplementary Material. $\square$

Theorem 2.4 is an immediate corollary of Proposition 3.5 and the above approximation steps.

3.5. *Outline of the proof of Theorem* 2.5. A standard tool from U-statistics theory is the Hoeffding decomposition of the kernel $h(x, y)$, given by

$$
h(x, y) = \theta^{(n)} + h_1^{(n)}(x) + h_1^{(n)}(y) + h_2^{(n)}(x, y),
$$

where

$$
\begin{aligned}
\theta^{(n)} &= \mathbb{E}(h(\sqrt{\ell_n}\tilde{S}'_1/\kappa, \sqrt{\ell_n}\tilde{S}'_2)/\kappa)), \\
h_1^{(n)}(x) &= \mathbb{E}(h(x, \sqrt{\ell_n}\tilde{S}'_1/\kappa)) - \theta^{(n)}, \\
h_2^{(n)}(x, y) &= h(x, y) - \theta^{(n)} - h_1^{(n)}(x) - h_1^{(n)}(y).
\end{aligned}
$$

Note that since we are dealing with a U-statistic of a triangular array, the decomposition depends upon the sample size n. Using the independence of $\tilde{S}'_j$ and Fubini's theorem, we obtain

$$
\mathbb{E}(h_1^{(n)}(\sqrt{\ell_n}\tilde{S}'_1/\kappa)) = 0,
$$
$$
\mathbb{E}(h_2^{(n)}(x, \sqrt{\ell_n}\tilde{S}'_1/\kappa)) = \mathbb{E}(h_2^{(n)}(\sqrt{\ell_n}\tilde{S}'_1/\kappa, y)) = 0
$$

for all $x, y \in \mathbb{R}$. Thus, the random variables $h_1^{(n)}(\sqrt{\ell_n}\tilde{S}'_1/\kappa)$ are independent and have mean 0. In addition, the kernels $h_2^{(n)}(x, y)$ are degenerate, and thus the random variables $h_2^{(n)}(\sqrt{\ell_n}\tilde{S}'_j)/\kappa, \sqrt{\ell_n}\tilde{S}'_k/\kappa)$ are pairwise uncorrelated. The Hoeffding decomposition of the kernel gives rise to the Hoeffding decomposition of $U_3(n)$,

$$
\begin{aligned}
\sqrt{b_n}\Big(\frac{\sqrt{\ell_n}}{\kappa}U_3(n) - \theta^{(n)}\Big) ={}& \frac{2}{\sqrt{b_n}}\sum_{j=1}^{b_n} h_1^{(n)}\Big(\sqrt{\ell_n}\frac{\tilde{S}'_j}{\kappa}\Big) \\
&+ \frac{\sqrt{b_n}}{b_n(b_n-1)}\sum_{1 \le j \ne k \le b_n} h_2^{(n)}\Big(\sqrt{\ell_n}\frac{\tilde{S}'_j}{\kappa}, \sqrt{\ell_n}\frac{\tilde{S}'_k}{\kappa}\Big).
\end{aligned}
$$

By the degeneracy of $h_2^{(n)}(x, y)$, we obtain

$$
\begin{aligned}
&\mathrm{Var}\Big(\frac{\sqrt{b_n}}{b_n(b_n-1)}\sum_{1 \le j \ne k \le b_n} h_2^{(n)}\Big(\sqrt{\ell_n}\frac{\tilde{S}'_j}{\kappa}, \sqrt{\ell_n}\frac{\tilde{S}'_k}{\kappa}\Big)\Big) \\
&= \frac{1}{b_n-1}\,\mathrm{Var}\Big(h_2^{(n)}\Big(\sqrt{\ell_n}\frac{\tilde{S}'_1}{\kappa}, \sqrt{\ell_n}\frac{\tilde{S}'_2}{\kappa}\Big)\Big),
\end{aligned}
$$

and some further calculations show that the right-hand side converges to 0. To handle the linear term in the Hoeffding decomposition, we can apply Lyapunov's central limit theorem for row-wise independent triangular arrays, and obtain convergence towards a normal law with mean 0 and variance

$$\psi^2 = 4 \lim_{n \to \infty} \operatorname{Var}\left(h_1^{(n)}\left(\sqrt{\ell_n} \frac{\tilde{S}_1'}{\kappa}\right)\right).$$

Since $\sqrt{\ell_n} \frac{\tilde{S}_1'}{\kappa} \to N(0, 1)$, one can prove $h_1^{(n)}(x) = \mathbb{E}(h(x, \sqrt{\ell_n} \frac{\tilde{S}_1'}{\kappa})) - \theta^{(n)} \to h_1(x) = \mathbb{E}(h(x, Z)) - \theta$ with $\theta = \mathbb{E}(h(Z, Z'))$ and the above variance equals $4 \operatorname{Var}(h_1(Z))$. In the end, we obtain the following central limit theorem for $U_3(n)$.

PROPOSITION 3.6. *Assume that $\ell_n = n^s$ with $s > 0.5$, $m_n = o(n^s \wedge n^{2s-1})$ and that there exist constants $\rho > 1 \vee \frac{9\delta}{(\delta+1)(\delta+2)}$ and $0 < \delta \leq 1$ such that $\mathbb{E}(|Y_1|^{4+2\delta}) < \infty$ and for all $k \in \mathbb{N}$, it holds $\beta_Y(k) \leq Ck^{-\rho(2+\delta)(1+\delta)/\delta^2}$. Then, we obtain under the null hypothesis*

$$\sqrt{n} \frac{U_3(n)}{\kappa} - \sqrt{b_n} \theta^{(n)} \xrightarrow{\mathcal{D}} \mathcal{N}(0, \psi^2),$$

where $\psi^2 = 4 \operatorname{Var}(h_1(Z))$, $h_1(x) = \mathbb{E}(|x - Z'|) - 2/\sqrt{\pi}$, where Z and Z' are two independent standard normal random variables.

Note that the centering $\theta^{(n)}$ likewise depends on the sample size, and is generally unknown. Thus, it is important to replace $\theta^{(n)}$ by its limit $\theta = 2/\sqrt{\pi}$.

LEMMA 3.7. *Assume that there exist constants $\rho > 1$ and $0 < \delta \leq 1$ such that $\mathbb{E}(|Y_1|^{4+2\delta}) < \infty$ and for all $k \in \mathbb{N}$ it holds $\beta_Y(k) \leq Ck^{-\rho(2+\delta)(1+\delta)/\delta^2}$. Moreover, let $\ell_n = n^s$ with $s > (1 + \delta \frac{\rho-1}{\rho+1})^{-1}$ and let $m_n = o(n^s)$. Then, under the null hypothesis, we have*

$$\sqrt{b_n}\left(\theta^{(n)} - \frac{2}{\sqrt{\pi}}\right) \to 0,$$

as $n \to \infty$.

The proof of Theorem 2.5 now follows from a combination of the former results. Proposition 3.1, Proposition 3.2, and Proposition 3.3 together imply

$$\sqrt{n}(U_3(n) - U(n)) \to 0.$$

Proposition 3.6 and Lemma 3.7 yield $\frac{\sqrt{n}}{\kappa} U_3(n) - \sqrt{b_n} \frac{2}{\sqrt{\pi}} \xrightarrow{\mathcal{D}} N(0, \psi^2)$. Thus, Theorem 2.5 is a consequence of Slutzky's lemma.

4. Extensions.

4.1. *Modifications of our test statistic.* Our test statistic corresponds to Gini's mean difference with entries given by the logarithms of local empirical variances. It explores the variability of estimated local variabilities. There are several possible extensions of this idea which might be useful not only for testing the constancy of the variance.

One possibility is the choice of other kernel functions $h : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ with suitable characteristics. In particular, we obtained similar limit results for the more general statistic

$$U(n) = \frac{1}{b_n(b_n - 1)} \sum_{1 \leq j \neq k \leq b_n} |\log(\hat{\sigma}_j^2) - \log(\hat{\sigma}_k^2)|^\alpha \quad \text{for } \alpha \in (0, 1]$$

but they come at the price of additional requirements regarding the growth conditions of the blocks, for instance $b_n/\ell_n^\alpha \to 0$. Moreover, additional simulations not reported here suggest that not much is gained by employing values other than $\alpha = 1$. The good performance obtained for $\alpha = 1$ can partly be explained by the rather large efficiency and better robustness of Gini's mean difference as compared to standard measures of variability like the standard deviation, see Gerstenberger, Vogel and Wendler [21] and the references cited therein.

Another modification is the evaluation of the variability of other local statistics, which shall be addressed in future work. More robust measures of scale as considered by Gerstenberger, Vogel and Wendler [21] in a setting with a single change point might replace the empirical variances. Using estimates of central location, kurtosis or tail behavior allows testing the stability of the corresponding characteristic. A hypothesis of particular interest is the assumption of the stability of second order characteristics.

REMARK 4.1. Davis, Huang and Yao [14] investigate likelihood ratio statistics for testing whether the time series can be described by a stable autoregressive process against the alternative of a single change. A change of the variance is not of direct interest but implicitly considered there since a change of the autoregressive parameters usually implies a change of the marginal variance if the variance of the innovations is constant. Within the somewhat restrictive parametric framework of Davis, Huang and Yao [14], one can apply a version of their test which allows for changing variances of the innovation process to check the assumption of a stable dependence structure underlying our test. If we are willing to accept the hypothesis of a stable autoregressive model after application of this preliminary test, testing the stability of the marginal variance using our approach is equivalent to testing the stability of the innovational variance.

Dette, Wu and Zhou [15, 16] investigate CUSUM-statistics for testing the constancy of second order characteristics, that is, of the variance or the correlation structure. They work on detrended data, allowing for a possibly time-varying mean under the null hypothesis, like we do. Given the complicated structure of their limiting process, they need bootstrap procedures for the calculation of critical values. Their test for the constancy of the correlation structure could be combined with our test, as outlined above for the test in Davis, Huang and Yao [14].

Tests for the null hypothesis of stable second order characteristics which are consistent not only against alternatives with a single change might be constructed by analyzing the variability of local variance and local correlation estimates jointly. Such multivariate extensions of the theory developed here are left to future work.

4.2. *Application to data with a discontinuous or a Hölder-continuous mean.* So far, we have assumed a time series setting with a Lipschitz-continuous mean function. This assumption might be too restrictive, for example, if structural breaks may occur not only in the variance but also in the mean. In such contexts, the hypothesis that the variance of the observed data $(X_i)_{i\in\mathbb{N}}$ is constant can be tested using the time series $(Z_i)_{i\in\mathbb{N}}$ of differences

$$Z_i = X_i - X_{i-1} = \sigma_H(Y_i - Y_{i-1}) + (\mu_i - \mu_{i-1}).$$

It inherits many properties from $(X_i)_{i\in\mathbb{N}}$ itself, like the existence of moments and the absolute regularity. Under the null hypothesis, the variance of the observations becomes

$$\mathrm{Var}(Z_i) = \sigma_H^2 \cdot (\mathrm{Var}(Y_i) + \mathrm{Var}(Y_{i-1}) + 2\,\mathrm{Cov}(Y_i, Y_{i-1}))$$

$$= 2\sigma_H^2 \cdot (\mathrm{Var}(Y_1) + \mathrm{Cov}(Y_1, Y_2)).$$

If the covariances are stationary, which comes along with the strict stationarity of $(Y_i)_{i\in\mathbb{N}}$ assumed in this entire work, changes in the variance of $(X_i)_{i\in\mathbb{N}}$ can thus be detected from

the sequence of differences $(Z_i)_{i \in \mathbb{N}}$. Moreover, the mean function of the differences not only preserves the Lipschitz-continuity and gets arbitrarily small, $|\mu_i - \mu_{i-1}| = O(1/n)$, but a close look at the proofs reveals that the results still hold in case of finitely many outliers in the mean (i.e., if μ is only piecewise Lipschitz-continuous on the intervals between finitely many jumps and we consider the time series $(Z_i)_{i \in \mathbb{N}}$). Hence, resorting to the differenced time series is advisable if abrupt changes in the mean are suspected.

EXAMPLE 4.2. Assume that we observe an increasing number of ℓ_n observations in each of an increasing number of b_n groups. One-way analysis of variance allows testing whether the group centers are identical, assuming all observations to be independent and identically distributed within each of the groups, with the same variance $\sigma_j^2 = \sigma^2$ for all groups $j = 1, \ldots, b_n$. The results derived in this paper allow testing the basic assumption of constant variance even if the observations within the different groups are dependent, as long as ℓ_n and b_n increase at appropriate rates. If the mean of the observations is not constant but only Lipschitz-continuous within each group, we can take differences of the observations within each group and proceed as described above. The condition of identical group sizes ℓ_n is restrictive but can presumably be relaxed as long as the group sizes $\ell_n^{(j)}$ increase at identical rates $\ell_n^{(j)} = O(\ell_n)$.

REMARK 4.3. A careful inspection of our proofs shows that one can moreover weaken the assumption of a Lipschitz-continuous mean function of $(X_i)_{i \in \mathbb{N}}$ to include a Hölder-continuous one with $|\mu_i - \mu_r| \leq L|i - r|/n^{s+1/4}$. This signifies that the overall variation of μ on the interval $[0, 1]$ is even allowed to grow at rate $n^{3/4-s}$. By reverting to the differenced time series $(Z_i)_{i \in \mathbb{N}}$, one is in addition able to deal with an increasing number of jumps $J(n)$ or a growing jump height $\Delta(n)$, as long as the condition $J(n)\Delta(n)^2 = o(n^{1/2})$ holds, and with a continuous part of the mean function fulfilling $|\mu_i - \mu_r| \leq L|i - r|/n^{1/4}$.

4.3. Estimation of the change-points and the variance function.

Another interesting issue is the estimation of the scale function $\sigma(\cdot)$ in case our test rejects the null hypothesis of σ being constant. If we assume the scale function to be piecewise constant, it is reasonable to estimate the change-point locations in advance and to calculate the empirical standard deviations of the observations in between subsequent change-points. Keeping in mind that our test is consistent also against smoothly varying alternatives, one could alternatively calculate local estimates using moving window techniques, possibly with an adaptive choice of the window width.

Under the assumption of a piecewise constant scale, we can adopt the approach from Wornowizki, Fried and Meintanis [31] as a first means to determine the number and location of the change-points. This approach is based on a recursive procedure which conducts in every step the change-point test described above. If the hypothesis is rejected, the dominant change point of the (sub-)sample is located by narrowing down the set of candidates to two blocks $B_{j^*} = \{(j^* - 1)\ell_n + 1, \ldots, j^*\ell_n\}$ and $B_{j^*+1} = \{j^*\ell_n + 1, \ldots, (j^* + 1)\ell_n\}$ via

$$j^* = \operatorname*{argmax}_{j \in \{1, \ldots, b_n-1\}} |\log(\hat{\sigma}_j^2) - \log(\hat{\sigma}_{j+1}^2)|.$$

Afterwards, the location of the change-point t^* is determined by

$$t^* = \operatorname*{argmax}_{t \in B_{j^*} \cup B_{j^*+1}} |\hat{\sigma}^2(X_{(j^*-1)\ell_n+1}, \ldots, X_t) - \hat{\sigma}^2(X_{t+1}, \ldots, X_{(j^*+1)\ell_n})|,$$

where $\hat{\sigma}^2(X_i, \ldots, X_k)$ for $i \leq k$ is the empirical variance based on the observations $X_i, \ldots, X_k$ and where we exclude values of t that are very close to the boundaries in order to

ensure a reliable estimation on both subsets $\{X_{(j^*-1)\ell_n+1}, \ldots, X_t\}$ and $\{X_{t+1}, \ldots, X_{(j^*+1)\ell_n}\}$. Once the change-point t^* is detected, we split the sample at t^* into two parts and the procedure is repeated on both subsamples. If the hypothesis can no longer be rejected, the procedure stops and it is assumed that all change points have been located. This procedure looks somewhat ad hoc and can probably be improved. Nevertheless, in our simulations it yielded similarly good results as the MOSUM approach described by Eichinger and Kirch [17], which in turn performed well as compared to several other state-of-the-art procedures in the simulations reported there.

5. Simulation study. We investigate the empirical size and power of the change-point test introduced in Section 2 and compare its performance to the procedure proposed in an earlier version [15] of Dette, Wu and Zhou [16]. Moreover, we discuss the performance of the long run variance estimator $\hat{\kappa}$ introduced in Section 2.3.

As data-generating processes, we consider two examples of independent observations, namely standard normal, $X_i \sim \mathcal{N}(0, 1)$, and exponential ones, $X_i \sim \text{Exp}(1)$, and four examples under dependence. The latter consist of two AR(1)-processes with $\alpha_1 = 0.4$ and $\alpha_1 = 0.7$, respectively, an ARMA(2, 2)-process

$$X_i = 0.8X_{i-1} - 0.4X_{i-2} + \varepsilon_i + 0.5\varepsilon_{i-1} + 0.34\varepsilon_{i-2}$$

and a GARCH(1, 1)-process

$$X_i = \sigma_i \varepsilon_i \quad \text{with } \sigma_i^2 = 0.1 + 0.1X_{i-1}^2 + 0.8\sigma_{i-1}^2,$$

each with independent standard normal innovations $(\varepsilon_i)_{i \in \mathbb{N}}$. The parameter choice for the GARCH-process is as in Andreou and Ghysels [3] and describes a high volatility persistence. ARMA- and GARCH-processes are, as mentioned in Example 2.2, mixing at an exponential rate. Moreover, all models considered possess finite sixth moments (see Theorem 5 in Lindner [25] for the GARCH-case). Based on extensive simulations not reported here, we recommend setting the tuning parameters to $s = 0.7$ and $q = 0.5$ with $\ell_n = n^s$ and $\tilde{\ell}_n = n^q$. Note that this choice is in line with all restrictions on the block lengths imposed by our asymptotic theory.

We investigate the empirical power of the tests under various (local) alternatives listed below. These include scenarios with one, two and four structural breaks, as well as a smoothly changing variance function, each with variance changes of magnitude $n^{-1/2}$.

$$\mathbb{A}1: \quad \sigma(x) = 1 \cdot \mathbb{1}_{\{0 \leq x < 1/2\}} + (1 + 0.2\sqrt{2000/n}) \cdot \mathbb{1}_{\{1/2 \leq x \leq 1\}},$$

$$\mathbb{A}2: \quad \sigma(x) = 1 \cdot \mathbb{1}_{\{0 \leq x < 1/3\}} + (1 + 0.2\sqrt{2000/n}) \cdot \mathbb{1}_{\{1/3 \leq x < 2/3\}} + 1 \cdot \mathbb{1}_{\{2/3 \leq x \leq 1\}},$$

$$\mathbb{A}3: \quad \sigma(x) = 1 \cdot \mathbb{1}_{\{0 \leq x < 1/5\}} + (1 + 0.2\sqrt{2000/n}) \cdot \mathbb{1}_{\{1/5 \leq x < 2/5\}} + 1 \cdot \mathbb{1}_{\{2/5 \leq x < 3/5\}}$$

$$+ (1 + 0.2\sqrt{2000/n}) \cdot \mathbb{1}_{\{3/5 \leq x < 4/5\}} + 1 \cdot \mathbb{1}_{\{4/5 \leq x \leq 1\}},$$

$$\mathbb{A}4: \quad \sigma(x) = 1 + 0.1 \cdot \sin(4\pi x) \cdot \sqrt{2000/n}.$$

All simulations are conducted in R [27]. The long run variance $\hat{\kappa}$ is estimated as described in Section 2.3, whereas the centering term $\mathbb{E}(|Z - Z'|) = 2/\sqrt{\pi}$ and the variance of the limit distribution $\psi^2 = 4/3 + 8/\sqrt{\pi}(\sqrt{3} - 2)$ can be calculated explicitly. All results are obtained for a nominal significance level of $\alpha = 5\%$. The analysis in Sections 5.1–5.3 is conducted for centered observations. We investigate the influence of nonconstant mean functions on our test results in Section 5.4 and on the performance of the long run variance estimator $\hat{\kappa}$ in Section 5.5.

5.1. *Comparison to the procedure of Dette*, *Wu and Zhou* [15, 16]. This section compares our test (henceforth denoted by SWFD) to the results of Dette, Wu and Zhou [15, 16] (denoted by DWZ), who employ a CUSUM-approach to detect change points in the variance. Since the asymptotic distribution of their test statistic crucially depends upon the dependence structure of the underlying time series, they use a bootstrap approach to obtain critical values.

The comparison is limited to the sample lengths $n = 500, 2000$ and the simulated rejection probabilities are based on only 4000 replications each, due to the computational cost of the DWZ procedure. As a rough comparison of the computation times of the respective procedures, we measured the overall computation time (on a 3.4 GHz Intel Core i7) required to obtain the results in Table 1, that is, the time required for $4000 \cdot 30$ executions of the respective test (including the time for the simulation of the data sets). For $n = 500$ ($n = 2000$), the SWFD procedure needed on average 0.00064 (0.00138) seconds per execution, while the DWZ procedure took 0.88813 (7.73488) seconds.

Table 1 shows the simulated rejection probabilities of the SWFD and the DWZ procedure under the hypothesis and the local alternatives $\mathbb{A}1$–$\mathbb{A}4$. Our SWFD test performs anti-conservative for $n = 500$ as we observe empirical sizes of about 10% or even higher instead of the nominal 5% level. For $n = 2000$, the empirical sizes fall below 10%, except for the GARCH(1, 1). A closer analysis of the SWFD procedure will be presented in the next subsection. For the DWZ test, we concentrate on the two sample sizes because of its much larger computation times. The size of the DWZ procedure is adequate for independent normal data but conservative for the exponential data and anti-conservative in the presence of higher positive dependencies. In case of the GARCH(1, 1), it performs worse than the SWFD test.

TABLE 1
Simulated rejection probabilities of the SWFD and DWZ test at the significance level $\alpha = 0.05$ for the sample sizes $n = 500, 2000$ under the null hypothesis $\mathbb{H}$ and various local alternatives $\mathbb{A}1$ to $\mathbb{A}4$ with effect sizes of magnitude $n^{-1/2}$ and for different data-generating processes. For the results under the alternatives, a size-correction has been conducted

		$N(0, 1)$	Exp(1)	AR(1), 0.4	AR(1), 0.7	ARMA(2, 2)	GARCH(1, 1)
				$n = 500$			
$\mathbb{H}$	SWFD	0.085	0.112	0.098	0.134	0.106	0.180
	DWZ	0.052	0.029	0.064	0.088	0.073	0.290
$\mathbb{A}1$	SWFD	0.734	0.318	0.630	0.356	0.400	0.336
	DWZ	0.999	0.649	0.980	0.824	0.836	0.680
$\mathbb{A}2$	SWFD	0.457	0.186	0.376	0.200	0.253	0.178
	DWZ	0.618	0.159	0.492	0.258	0.252	0.134
$\mathbb{A}3$	SWFD	0.221	0.126	0.202	0.128	0.154	0.114
	DWZ	0.088	0.047	0.069	0.044	0.061	0.042
$\mathbb{A}4$	SWFD	0.481	0.194	0.399	0.222	0.255	0.199
	DWZ	0.372	0.118	0.274	0.144	0.160	0.096
				$n = 2000$			
$\mathbb{H}$	SWFD	0.073	0.091	0.074	0.096	0.084	0.148
	DWZ	0.052	0.039	0.058	0.108	0.074	0.394
$\mathbb{A}1$	SWFD	0.891	0.344	0.784	0.439	0.505	0.375
	DWZ	1	0.760	0.995	0.875	0.920	0.727
$\mathbb{A}2$	SWFD	0.734	0.246	0.610	0.310	0.343	0.251
	DWZ	0.869	0.239	0.748	0.382	0.446	0.228
$\mathbb{A}3$	SWFD	0.805	0.269	0.674	0.368	0.397	0.294
	DWZ	0.318	0.066	0.261	0.127	0.140	0.070
$\mathbb{A}4$	SWFD	0.644	0.185	0.504	0.256	0.279	0.218
	DWZ	0.516	0.136	0.392	0.192	0.248	0.138

To achieve a fair comparison, we report size-corrected empirical powers under the four alternatives, using the empirical 95% percentile of the test results for the same distribution and the same sample size as critical values (see Table 1 in the Supplementary Material ([28], Supplement B) for the rejection rates at the nominal (asymptotical) 5%-level). While the DWZ procedure obtains higher empirical power in case of one or few structural breaks ($\mathbb{A}1$ and $\mathbb{A}2$), the SWFD does so in case of multiple breaks and for the sine function ($\mathbb{A}3$ and $\mathbb{A}4$). Moreover, both tests show difficulties in coping with models of strong dependence, especially the GARCH(1, 1)-model, as well as with exponentially distributed observations.

5.2. *Analysis of the empirical size.*   In the next subsections, we will analyse the performance of the SWFD-test in more detail. Throughout, we will base our results on 6000 replications. Figure 1 depicts the empirical size as a function of the sample length, where we used $n = 500, 1000, 2000, 3000, 4000, 5000, 8000, 12{,}000$ and $16{,}000$.

The empirical size is about 10% or even larger in case of moderately large samples ($n = 500$). It approaches the nominal significance level alpha $= 5\%$ as the sample size increases, though the test stays liberal in case of the scenarios considered here. Moreover, the dependence as well as the distribution of the data seem to be important. The empirical sizes are smaller if the dependences among the observations are low ($N(0, 1)$, AR(1) with parameter $\alpha_1 = 0.4$), while stronger dependences as in the GARCH(1, 1) process lead to much higher rejection rates. The rejection rates are also rather large in case of the exponentially distributed observations, despite their independence.

5.3. *Analysis of the empirical power.*   Figure 2 shows the empirical power at the nominal significance level $\alpha = 0.05$ as a function of the sample size separately for each of the local alternatives $\mathbb{A}1$–$\mathbb{A}4$ and for the same choices of n as considered before (see Figure 1 in Supplement B for the size-corrected version). All simulated rejection probabilities are based on 6000 replications. Across all alternatives, the graphs for the different underlying distributions show the same order, which is almost a reversed image of the order in Figure 1. The highest empirical power is obtained for the distributions with the lowest dependence ($N(0, 1)$, AR(1) with parameter $\alpha_1 = 0.4$), while all other processes lead to somewhat smaller rejection rates. The performance for small sample sizes is significantly worse compared to the empirical
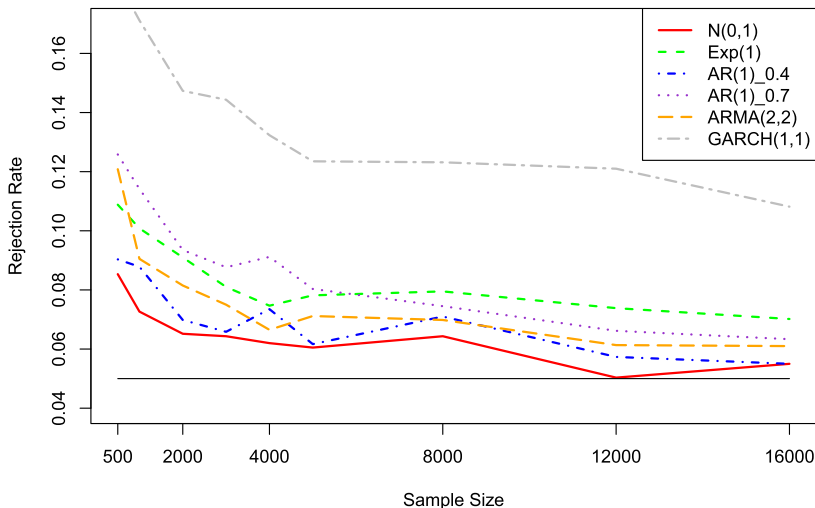

FIG. 1.   *Empirical rejection rates of the SWFD test under the null hypothesis as a function of the sample size for different distributions of the data-generating process. The nominal significance level $\alpha = 0.05$ is represented by the solid black line.*
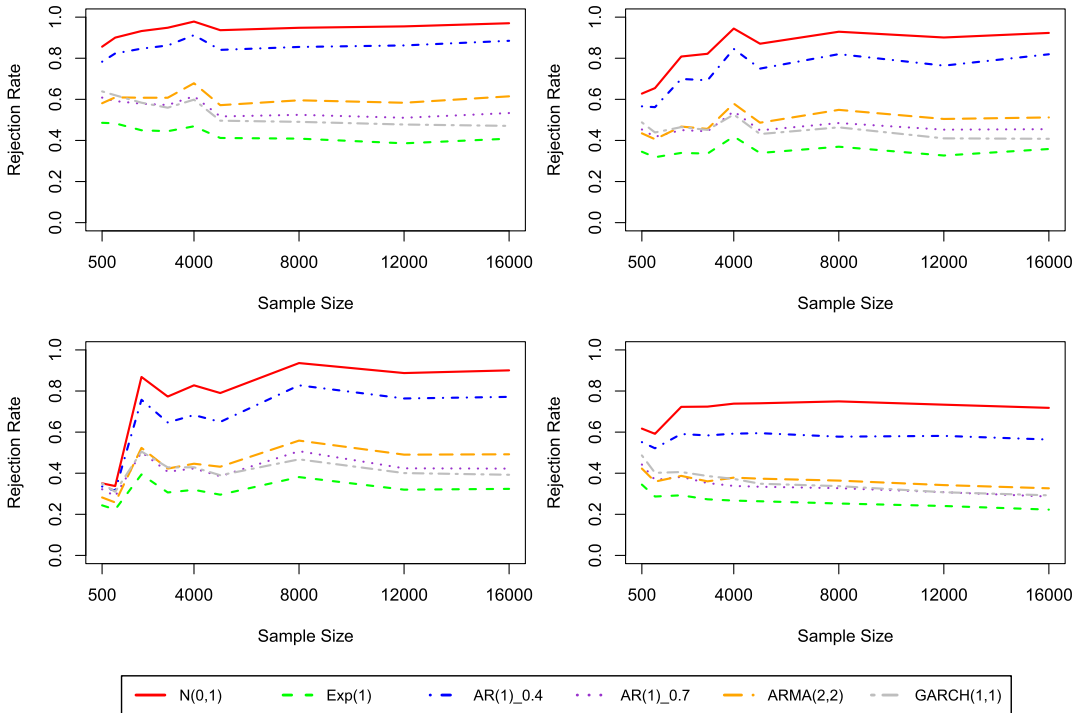
FIG. 2. *Empirical power of the SWFD test for the local alternatives $\mathbb{A}1$(top left)-$\mathbb{A}4$(bottom right) with effect sizes of magnitude $n^{-1/2}$ as a function of the sample size n at the nominal significance level $\alpha = 0.05$.*

power for larger $n \geq 2000$, but once an appropriate sample length is reached, the rejection rates stabilize at a certain level. Besides, the partitioning into blocks of length $\ell_n = n^{0.7}$ might have a certain influence, being less ideal for some n with respect to the location of the break points. As an example, observe the peak at $n = 4000$ under alternative $\mathbb{A}2$ (top right). The partition there is close to ideal due to $(4000^{0.7}) \cdot 4 \approx 1328$ while the break points are located at $4000/3 \approx 1333$ and at 2666.

5.4. *Performance for noncentred data.* Our limit theory likewise holds for time series that are not stationary in the mean. Table 2 shows the simulated rejection rates at the nominal significance level $\alpha = 0.05$ for samples of length $n = 3000$ and for the mean functions $\mu(x) = x$, $\mu(x) = \sin(2\pi x)$ and $\mu(x) = 0 \cdot \mathbb{1}_{\{0 \leq x < 1/2\}} + 1 \cdot \mathbb{1}_{\{1/2 \leq x \leq 1\}}$ (see Table 3 in Supplement B for the size-corrected version). The theory developed here requires that our test is applied to data with a Lipschitz-continuous mean function, implying that jumps in the mean cause problems. As pointed out in Section 4, this issue can be resolved by analysing the differenced time series $(Z_i)_{i \in \mathbb{N}}$, $Z_i = X_i - X_{i-1}$. The results for the last mean function demonstrate that this approach is indeed also effective in practice.

5.5. *Performance of the long run variance estimator $\hat{\kappa}$.* We briefly discuss the performance of the estimator $\hat{\kappa}$ introduced in Section 2.3 on the basis of its empirical bias and root mean square error (RMSE). To facilitate a comparison, we standardize the data-generating processes introduced above to yield a theoretical long run variance of 1. All results are based on 6000 replications. Table 3 shows the results for the sample lengths $n = 500, 1000, 3000$ and for the mean functions $\mu(x) = 0$, $\mu(x) = \sin(2\pi x)$ and $\mu(x) = \mathbb{1}_{\{1/2 < x \leq 1\}}$. Note that for the latter mean function, we consider the differenced time series $(Z_i)_{i \in \mathbb{N}}$. The dependence structure and underlying distribution seem to influence the quality of the estimate, with high

TABLE 2

Simulated rejection probabilities of the test SWFD at the nominal significance level $\alpha = 0.05$ for the sample size $n = 3000$ under the hypothesis and various local alternatives with effect sizes of magnitude $n^{-1/2}$. Results for different data-generating processes and different mean functions μ are given

	$N(0,1)$	Exp(1)	AR(1), 0.4	AR(1), 0.7	ARMA(2, 2)	GARCH(1, 1)
			$\mu(x) = x$			
$\mathbb{H}$	0.062	0.081	0.066	0.085	0.076	0.134
$\mathbb{A}1$	0.954	0.441	0.863	0.578	0.619	0.561
$\mathbb{A}2$	0.822	0.330	0.703	0.426	0.463	0.464
$\mathbb{A}3$	0.783	0.312	0.639	0.400	0.420	0.418
$\mathbb{A}4$	0.723	0.267	0.582	0.354	0.368	0.387
			$\mu(x) = \sin(2\pi x)$			
$\mathbb{H}$	0.062	0.086	0.077	0.094	0.070	0.148
$\mathbb{A}1$	0.940	0.451	0.854	0.579	0.607	0.569
$\mathbb{A}2$	0.835	0.360	0.697	0.436	0.466	0.457
$\mathbb{A}3$	0.696	0.278	0.576	0.382	0.409	0.383
$\mathbb{A}4$	0.722	0.267	0.574	0.360	0.370	0.367
			$\mu(x) = 0 \cdot \mathbb{1}_{\{0 \leq x < 1/2\}} + 1 \cdot \mathbb{1}_{\{1/2 \leq x \leq 1\}}$ for time series $(Z_i)_{i \in \mathbb{N}}$			
$\mathbb{H}$	0.067	0.091	0.064	0.062	0.079	0.124
$\mathbb{A}1$	0.814	0.390	0.889	0.924	0.664	0.491
$\mathbb{A}2$	0.648	0.306	0.737	0.783	0.512	0.391
$\mathbb{A}3$	0.587	0.279	0.692	0.733	0.454	0.369
$\mathbb{A}4$	0.538	0.234	0.628	0.687	0.405	0.340

dependence leading to a larger bias and a higher RMSE. The empirical bias is usually negative, thereby providing an explanation for the liberal behaviour of our test in Figure 1. In contrast, the mean function seems to have little influence. These results suggest that an improved estimator $\hat{\kappa}$ might yield even better results for our test procedure. However, the task of

TABLE 3

Simulated Bias and RMSE of $\hat{\kappa}$ for the mean functions $\mu(x) = 0$ (top row for each n), $\mu(x) = \sin(2\pi x)$ and $\mu(x) = \mathbb{1}_{\{1/2 < x \leq 1\}}$ (bottom row for each n) for different sample lengths n, $q = 0.5$ and $s = 0.7$. For the last mean function, the time series $(Z_i)_{i \in \mathbb{N}}$ is used. The observations are standardized to yield a theoretical long run variance $\kappa_{Y^2} = 1$

$N(0,1)$		Exp(1)		AR(1), 0.4		AR(1), 0.7		ARMA(2, 2)		GARCH(1, 1)	
Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
					$n = 500$						
−0.045	0.172	−0.130	0.261	−0.089	0.192	−0.199	0.267	−0.096	0.200	−0.285	0.350
−0.006	0.174	−0.127	0.255	−0.054	0.185	−0.183	0.257	−0.089	0.198	−0.275	0.344
−0.017	0.179	−0.120	0.256	−0.030	0.175	−0.033	0.173	−0.087	0.200	−0.273	0.344
					$n = 1000$						
−0.029	0.147	−0.101	0.206	−0.065	0.160	−0.137	0.209	−0.063	0.163	−0.219	0.287
−0.004	0.147	−0.101	0.209	−0.035	0.155	−0.129	0.201	−0.063	0.165	−0.219	0.283
−0.001	0.154	−0.091	0.204	−0.024	0.150	−0.024	0.149	−0.059	0.168	−0.223	0.289
					$n = 3000$						
−0.014	0.106	−0.064	0.141	−0.032	0.108	−0.075	0.132	−0.035	0.111	−0.145	0.195
−0.002	0.106	−0.060	0.137	−0.016	0.107	−0.066	0.131	−0.032	0.114	−0.140	0.192
0.012	0.110	−0.051	0.135	−0.012	0.105	−0.012	0.106	−0.033	0.115	−0.152	0.198

finding a suitable estimator is rather intricate since we only observe noncentred data $(X_i)_{i \in \mathbb{N}}$ and additionally require a convergence rate of $\sqrt{b_n}$.

6. Data example. As a data example, we consider the worldwide relative search interest for the topic "global warming" retrieved from Google Trends (www.google.com/trends). Data from Google Trends has frequently been used in environmental research as a measure for public interest; see, for instance, Anderegg and Goldsmith [2] and Burivalova, Butler and Wilcove [10]. Google Trends does not provide the absolute search volume but adjusts the data in two steps: Each time period, the number of searches for the topic "global warming" as a proportion of the total searches within that time period is calculated. Afterwards, these proportions of the total searches are scaled, with the time period of the highest proportion being assigned the value 100 (for details, see the Google Trends FAQs (https://support.google.com/trends/answer/4365533?hl=en&ref_topic=6248052).

Figure 3 shows the weakly worldwide relative search interest from January 2004 to December 2019. The increase in relative search interest between 2006/2007 and 2010/2011 is often related to the release of the documentary film "An Inconvenient Truth" in 2006 as well as the two media events colloquially often referred to as "climategate" and "glaciergate" in November 2009 and January 2010, respectively (see [2] and [10]). While a rise in the mean can be interpreted as a general rise of the public interest in the topic, an increase in the variance might hint at frequent news publication and an increased media coverage causing spikes and overall more fluctuation in the search interest.

To eliminate seasonal effects, we work with the time series of annual differences $(Z_i)_{i \in \mathbb{N}}$, that is, $Z_i = X_i - X_{i-52}$, starting in January 2005, and thus consider changes in the relative search interest. After eliminating the last week in all years with 53 weeks (2006, 2012 and 2017) in the original time series X, we are left with $n = 780$ differences. In the following, we test for a stationary variance in the series of annual differences.

As we find our test to behave rather liberal for small sample sizes in the simulation study in Section 5.2, we consider not only the usual significance level $\alpha = 5\%$ but also $\alpha = 1\%$. Figure 4 shows the seasonally differenced observations $Z_1, \ldots, Z_{780}$ as well as the detected changes in the variance for $\alpha = 5\%$, which are estimated as described in Section 4.3 and located at the weeks of the 18.12.2005, 24.12.2006, 18.09.2011 and 08.11.2015. The estimated change points seem to capture the periods of increased variance of $(Z_i)_{i \in \mathbb{N}}$ quite well, thereby not only indicating changes during the time period of increased interest mentioned above, but additionally detecting a change in late 2015. For $\alpha = 1\%$, we detect the same except for the last change point.
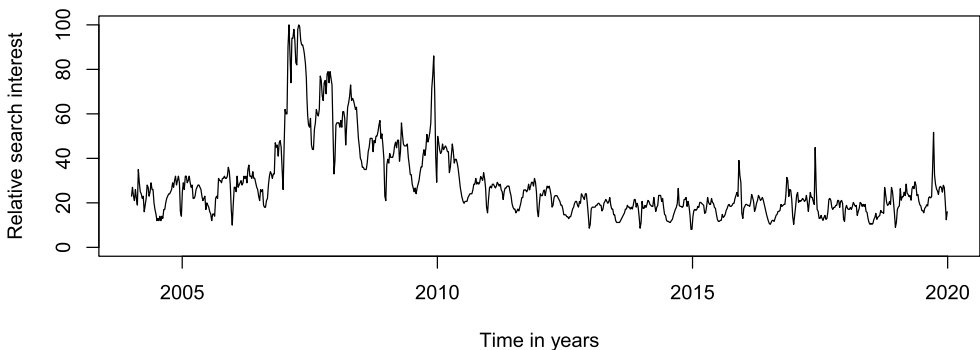

FIG. 3. *Weekly worldwide relative search interest for the topic "global warming" from January* 2004 *to December* 2019 *obtained from Google Trends.*
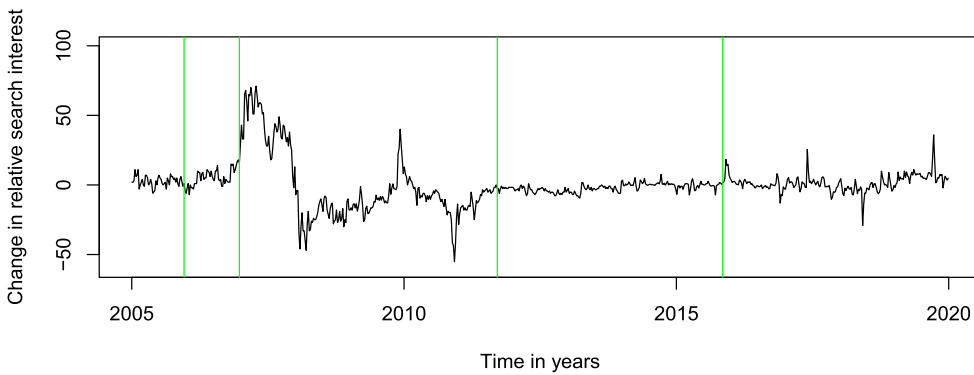
FIG. 4.   *Change in the weekly worldwide relative search interest for the topic "global warming" from January* 2005 *to December* 2019 *obtained from Google Trends. Detected variance change points for α = 5% are marked by the vertical green lines.*

Moreover, a close look at the data hints at a possibly nonstationary mean, with which our procedure seems to cope quite well due to the centering via the block means. A cumulative sum test for a change in mean does not reject the null hypothesis at $\alpha = 1\%$, while at $\alpha = 5\%$, it sequentially detects a very large number of 11 changes when being combined with binary segmentation. There is hence considerable uncertainty concerning mean stationarity when applying such a standard test. In contrast, our test consistently indicates a nonstationary variance, irrespective of the chosen significance level.

SUPPLEMENTARY MATERIAL

Supplement A: Proofs (DOI: 10.1214/21-AOS2092SUPPA; .pdf). A supplement containing detailed proofs and additional results from the literature.

Supplement B: Additional simulation results (DOI: 10.1214/21-AOS2092SUPPB; .pdf). A supplement containing additional simulation results from Section 5.

REFERENCES

[1] ABRAHAM, B. and WEI, W. W. S. (1984). Inferences about the parameters of a time series model with changing variance. *Metrika* **31** 183–194. MR0754960 https://doi.org/10.1007/BF01915199

[2] ANDEREGG, W. R. L. and GOLDSMITH, G. R. (2014). Public interest in climate change over the past decade and the effects of the 'climategate' media event. *Environ. Res. Lett.* **9** 54005.

[3] ANDREOU, E. and GHYSELS, E. (2002). Detecting multiple breaks in financial market volatility dynamics. *J. Appl. Econometrics* **17** 579–600.

[4] AUE, A., HÖRMANN, S., HORVÁTH, L. and REIMHERR, M. (2009). Break detection in the co-variance structure of multivariate time series models. *Ann. Statist.* **37** 4046–4087. MR2572452 https://doi.org/10.1214/09-AOS707

[5] BAUFAYS, P. and RASSON, J.-P. (1985). Variance changes in autoregressive model. In *Time Series Analysis*: *Theory and Practice*, 7 (*Toronto, Ont.*, 1983) 119–128. North-Holland, Amsterdam. MR0787816

[6] BLOOMFIELD, P., HURD, H. L. and LUND, R. B. (1994). Periodic correlation in stratospheric ozone data. *J. Time Series Anal.* **15** 127–150. MR1263886 https://doi.org/10.1111/j.1467-9892.1994.tb00181.x

[7] BOROVKOVA, S., BURTON, R. and DEHLING, H. (2001). Limit theorems for functionals of mixing processes with applications to *U*-statistics and dimension estimation. *Trans. Amer. Math. Soc.* **353** 4261–4318. MR1851171 https://doi.org/10.1090/S0002-9947-01-02819-7

[8] BRADLEY, R. C. (2005). Basic properties of strong mixing conditions. A survey and some open questions. *Probab. Surv.* **2** 107–144. MR2178042 https://doi.org/10.1214/154957805100000104

[9] BRADLEY, R. C. (2007). *Introduction to Strong Mixing Conditions. Vol.* 2. Kendrick Press, Heber City, UT. MR2325295

[10] BURIVALOVA, Z., BUTLER, R. A. and WILCOVE, D. S. (2018). Analyzing Google search data to debunk myths about the public's interest in conservation. *Front. Ecol. Environ.* **16** 509–514.

[11] CARLSTEIN, E. (1986). The use of subseries values for estimating the variance of a general statistic from a stationary sequence. *Ann. Statist.* **14** 1171–1179. MR0856813 https://doi.org/10.1214/aos/1176350057

[12] CHEN, J. and GUPTA, A. K. (1997). Testing and locating variance changepoints with application to stock prices. *J. Amer. Statist. Assoc.* **92** 739–747. MR1467863 https://doi.org/10.2307/2965722

[13] DAHLHAUS, R. (1997). Fitting time series models to nonstationary processes. *Ann. Statist.* **25** 1–37. MR1429916 https://doi.org/10.1214/aos/1034276620

[14] DAVIS, R. A., HUANG, D. W. and YAO, Y.-C. (1995). Testing for a change in the parameter values and order of an autoregressive model. *Ann. Statist.* **23** 282–304. MR1331669 https://doi.org/10.1214/aos/1176324468

[15] DETTE, H., WU, W. and ZHOU, Z. (2015). Change point analysis of second order characteristics in nonstationary time series. Available at arXiv:1503.08610.

[16] DETTE, H., WU, W. and ZHOU, Z. (2019). Change point analysis of correlation in non-stationary time series. *Statist. Sinica* **29** 611–643. MR3931381

[17] EICHINGER, B. and KIRCH, C. (2018). A MOSUM procedure for the estimation of multiple random change points. *Bernoulli* **24** 526–564. MR3706768 https://doi.org/10.3150/16-BEJ887

[18] GALEANO, P. and PEÑA, D. (2007). Covariance changes detection in multivariate time series. *J. Statist. Plann. Inference* **137** 194–211. MR2292851 https://doi.org/10.1016/j.jspi.2005.09.003

[19] GAO, Z., SHANG, Z., DU, P. and ROBERTSON, J. L. (2019). Variance change point detection under a smoothly-changing mean trend with application to liver procurement. *J. Amer. Statist. Assoc.* **114** 773–781. MR3963179 https://doi.org/10.1080/01621459.2018.1442341

[20] GERSTENBERGER, C. and VOGEL, D. (2015). On the efficiency of Gini's mean difference. *Stat. Methods Appl.* **24** 569–596. MR3421674 https://doi.org/10.1007/s10260-015-0315-x

[21] GERSTENBERGER, C., VOGEL, D. and WENDLER, M. (2020). Tests for scale changes based on pairwise differences. *J. Amer. Statist. Assoc.* **115** 1336–1348. MR4143469 https://doi.org/10.1080/01621459.2019.1629938

[22] GOMBAY, E., HORVÁTH, L. and HUŠKOVÁ, M. (1996). Estimators and tests for change in variances. *Statist. Decisions* **14** 145–159. MR1406521

[23] INCLÁN, C. and TIAO, G. C. (1994). Use of cumulative sums of squares for retrospective detection of changes of variance. *J. Amer. Statist. Assoc.* **89** 913–923. MR1294735

[24] LEE, S. and PARK, S. (2001). The cusum of squares test for scale changes in infinite order moving average processes. *Scand. J. Stat.* **28** 625–644. MR1876504 https://doi.org/10.1111/1467-9469.00259

[25] LINDNER, A. M. (2009). Stationarity, mixing, distributional properties and moments of GARCH(p, q)-processes. In *Handbook of Financial Time Series* 43–69. Springer, Berlin.

[26] MOKKADEM, A. (1988). Mixing properties of ARMA processes. *Stochastic Process. Appl.* **29** 309–315. MR0958507 https://doi.org/10.1016/0304-4149(88)90045-2

[27] R CORE TEAM (2019). *R: a Language and Environment for Statistical Computing*. Vienna, Austria.

[28] SCHMIDT, S. K., WORNOWIZKI, M., FRIED, R. and DEHLING, H. (2021). Supplement to "An asymptotic test for constancy of the variance under short-range dependence." https://doi.org/10.1214/21-AOS2092SUPPA, https://doi.org/10.1214/21-AOS2092SUPPB

[29] WICHERN, D. W., MILLER, R. B. and HSU, D.-A. (1976). Changes of variance in first-order autoregressive time series models—with an application. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **25** 248–256.

[30] WIED, D., ARNOLD, M., BISSANTZ, N. and ZIGGEL, D. (2012). A new fluctuation test for constant variances with applications to finance. *Metrika* **75** 1111–1127. MR2989322 https://doi.org/10.1007/s00184-011-0371-7

[31] WORNOWIZKI, M., FRIED, R. and MEINTANIS, S. G. (2017). Fourier methods for analyzing piecewise constant volatilities. *AStA Adv. Stat. Anal.* **101** 289–308. MR3679347 https://doi.org/10.1007/s10182-017-0288-1

[32] WU, W. B. and ZHAO, Z. (2007). Inference of trends in time series. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **69** 391–410. MR2323759 https://doi.org/10.1111/j.1467-9868.2007.00594.x