SFB 823

An asymptotic test on the stationarity of the variance

Herold Dehling, Roland Fried,
Max Wornowizki

Discussion Paper

Nr. 71/2016

November 8, 2016

AN ASYMPTOTIC TEST ON THE STATIONARITY OF THE VARIANCE

HEROLD DEHLING, ROLAND FRIED AND MAX WORNOWIZKI

ABSTRACT. We reconsider a statistic introduced in Wornowizki et al. (2016) allowing to test the stationarity of the variance for a sequence of independent random variables. Instead of determining rejection regions via the permutation principle as proposed before, we provide asymptotic critical values leading to huge savings in computation time. To prove the required limit theorems, the test statistic is viewed as a U-statistic constructed from blockwise variance estimates. Since the distribution of the test statistic depends on the sample size, a suitable new law of large numbers as well as a central limit theorem are developed. These asymptotic results are illustrated on artificial data. The permutation and asymptotic version of the test are compared to alternative procedures in extensive Monte Carlo experiments. The simulation results suggest that the methods offer similar results and high power when compared to their competitors, particularly in the case of multiple structural breaks. They also estimate the structural break positions adequately.

1. Introduction

Consider a sequence of random variables $X_1, \ldots, X_n, n \in \mathbb{N}$, with existing variances denoted by $\sigma_i^2 = Var(X_i), i = 1, \ldots, n$. In this work, we propose an asymptotic procedure testing the constancy of the variance. We investigate the hypotheses

$$\mathbb{H}_0 : \forall i \in \{1, \ldots, n-1\} : \sigma_i^2 = \sigma_{i+1}^2 \text{ vs. } \mathbb{H}_1 : \exists i \in \{1, \ldots, n-1\} : \sigma_i^2 \neq \sigma_{i+1}^2.$$

The task of detecting changes in the volatility process received a lot of attention, particularly in the last forty years. Early approaches to derive appropriate decision rules often relied on distributional assumptions. Examples are among others Hsu (1977), using cumulative sums of χ^2-type random variables under the assumption of Gaussianity, as well as the likelihood-based procedures presented in Chen & Gupta (1997) and Jandhyala et al. (2002). More recent papers weakened the distributional assumptions proposing asymptotical CUSUM-type tests and nonparametric procedures, see for example Wied et al. (2012) and Ross (2013). Keep in mind that most such methods compute their test statistic by splitting the sample into two blocks of consecutive observations. They are constructed to detect at most one change at a time and typically perform well in such cases. However, in case of several variance changes they can lose a considerable amount of power.

The statistic considered in the following tries to circumvent this problem by splitting the data into several blocks. It is based on a Fourier-type transformation of the blockwise variance estimates along with the weighting scheme proposed in Matteson & James (2014). The corresponding permutation test investigated in Wornowizki et al. (2016) attains competitive rejection rates in the case of one volatility change and outperforms classical procedures in case of multiple change points. In addition, it does not make any assumptions on the distribution of the data and keeps the significance level for any sample size. However, its high computational costs limit its applicability for large sample sizes.

Key words and phrases. Change point analysis, variance, piecewise identical distribution and U-statistic.

The authors were supported in part by the Collaborative Research Center 823, Project C3 *Analysis of Structural Change in Dynamic Processes*, of the German Research Foundation.

In this paper, we resolve this issue by developing asymptotic theory for the statistic investigated with the permutation test. This allows us to construct an asymptotic test with the same advantages with regard to change point detection as its permutation version but at much lower computational cost. In analogy to Wornowizki et al. (2016), we adopt the framework of locally stationary, but globally non-constant variance considered, among others, in Mercurio & Spokoiny (2004), Stărică & Granger (2005), Vassiliou & Demetriou (2005), Spokoiny (2009), Davies et al. (2012) and Fried (2012). In particular, for the random variables $X_1, \ldots, X_n$ we mainly require independence, identical distribution under $\mathbb{H}_0$ and existing fourth moments. A more detailled discussion on our assumptions is given at the beginning of Section 2.
To construct the test statistic under study presented, we divide the indices $1, \ldots, n$ into N consecutive, equally sized blocks. These blocks, denoted by $B_{N,1}, \ldots, B_{N,N}$, have common length $\tau_N = n/N$, which is for simplicity assumed to be an integer. Thus, block $B_{N,j}$ consists of the indices $(j-1)\,\tau_N + 1, \ldots j\,\tau_N$ for $j = 1, \ldots, N$. The corresponding empirical variance, calculated by $\hat{\sigma}_j^2 = \frac{1}{\tau_N}\sum_{i \in B_{N,j}} X_i^2$ for $j = 1, \ldots, N$, already takes into account the zero mean. The blockwise variance estimates are then compared to one another via the statistic

$$T_N = \frac{1}{N(N-1)} \sum_{1 \leq j \neq k \leq N} \left|\log\left(\hat{\sigma}_j^2\right) - \log\left(\hat{\sigma}_k^2\right)\right|^a$$

for some $a \in (0, 2)$. Note that the case $a = 1$ corresponds to Gini's mean difference applied to the logarithmised variance estimates. A high value of T_N reflects potential changes in the variance process and will therefore lead to a rejection of the corresponding test.

The paper is structured as follows: In Section 2 we formulate and prove a suitable law of large numbers (LLN) as well as a central limit theorem (CLT) for T_N along with some auxiliary lemmas. These results are supported by the simulations presented in Section 3. In particular, we illustrate that the size of the asymptotic test for the stationarity of the variance converges to the predefined significance level. Section 4 investigates the power of this test under various alternatives by comparing it to several procedures from the literature. Section 5 summarizes the main results and gives an outlook on possible future work. Some minor computations and proofs of auxiliary results are given in the appendix.

2. Theoretical Results

In this section, we study a LLN and a CLT for T_N (Theorem 2.1 and Theorem 2.2, respectively). For the proofs, we make use of several auxiliary results: Lemma 2.3 approximates T_N by means of a Taylor expansion, Proposition 2.4 shows the LLN for this approximation and Proposition 2.5 makes use of the Hoeffding decomposition to prove a CLT for the approximation. Hereby, we center by the sample size dependent expected value of the approximation rather than by the corresponding limit. To justify this step Lemma 2.6 investigates the convergence rate of the centering terms.

Before we proceed, we present our assumptions on the data and introduce some further notation.

In what follows, we always assume that $X_1, \ldots, X_n$ fulfill the following properties:

- Identical distribution under the null hypothesis
- Independence
- Zero mean
- Existing fourth moments

We like to note that the independence assumption can probably be weakened. Furthermore, zero means can be obtained by centering and thus the assumption can be dropped. However, the last property is necessary, because we work towards a CLT involving empirical variances. We therefore have to handle second moments of the squared random variables. In the following, we refer to this set of assumptions by (A).
Under $\mathbb{H}_0$, the random variables $X_1, \ldots, X_n$ are i.i.d.. Let in this case $\sigma = \sqrt{\mathrm{Var}(X_1)}$ and $\gamma = \sqrt{\mathrm{Var}(X_1^2/\sigma^2)}$ denote their common standard deviation and the common standard deviation of their scaled squares, respectively. We also frequently make use of the auxiliary random variables

$$S_{N,j} = \frac{1}{\sqrt{\tau_N}} \sum_{i \in B_{N,j}} ((X_i/\sigma)^2 - 1)$$

for $j = 1, \ldots, N$ to approximate $\tau_N^{a/2} T_N$ by

$$\tilde{T}_N = \frac{1}{N(N-1)} \sum_{1 \le j \ne k \le N} |S_{N,j} - S_{N,k}|^a ,$$

see Lemma 2.3. Under $\mathbb{H}_0$ the $S_{N,1}, \ldots, S_{N,N}$ are i.i.d. with $E(S_{N,1}) = 0$ and $Var(S_{N,1}) = \gamma^2$.

Let us start by formulating our main results:

Theorem 2.1 (Law of Large Numbers). *Let $X_1, \ldots, X_n$ fulfill the properties (A) and let $\frac{N}{\tau_N} \to 0$ as $n \to \infty$. For $a \in (1, 2)$, let additionally $E|X_1|^{4a} < \infty$ be fulfilled. Then, under $\mathbb{H}_0$ we get*

$$\frac{\tau_N^{a/2}}{\gamma^a} T_N \longrightarrow 2^{a/2} E|Z|^a, \tag{1}$$

in probability as $n \to \infty$, where Z denotes a standard normal random variable.

The proof of this theorem is done in two steps. First, in Lemma 2.3 i) we take a Taylor expansion of the logarithm. This allows to show that the left hand side of (1) can be approximated by the first order term in the Taylor expansion, which is $\tilde{T}_N/\gamma^a$. In a second step, Proposition 2.4 shows that the latter expression converges in probability to the desired limit.

Theorem 2.2 (Central Limit Theorem). *Let $X_1, \ldots, X_n$ fulfill the properties (A) and let $E|X_1|^6$ be finite. If $a \in (1.5, 2)$, in addition assume that there exists a $\delta > 0$ such that $E|X_1|^{4a+\delta} < \infty$. If $\frac{N}{\tau_N^{\min(a,2-a)}} \longrightarrow 0$, then for $a \in (0, 2)$, under $\mathbb{H}_0$, we obtain*

$$\sqrt{N} \left(\frac{\tau_N^{a/2}}{\gamma^a} T_N - 2^{\frac{a}{2}} E|Z|^a \right) \longrightarrow N(0, \tilde{\sigma}^2)$$

in distribution as $n \to \infty$, where $\tilde{\sigma}^2 = 4\mathrm{Var}(h_1(Z))$ for $h_1(x) = E|x - Z'|^a$ and Z, Z' denote independent standard normal random variables.

This CLT allows us to construct an asymptotic test for the stationarity of the variance in a straightforward way, see page 6. It is proven in three steps. First, we apply the Taylor expansion of the logarithm conducted in Lemma 2.3 ii). Afterwards, this approximation is regarded as a U-statistic based on the random variables $S_{N,j}, j = 1, \ldots, N$. Its distribution depends on the sample size. To obtain an adequate CLT, we modify the classical CLT for U-statistics in Proposition 2.5, making use of the Hoeffding decomposition. Finally, Lemma 2.6 allows to show that the asymptotic Gaussianity still holds when we replace the sample size dependent centering $E(\tilde{T}_N/\gamma^a)$ used in Proposition 2.5 by its limit $2^{\frac{a}{2}} E|Z|^a$.

Next, we state several lemmas as well as propositions and apply them to prove the above theorems. The proofs of the auxiliary results are given in the appendix.

The following result provides a Taylor expansion of T_N allowing to handle the otherwise bothersome logarithms in T_N.

Lemma 2.3 (Taylor Expansion)**.** *Let the assumptions of Theorem 2.1 be fulfilled.*

i) It then holds that

$$\tau_N^{a/2} T_N - \tilde{T}_N \longrightarrow 0$$

in probability as $n \to \infty$.

ii) If $\frac{N}{\tau_N^{\min(a,2-a)}} \to 0$, it holds that

$$\sqrt{N}\left(\tau_N^{a/2} T_N - \tilde{T}_N\right) \longrightarrow 0$$

in probability as $n \to \infty$.

Due to Lemma 2.3, our test statistic $\frac{\tau_N^{a/2}}{\gamma^a} T_N$ can be approximated by the scaled U-statistic

$$\tilde{T}_N/\gamma^a = \frac{1}{N(N-1)} \sum_{1 \le j \ne k \le N} h(S_{N,j}/\gamma, S_{N,k}/\gamma) \quad \text{for } a \in (0,2)$$

with kernel $h(x,y) = |x-y|^a$. In the following, we therefore give some results on the asymptotic behavior of $\tilde{T}_N/\gamma^a$. Note that the underlying random variables form a triangular array obtained via the normalized block sums.

Proposition 2.4 (U-Statistics of Renormalized Block Sums)**.** *Let $X_1, \ldots, X_n$ fulfill the assumptions (A) and let Z denote a standard normal random variable. If $\tau_N \to \infty$ as $n \to \infty$, under $\mathbb{H}_0$ it holds:*

i) $\lim_{N\to\infty} E(\tilde{T}_N/\gamma^a) = 2^{a/2} E|Z|^a$.

ii) If $E|X_1|^{4a} < \infty$, then $\tilde{T}_N/\gamma^a$ converges in probability towards $2^{a/2}E|Z|^a$ as $n \to \infty$.

We now are able to prove the LLN stated above:

Proof of Theorem 2.1. The desired convergence in probability, $\frac{\tau_N^{a/2}}{\gamma^a} T_N \longrightarrow 2^{a/2}E|Z|^a$, follows from combining the fact that $\tau_N^{a/2} T_N - \tilde{T}_N \longrightarrow 0$ (see Lemma 2.3 i)) and that it holds $\tilde{T}_N/\gamma^a \longrightarrow 2^{a/2}E|Z|^a$ (see Proposition 2.4 ii)). □

We now state the version of the CLT relying on the sample size dependent centering:

Proposition 2.5. *Let the assumptions of Theorem 2.2 be fulfilled. It then holds that*

$$\sqrt{N}\left(\tilde{T}_N/\gamma^a - E(\tilde{T}_N/\gamma^a)\right) \longrightarrow N(0, \tilde{\sigma}^2)$$

in distribution as $n \to \infty$.

We also determine the convergence rate for the sequence of centering terms used in the CLT above:

Lemma 2.6. *Let $X_1, \dots, X_n$ fulfill the properties (A) and additionally let $E|X_1|^6 < \infty$ hold. Also, consider a standard Gaussian random variable Z. Then for any $a \in (0,2)$, it holds that*

$$\sqrt{N}\left|E(\tilde{T}_N/\gamma^a) - 2^{\frac{a}{2}} E|Z|^a\right| = \mathcal{O}\left(\frac{\log(\tau_N)\sqrt{n}}{\tau_N}\right).$$

Proposition 2.5, Lemma 2.6 and Lemma 2.3 ii) allow us to prove the desired CLT:

Proof of Theorem 2.2. We rewrite

$$\sqrt{N}\left(\frac{\tau_N^{a/2}}{\gamma^a}T_N - 2^{\frac{a}{2}}E|Z|^a\right) = \underbrace{\sqrt{N}\left(\tau_N^{a/2}T_N - \tilde{T}_N\right)}_{(*)}/\gamma^a + \underbrace{\sqrt{N}\left(\tilde{T}_N/\gamma^a - E(\tilde{T}_N/\gamma^a)\right)}_{(**)}$$
$$+ \underbrace{\sqrt{N}\left(E(\tilde{T}_N/\gamma^a) - 2^{\frac{a}{2}}E|Z|^a\right)}_{(***)}$$

and see that $(*) \longrightarrow 0$ in probability by Lemma 2.3 ii), $(**) \longrightarrow N(0, \tilde{\sigma}^2)$ in distribution by Proposition 2.5 and $(***) \longrightarrow 0$ by Lemma 2.6 since $\frac{N}{\tau_N^{\min(a,2-a)}} \longrightarrow 0$. □

3. Simulation Study of the Limit Theorems

In this section we present two simulation studies illustrating our theoretical results.

First, we check the LLN stated in Theorem 2.1 for several distributions generating the data and different choices of N, the number of blocks. More precisely, we investigate the standard Gaussian distribution, the t-distribution with 10 degrees of freedom and the exponential distribution with parameter $\lambda = 1$. Data sampled from the latter two distributions is standardized to have expectation 0 and variance 1, in order to match our assumption and allow for comparability, respectively. For each of the sample sizes $n = 100, 500, 1000, 2500, 5000, 10000, 15000$, we determine five numbers of blocks via $N = \lfloor n^{1-s} \rfloor$, choosing $s = 0.4, 0.5, \dots, 0.8$. Here, $\lfloor x \rfloor$ denotes the largest integer smaller or equal to x. Given a fixed number of blocks, the corresponding data of size n is split into N blocks of, as far as possible, equal length τ_N, that is, the actual block lengths differ by at most one. Hence, $\tau_N \approx n^s$ follows immediately. The assumption $\frac{N}{\tau_N} \longrightarrow 0$ in Theorem 2.1 therefore translates to $s > 0.5$. For each sample size n and each distribution, 100 000 samples of size n are generated. For each sample, the left hand side of (1) is calculated for each appropriate choice of τ_N and for $a = 1$. Thereby, γ is replaced by its empirical counterpart, which in turn involves the estimation of σ by the empirical standard deviation. Finally, the empirical bias and root mean squared error (RMSE) when compared to the limit $\sqrt{2}E|Z| \overset{6.1}{=} 2/\sqrt{\pi}$ are calculated from the 100 000 replications for each scenario. The results are presented in Figure 1.

We observe several interesting properties of the method in these plots: First, the precision of the estimation in general increases with the sample size. This holds even for $s \leq 0.5$, except for the bias in case of t-distributed data. Second, for each individual plot, the graphs corresponding to the different block lengths $N = \lfloor n^{1-s} \rfloor$ are almost perfectly ordered according to the parameter s for both the bias and the RMSE. In general the absolute bias decreases in s, so that a large number of comparatively small blocks leads to a small bias. The RMSE behaves in the opposite way, indicating a larger variance in case of many small blocks. We thus observe a classical bias–variance trade-off. Finally, the distribution of the

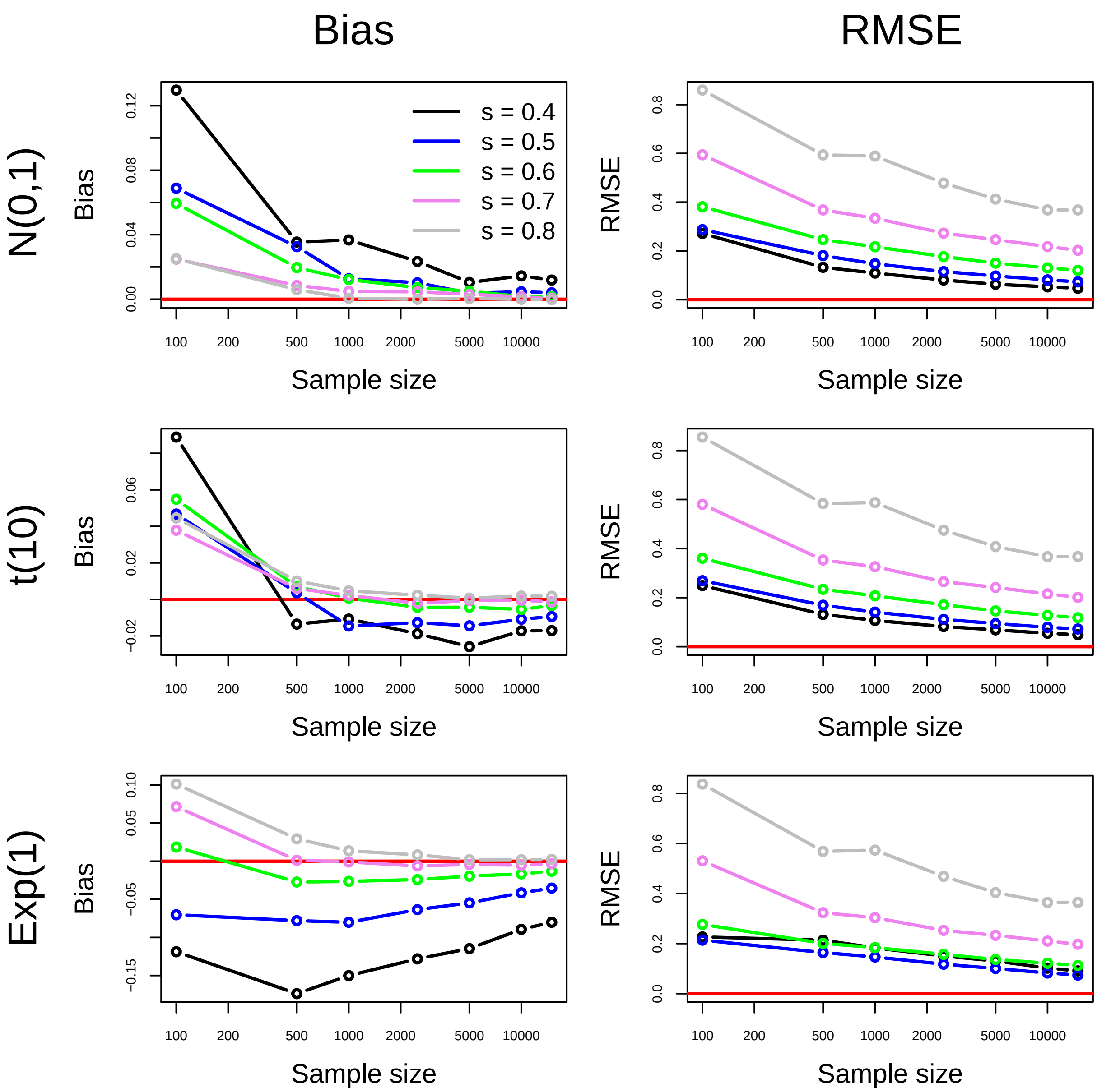

FIGURE 1. Empirical bias and RMSE for the LHS of (1) as a function of the sample size, using data generated from the standard Gaussian distribution (N(0,1)), the standardized t-distribution with 10 degrees of freedom (t(10)) and the standardized exponential distribution with parameter $\lambda = 1$ (Exp(1)) and choosing different numbers of blocks $N = \lfloor n^{1-s} \rfloor$. The x-axis is presented in log scale. The neutral axis is represented by the red line.

data appears to have some influence on the bias, but little on the RMSE. In the Gaussian case, the limit is on average overestimated for all choices of blocks considered. For the other two distributions, it is quite often underestimated for $s = 0.4, 0.5, 0.6$. In summary, these results support the validity of the LLN, at least for $s > 0.5$.

Next, we illustrate the CLT stated in Theorem 2.2. For this purpose, a test for the stationarity of the variance is constructed based on Theorem 2.2 and its size for various sample sizes n is investigated. Under $\mathbb{H}_0$, the statistic

$$T_{asy} = \sqrt{N}\left(\frac{\tau_N^{a/2}}{\gamma^a}T_N - 2^{\frac{a}{2}}E|Z|^a\right)$$

asymptotically follows a Gaussian distribution with expectation 0 and variance $\tilde{\sigma}^2$, due to the CLT. Under the alternative of piecewise stationarity, the variance estimates $\hat{\sigma}_j^2, j = 1, \ldots, N$, typically differ more than under the null hypothesis. This results in comparatively large values for T_N and consequently also for T_{asy}. Thus, the procedure rejecting $\mathbb{H}_0$ if and only if $T_{asy} > \tilde{\sigma}u_{1-\alpha}$ provides an asymptotic test for the significance level α suitable for our problem, where u_β denotes the β-quantile of the standard Gaussian distribution. For the case $a = 1$ considered in the following it holds that $\tilde{\sigma} = \frac{4}{3} + \frac{8}{\pi}(\sqrt{3} - 2)$, cf. Gerstenberger & Vogel (2015).

The test introduced above is applied on samples of sizes $n = 100, 500, 1000, 2500, 5000, 10000, 15000$. Thus, we generated data from each of the three distributions used in the simulations concerning the LLN. We also consider the same splitting procedure for the blocks and the same values of s as before, repeating each scenario 100 000 times. The corresponding rejection rates for $\alpha = 5\%$ are given in percent and presented in Figure 2:

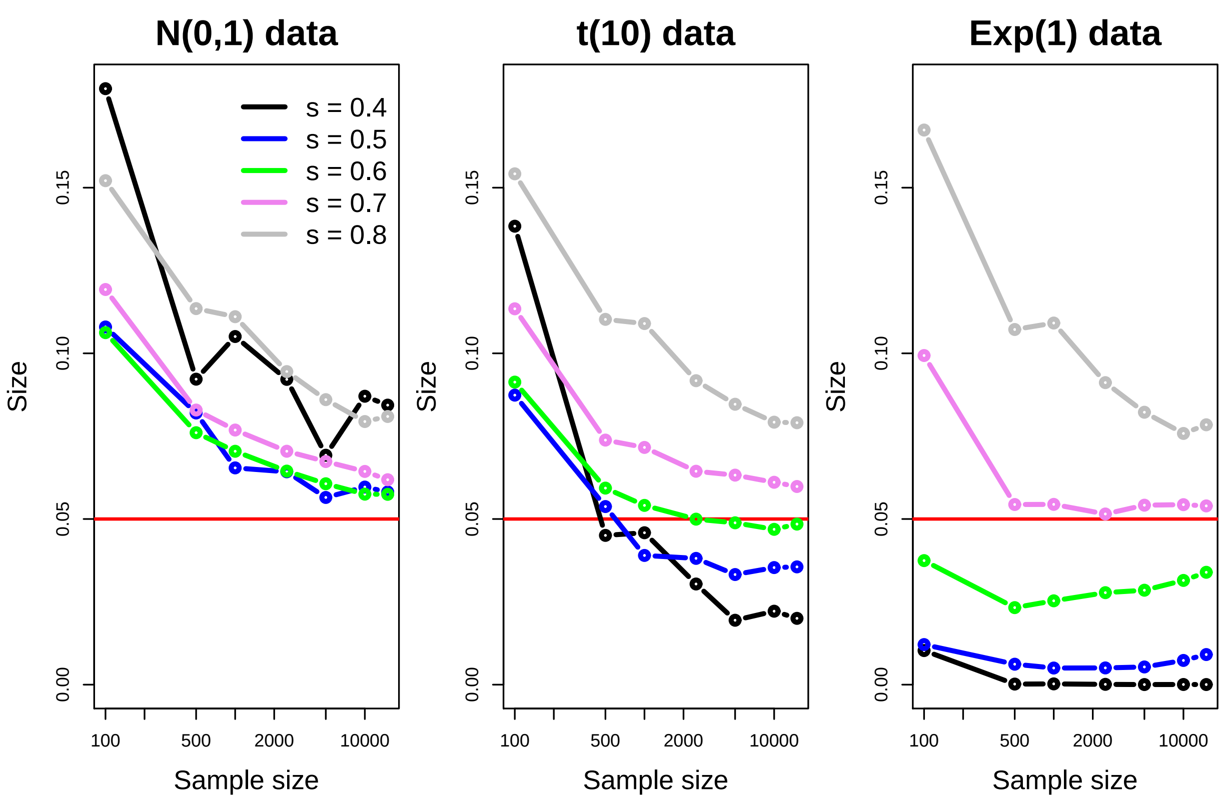

FIGURE 2. Empirical size of the asymptotic test based on T_{asy} as a function of the sample size, using data generated from the standard Gaussian distribution (N(0,1)), the standardized t-distribution with 10 degrees of freedom (t(10)) and the standardized exponential distribution with parameter $\lambda = 1$ (Exp(1)) and choosing different numbers of blocks $N = \lfloor n^{1-s} \rfloor$. The x-axis is presented in log scale. The significance level $\alpha = 0.05$ is represented by the horizontal red line.

The results differ considerably among the different choices of the number of blocks and also depend on the distribution of the data. Overall, $s = 0.6$ and $s = 0.7$ lead to rejection rates closest to the significance level $\alpha = 0.05$. The block choices induced by $s = 0.4$ and $s = 0.8$ seem inappropriate, so that a compromise between the block size and the number of blocks appears to be crucial for our procedure. This may be caused by the fact that our procedure relies on the convergence both in terms of τ_N and N. The latter also explains the overall rather slow convergence of the method. For $s = 0.5$ the test performs quite well in the Gaussian case, but is disadvantageous otherwise. The impact of the distribution choice is quite similar to the one for the bias of the LLN presented in Figure 1.

4. Comparison to Alternative Procedures

There exist many approaches in the literature for checking a series of observations for constant variance. In this section, we first introduce a representative collection of such methods. Afterwards, these procedures are applied to artificial data along with the asymptotic test studied in Section 3, allowing to evaluate the performance of the latter.

One classical tool for the detection of structural breaks is the CUSUM procedure. Among others, Wied et al. (2012) utilize it in order to detect changes in the variance structure. They focus on the CUSUM statistic

$$T_{CUS} = \widehat{D} \max_{1 \le t \le n} \left| \frac{t}{\sqrt{n}} (\widehat{\sigma}^2_{1:t} - \widehat{\sigma}^2_{1:n}) \right|,$$

which asymptotically has the same distribution as $\sup_{z \in [0,1]} |B(z)|$. Here, $B(\cdot)$ denotes a one-dimensional Brownian bridge and $\widehat{\sigma}^2_{1:t}$ describes the empirical variance of the first t observations. The normalising data dependent scalar $\widehat{D}$ takes into account the asymptotic variance necessary to attain the asymptotic distribution. The method determines the difference between the variance estimated on the complete sample and the variance estimated on the subsample up to observation t. It then considers the maximal absolute deviation to test for a structural break via critical values derived from asymptotics. The integral part of the CUSUM procedure is the cumulative sum arising in the computation of $\hat{\sigma}^2_{1:t}$. Unfortunately, this quantity is prone to masking effects in case of several change points. In other words, multiple changes of the variance can lead to low detection rates, if their effects cancel each other out in the cumulative sums. Therefore, in principle, the CUSUM procedure is designed to test for at most one structural change at a time.

Essentially, the problem sketched above arises because $\hat{\sigma}^2_{1:t}$ reflects a sort of averaged variance up to observation t rather than the local variance close to observation t. Therefore, many authors try to alleviate this masking effect by estimating the variance locally. More precisely, they perform the variance estimation on a moving window of observations. One way to adopt this strategy is the MOSUM test based on the statistic

$$T_{MOS} = \tilde{D} \sqrt{\lfloor hn \rfloor} \max_{0 \le t \le n - \lfloor hn \rfloor} \left| \tilde{\sigma}^2_{(t+1):(t+\lfloor hn \rfloor)} - \tilde{\sigma}^2_{1:n} \right|,$$

where $\tilde{\sigma}^2_{i:j} = \frac{1}{j-i} \sum_{k=i}^{j} X_k^2$ for $j > i$, c.f. the approach of Chu et al. (1995) applied to the squared observations. The method estimates local variances using a sliding window of fixed size $\lfloor nh \rfloor$. It then compares the local estimates to the global one, similar to the CUSUM procedure. The normalizing data dependent factor $\tilde{D}$ as well as appropriate critical values are derived from asymptotics. The procedure is implemented in the *strucchange* package (Zeiles et al., 2002) for the software R (R Core Team, 2016). In the following, it is always conducted choosing the default bandwidth $h = 0.15$, unless stated otherwise.

Ross (2013) chooses a different approach to test the stationarity of the variance. Motivated by the classical distribution-free Mood test, he considers the ranks corresponding to the random variables $X_1, \dots, X_n$. Similar to the CUSUM procedure, for every $t = 1, \dots, n$ the standardized statistic of the Mood test

$$T_{Mood,t} = \frac{\left| \sum_{h=1}^{t} \left(r_h - \frac{n+1}{2} \right)^2 - \mu_t^* \right|}{\sigma_t^*}$$

is then computed on the first t of these ranks $r_1, \dots, r_n$. The involved quantities $\mu_t^* = t(n^2-1)/12$ and $\sigma_t^* = \sqrt{t(n-t)(n+1)(n^2-4)/180}$ correspond under the null hypothesis to the expected value and the standard deviation of the sum appearing in $T_{Mood,t}$, respectively. In analogy to the CUSUM test the final test statistic is constructed considering the maximum over $t = 1, \dots, n$:

$$T_{Mood} = \max_{t=1,\dots,n} T_{Mood,t}.$$

Corresponding critical values can be derived by simulations and are listed for several sample sizes in Ross (2013). The procedure is based solely on the ranks of the observations and is therefore distribution-free. However, it suffers from masking problems for the same reasons as the CUSUM test.

In addition to these methods, we consider the asymptotic procedure based on T_{asy} investigated in the second part of Section 3. Another test is constructed by applying the permutation approach to T_N. Thereby, we always conduct 10 000 permutations. As shown in Wornowizki et al. (2016), the performance of this method is hardly influenced by the weight parameter a and particularly fit for detecting multiple structural changes.

We apply all five tests in five data scenarios listed below. To give a clear and compact overview, let $|n = n_1, \sigma = \sigma_1 | n = n_2, \sigma = \sigma_2|$ denote n_1 observations with standard deviation σ_1, followed by n_2 observations with standard deviation σ_2. The data cases under study are:

1) $|n = 6000, \sigma = 1|$
2) $|n = 3000, \sigma = 1| n = 3000, \sigma = 1.1|$
3) $|n = 2000, \sigma = 1| n = 2000, \sigma = 1.1 |n = 2000, \sigma = 1|$
4) $|n = 1200, \sigma = 1| n = 1200, \sigma = 1.1 |n = 1200, \sigma = 1| n = 1200, \sigma = 1.1 |n = 1200, \sigma = 1|$
5) $|n = 1200, \sigma = 1| n = 600\ \ , \sigma = 1.1 |n = 1200, \sigma = 1| n = 1800, \sigma = 1.1 |n = 1200, \sigma = 1|$

These correspond to the null hypothesis (1)), to one structural break (2)), to two structural breaks (3)), to four equidistant structural breaks (4)) and to four non-equidistant structural breaks (5)). The data for each of the five scenarios is generated from four different distributions. These are the standard Gaussian distribution, the t-distribution with 10 and 5 degrees of freedom and the exponential distribution with parameter $\lambda = 1$. Each of them is standardized to have expectation zero and variance one. We thus end up with 20 data cases in total. For each of them 10 000 replications are conducted and the five test procedures introduced above are applied to the data set of each replication, where we choose a significance level of $\alpha = 5\%$. The permutation tests are executed with 2000 permutations. In order to reduce the computational burden, both tests based on T_N are carried out for $a = 1$ and $s = 0.7$ only. The corresponding rejection rates are presented in Table 1.

Under the null hypothesis of constant variance, the newly proposed test behaves somewhat liberal for Gaussian and $t(10)$ data. Besides that, the only procedure clearly violating the five percent bound is the one based on T_{MOS} in case of $t(5)$ data. As expected, the CUSUM and Mood-type procedures perform much better than their competitors in case of one structural change. However, as the number of change point increases, both loose a tremendous amount of power. The only exception is the exponential case, where the rank-based procedure universally outperforms all remaining tests by a considerably amount. All in all, the methods

TABLE 1. Rejection rates in percent for the five tests under study for the five different data scenarios presented on page 9, and for Gaussian data (N(0,1)), data from the t-distribution with 10 and 5 degrees of freedom (t(10) and t(5)) and exponential data (Exp(1)). The results of the asymptotic test derived from the CLT in Theorem 2.2 and for its equivalent based on permutations for $s = 0.7$ and $a = 1$ are denoted by T_{asy} and T_{perm}, respectively.

		T_{CUS}	T_{Mood}	T_{MOS}	T_{asy}	T_{perm}
1) $\mathbb{H}_0$:	N(0,1)	4.3	5.0	5.0	6.3	4.6
	t(10)	4.6	5.0	5.4	6.3	5.1
	t(5)	3.7	5.2	6.4	4.2	5.1
	Exp(1)	3.9	4.8	5.5	5.2	4.9
2) 1 break:	N(0,1)	99.7	94.7	89.6	95.6	94.3
	t(10)	96.7	91.0	72.3	82.3	80.0
	t(5)	71.7	86.2	38.7	43.3	47.8
	Exp(1)	60.8	100.0	30.9	34.5	33.2
3) 2 breaks:	N(0,1)	15.6	10.8	35.4	30.9	25.6
	t(10)	11.3	9.8	24.8	21.2	17.9
	t(5)	5.7	9.6	15.3	9.6	11.1
	Exp(1)	5.7	44.3	12.3	9.2	9.7
4) 4 breaks:	N(0,1)	10.3	26.6	86.4	87.3	83.6
	t(10)	7.8	23.6	66.9	68.3	64.2
	t(5)	4.4	20.3	35.9	32.7	36.7
	Exp(1)	4.4	91.9	28.6	26.5	24.9
5) 4 noneq. breaks:	N(0,1)	27.3	31.5	83.4	88.3	85.1
	t(10)	18.7	27.9	63.5	68.0	63.9
	t(5)	9.3	23.8	34.3	32.5	36.5
	Exp(1)	8.7	96.3	27.9	26.5	25.2

referred to as T_{asy} and T_{perm} lead to competitive results across all scenarios considered. Their rejection rates are usually quite similar with the exception of $t(5)$-data scenarios. Here, the asymptotic test looses a small amount of power because the respective requirement of existing fourth moments is not met. Overall, the MOSUM-type procedure based on T_{MOS} slightly outperforms the tests relying on T_{asy} and T_{perm}.
In addition to high rejection rates under various alternatives, suitable change point detection procedures allow an adequate estimation of the position of the structural breaks. We thus apply the corresponding estimation scheme presented in Wornowizki et al. (2016) to the tests relying on T_{asy} and T_{MOS}. The results for the situations 3) and 4) in case of Gaussian data are depicted in Figure 3. The plots shows quite clearly that T_{MOS}, even though occasionally having higher rejection rates, has much more difficulties to allocate the individual structural breaks. The method therefore leads to a considerably worse estimation of the position and

number of the structural changes. In comparison, T_{asy} is able to reconstruct the true variance structure much more reliably.

Since the tests based on T_{asy} and T_{MOS} overall performed best in terms of power, we reinvestigate them in further data cases. As before, we work with 10 000 samples of size $n = 6000$. However, this time we create nine variance changes which are as follows:

I) Equidistant blocks of length 600; standard deviations alternating between 1 and 1.1
II) Randomly positioned blocks with minimum block length 120; standard deviations alternating between 1 and 1.1
III) Randomly positioned blocks with minimum block length 120; standard deviations randomly chosen from the candidates 0.9, 0.95, 1, 1.05, 1.1, where adjacent blocks must have different variances.

For the test using T_{asy}, we split the data into 5, 10, 20 or 80 equidistant blocks of lengths 1200, 600, 300 or 75. For T_{MOS}, we consider the window sizes $\lfloor nh \rfloor = 75, 300, 600, 1200$. The results for both procedures are collected in Table 2. As expected, both methods perform best in data case I), if they are applied using the correct data partition. Also, quite unsurprisingly, too large blocks of length 1200 substantially decrease the rejection rate of the T_{asy}-test. The same holds for large window sizes for the MOSUM procedure. We also note that this time our newly proposed test outperforms its competitor based on T_{MOS} in each case but the one, often showing considerably higher rejection rates. Combined with the above results, this stresses the notion that the procedure introduced in this work is certainly worth investigating.

Table 2. Rejection rates in percent of the tests based on T_{asy} and T_{MOS} for different block lengths τ_N in the data cases I), II), and III, see page 11.

	T_{asy}				T_{MOS}				
τ_N	75	300	600	1200	75	300	600	900	1200
I)	62.5	93.0	98.0	8.3	0.0	47.0	50.2	15.3	4.9
II)	54.5	74.4	68.8	54.3	0.0	46.8	53.5	54.2	52.6
III)	88.0	95.7	94.6	87.1	0.3	85.6	91.5	91.8	89.7

5. Conclusion

The present paper introduces an asymptotic test on the stationarity of the variance for a sequence of independent random variables. The statistic we focus on is up to now only utilized for testing via a permutation approach, see Wornowizki et al. (2016). Our asymptotic pendant to this promising procedure offers similar power along with a tremendous reduction of computational complexity. As shown in various scenarios in our simulation study, the new method also performs well when compared to different tests suggested in the literature in case of symmetrical distributions. In particular, our method estimates the structural break positions adequately in case of a rejection of the null hypothesis. To derive the asymptotic distribution under the null hypothesis, the statistic of interest is represented as a U-statistic computed from blockwise variance estimates. Its distribution depends on the sample size, so that the classical LLN and CLT for U-statistics are not applicable. We thus develop new versions of such limit theorems corresponding to our setting and illustrate their correctness in a simulation study. Future work in this area could treat the case of dependent data and possibly make use of appropriate resampling techniques. Also, proper strategies for choosing a suitable splitting of the data into blocks may further increase the method's performance.

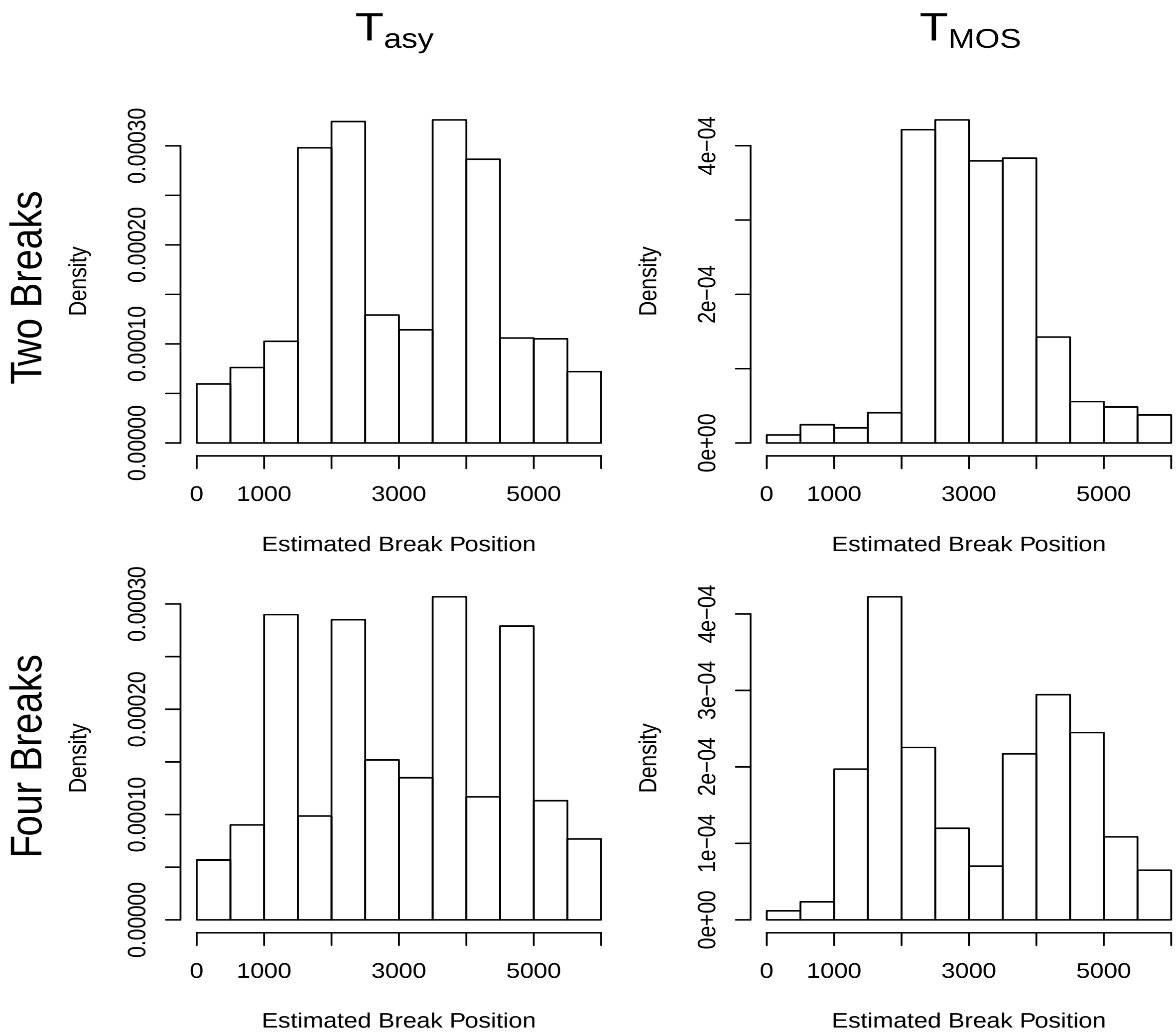

FIGURE 3. Estimated change point positions computed on Gaussian data for the tests based on T_{asy} (left colum) and T_{MOS} (right column) in the case of two (upper row) and four (lower row) equidistant variance changes.

6. Appendix

Proof of Lemma 2.3. i) The lemma essentially provides an approximation of T_N. It is established using the fact that $|\log(x+1) - x| = O(x^2)$ holds for $|x| < 1$. To ensure the latter assumption holds, we introduce the set $M_\varepsilon = \{\max_{1 \le j \le N} |S_{N,j}|/\sqrt{\tau_N} \le \varepsilon\}$ for any $\varepsilon \in (0, 1)$. First we prove that its complement M_ε^C has vanishing probability. We then show that the desired approximation indeed holds on the set M_ε, yielding the claim. In the following, the indicator functions corresponding to the sets M_ε and M_ε^C are denoted by I_{M_ε} and $I_{M_\varepsilon^C}$, respectively.

Rewriting $\tau_N^{a/2}T_N - \tilde{T}_N$ in terms of the $S_{N,j}$ yields

$$
\begin{aligned}
&|\tau_N^{a/2}T_N - \tilde{T}_N| \\
\leq & \frac{\tau_N^{a/2}}{N(N-1)} \sum_{1\leq j\neq k\leq N} \Bigg| \left| \log\left(\frac{1}{\tau_N}\sum_{i\in\mathcal{B}_{N,j}} X_i^2\right) - \log\left(\frac{1}{\tau_N}\sum_{i\in\mathcal{B}_{N,k}} X_i^2\right)\right|^a \\
& - \left|\frac{1}{\tau_N}\sum_{i\in B_{N,j}}((X_i/\sigma)^2-1) - \frac{1}{\tau_N}\sum_{i\in B_{N,k}}((X_i/\sigma)^2-1)\right|^a \Bigg| \\
= & \frac{\tau_N^{a/2}}{N(N-1)} \sum_{1\leq j\neq k\leq N} \Bigg| \left| \log\left(\frac{1}{\tau_N}\sum_{i\in\mathcal{B}_{N,j}} (X_i/\sigma)^2\right) - \log\left(\frac{1}{\tau_N}\sum_{i\in\mathcal{B}_{N,k}} (X_i/\sigma)^2\right)\right|^a \\
& - \left|\frac{1}{\tau_N}\sum_{i\in B_{N,j}}((X_i/\sigma)^2-1) - \frac{1}{\tau_N}\sum_{i\in B_{N,k}}((X_i/\sigma)^2-1)\right|^a \Bigg| \\
= & \frac{\tau_N^{a/2}}{N(N-1)} \sum_{1\leq j\neq k\leq N} \left| \left| \log\left(\frac{1}{\sqrt{\tau_N}}S_{N,j}+1\right) - \log\left(\frac{1}{\sqrt{\tau_N}}S_{N,k}+1\right)\right|^a - \left|\frac{1}{\sqrt{\tau_N}}S_{N,j} - \frac{1}{\sqrt{\tau_N}}S_{N,k}\right|^a \right|.
\end{aligned}
$$

We now consider the probability of the above expression restricted to the set M_ε^C:

$$
\begin{aligned}
& P\Bigg(I_{M_\varepsilon^C} \cdot \frac{\tau_N^{a/2}}{N(N-1)} \sum_{1\leq j\neq k\leq N} \Bigg| \left|\log\left(\frac{1}{\sqrt{\tau_N}}S_{N,j}+1\right) - \log\left(\frac{1}{\sqrt{\tau_N}}S_{N,k}+1\right)\right|^a \\
& - \left|\frac{1}{\sqrt{\tau_N}}S_{N,j} - \frac{1}{\sqrt{\tau_N}}S_{N,k}\right|^a \Bigg| > 0 \Bigg) \\
\leq & P\left(M_\varepsilon^C\right) = P\left(\max_{1\leq j\leq N}\left|\frac{1}{\sqrt{\tau_N}}S_{N,j}\right| > \varepsilon\right) \leq N\cdot P\left(\left|\frac{1}{\sqrt{\tau_N}}S_{N,1}\right| > \varepsilon\right) \leq N\frac{\gamma^2}{\tau_N\varepsilon^2} \longrightarrow 0
\end{aligned}
$$

Hereby, we used Chebychev's inequality and the fact that $\frac{N}{\tau_N} \to 0$ holds by assumption. Thus, as desired, the probability of M_ε^C vanishes implying the proposition on the set M_ε^C.

Next, we consider the case of M_ε. Our goal is to bound

$$
I_{M_\varepsilon} \cdot \frac{\tau_N^{a/2}}{N(N-1)} \sum_{1\leq j\neq k\leq N} \left| \left|\log\left(\frac{1}{\sqrt{\tau_N}}S_{N,j}+1\right) - \log\left(\frac{1}{\sqrt{\tau_N}}S_{N,k}+1\right)\right|^a - \left|\frac{1}{\sqrt{\tau_N}}S_{N,j} - \frac{1}{\sqrt{\tau_N}}S_{N,k}\right|^a \right|
$$

from above by a nonnegative random variable with vanishing mean, which would complete the proof. The actual calculations differ for $a \in (0,1)$ and $a \in [1,2)$. However, for both of them the approximation $|\log(x+1) - x| = O(x^2)$ is valid in case of $|x| < 1$ and implies the useful inequality

$$
\left|\log\left(\frac{1}{\sqrt{\tau_N}}S_{N,j}+1\right) - \frac{1}{\sqrt{\tau_n}}S_{N,j}\right| \cdot I_{M_\varepsilon} \leq c\left(\frac{1}{\sqrt{\tau_n}}S_{N,j}\right)^2 \tag{2}
$$

for any $1 \leq j \leq N$ and some $c \in \mathbb{R}^+$. For $a \in (0,1)$ Lemma 6.2 i) allows us to show

$$
\begin{aligned}
&I_{M_\varepsilon} \cdot \frac{\tau_N^{a/2}}{N(N-1)} \sum_{1\leq j\neq k\leq N} \left| \left| \log\left(\frac{1}{\sqrt{\tau_N}} S_{N,j}+1\right) - \log\left(\frac{1}{\sqrt{\tau_N}} S_{N,k}+1\right)\right|^a - \left|\frac{1}{\sqrt{\tau_N}} S_{N,j} - \frac{1}{\sqrt{\tau_N}} S_{N,k}\right|^a \right| \\
\overset{(6)}{\leq} &I_{M_\varepsilon} \cdot \frac{\tau_N^{a/2}}{N(N-1)} \sum_{1\leq j\neq k\leq N} \left| \log\left(\frac{1}{\sqrt{\tau_N}} S_{N,j}+1\right) - \log\left(\frac{1}{\sqrt{\tau_N}} S_{N,k}+1\right) - \frac{1}{\sqrt{\tau_N}} S_{N,j} + \frac{1}{\sqrt{\tau_N}} S_{N,k}\right|^a \\
\leq &I_{M_\varepsilon} \cdot \frac{\tau_N^{a/2}}{N(N-1)} \sum_{1\leq j\neq k\leq N} \left(\left| \log\left(\frac{1}{\sqrt{\tau_N}} S_{N,j}+1\right) - \frac{1}{\sqrt{\tau_N}} S_{N,j}\right| + \left| \log\left(\frac{1}{\sqrt{\tau_N}} S_{N,k}+1\right) - \frac{1}{\sqrt{\tau_N}} S_{N,k}\right| \right)^a \\
\overset{(2)}{\leq} &I_{M_\varepsilon} \cdot \frac{\tau_N^{a/2}}{N(N-1)} \sum_{1\leq j\neq k\leq N} c^a \left(\left(\frac{1}{\sqrt{\tau_N}} S_{N,j}\right)^2 + \left(\frac{1}{\sqrt{\tau_N}} S_{N,k}\right)^2 \right)^a \\
\overset{(6)}{\leq} &I_{M_\varepsilon} \cdot \frac{c^a\tau_N^{a/2}}{N(N-1)} \sum_{1\leq j\neq k\leq N} \left(\left(\frac{1}{\sqrt{\tau_N}} S_{N,j}\right)^{2a} + \left(\frac{1}{\sqrt{\tau_N}} S_{N,k}\right)^{2a} \right) \\
= &I_{M_\varepsilon} \cdot \frac{2c^a\tau_N^{a/2}}{N} \sum_{1\leq j\leq N} \left(\frac{1}{\sqrt{\tau_N}} S_{N,j}\right)^{2a} = I_{M_\varepsilon} \cdot \frac{2c^a\tau_N^{a/2-a}}{N} \sum_{1\leq j\leq N} (S_{N,j})^{2a}.
\end{aligned}
$$

For the nonnegative random variable above, we get

$$
\begin{aligned}
&\mathrm{E}\left(I_{M_\varepsilon} \cdot \frac{2c^a\tau_N^{a/2-a}}{N} \sum_{1\leq j\leq N} (S_{N,j})^{2a} \right) \leq \mathrm{E}\left(\frac{2c^a\tau_N^{a/2-a}}{N} \sum_{1\leq j\leq N} (S_{N,j})^{2a} \right) \\
= &2c^a\tau_N^{a/2-a}\mathrm{E}\left((S_{N,1})^{2a}\right) \leq 2c^a\tau_N^{a/2-a}\mathrm{E}\left((S_{N,1})^2\right)^a \\
= &2c^a\tau_N^{a/2-a}\mathrm{Var}\left(S_{N,1}\right)^a = 2c^a\tau_N^{a/2-a}\gamma^{2a} \longrightarrow 0
\end{aligned}
$$

using Jensen's inequality. Thus, this random variable converges to 0 in probability. Consequently, the claim holds for $a \in (0,1)$.
In case of $a \in [1,2)$, we take a similar approach using Lemma 6.2 ii):

$$
\begin{aligned}
&I_{M_\varepsilon} \cdot \frac{\tau_N^{a/2}}{N(N-1)} \sum_{1\leq j\neq k\leq N} \left| \left| \log\left(\frac{1}{\sqrt{\tau_N}} S_{N,j}+1\right) - \log\left(\frac{1}{\sqrt{\tau_N}} S_{N,k}+1\right)\right|^a - \left|\frac{1}{\sqrt{\tau_N}} S_{N,j} - \frac{1}{\sqrt{\tau_N}} S_{N,k}\right|^a \right| \\
\overset{(7)}{\leq} &I_{M_\varepsilon} \cdot \frac{\tau_N^{a/2}}{N(N-1)} \sum_{1\leq j\neq k\leq N} a \left| \log\left(\frac{1}{\sqrt{\tau_N}} S_{N,j}+1\right) - \log\left(\frac{1}{\sqrt{\tau_N}} S_{N,k}+1\right) - \frac{1}{\sqrt{\tau_N}} S_{N,j} + \frac{1}{\sqrt{\tau_N}} S_{N,k}\right| \\
\leq &I_{M_\varepsilon} \cdot \frac{a\tau_N^{a/2}}{N(N-1)} \sum_{1\leq j\neq k\leq N} \left(\left| \log\left(\frac{1}{\sqrt{\tau_N}} S_{N,j}+1\right) - \frac{1}{\sqrt{\tau_N}} S_{N,j}\right| + \left| \log\left(\frac{1}{\sqrt{\tau_N}} S_{N,k}+1\right) - \frac{1}{\sqrt{\tau_N}} S_{N,k}\right| \right) \\
\overset{(2)}{\leq} &I_{M_\varepsilon} \cdot \frac{a\tau_N^{a/2}}{N(N-1)} \sum_{1\leq j\neq k\leq N} c \left(\left(\frac{1}{\sqrt{\tau_N}} S_{N,j}\right)^2 + \left(\frac{1}{\sqrt{\tau_N}} S_{N,k}\right)^2 \right) \\
= &I_{M_\varepsilon} \cdot \frac{2ca\tau_N^{a/2-1}}{N} \sum_{1\leq j\leq N} (S_{N,j})^2,
\end{aligned}
$$

and the latter converges to 0 in probability because of

$$\mathrm{E}\left(I_{M_\varepsilon} \cdot \frac{2ac\tau_N^{a/2-1}}{N} \sum_{1\le j\le N} (S_{N,j})^2 \right) \le \mathrm{E}\left(\frac{2ac\tau_N^{a/2-1}}{N} \sum_{1\le j\le N} (S_{N,j})^2 \right)$$
$$=2ca\tau_N^{a/2-1}\mathrm{E}\left((S_{N,1})^2\right) = 2ca\tau_N^{a/2-1}\mathrm{Var}\left(S_{N,1}\right) = 2ca\tau_N^{a/2-1}\gamma^2 \longrightarrow 0.$$

Thus the desired result follows for $a \in [1,2)$, which completes the proof.

ii) We use the same arguments as before and multiply the expression in i) by the factor $\sqrt{N}$. Consequently, one obtains the desired convergence for $a \in (0,1)$ if $\sqrt{N}2c^a\tau_N^{-a/2}\gamma^{2a} \longrightarrow 0$ holds. The latter is ensured by $\frac{N}{\tau_N^a} \to 0$. In the case $a \in [1,2)$ we take a similar approach exploiting $\frac{N}{\tau_N{}^{2-a}} \to 0$.

□

Proof of Proposition 2.4. i) By the classical central limit theorem $S_{N,j}/\gamma$ converges in distribution to a standard normal random variable as $n \to \infty$ (and thereby $\tau_N \to \infty$) for any $1 \le j \le N$. For $j \neq k$ the normalized block sums $S_{N,j}/\gamma$ and $S_{N,k}/\gamma$ are independent. Thus $(S_{N,j} - S_{N,k})/\gamma$ converges in distribution to a Gaussian random variable with mean zero and variance 2. Consequently, for $a \in (0,2)$, we obtain convergence of the expectations $E|(S_{N,j} - S_{N,k})/\gamma|^a$ towards $E|2^{1/2}Z|^a$, and hence

$$E(\tilde{T}_N/\gamma^a) = E|(S_{N,j} - S_{N,k})/\gamma|^a \longrightarrow 2^{a/2}E(|Z|^a) \text{ as } n \to \infty.$$

ii) Since $E(\tilde{T}_N/\gamma^a) \longrightarrow 2^{a/2}E|Z|^a$ holds by (i), it suffices to show that $\mathrm{Var}(\tilde{T}_N) \longrightarrow 0$. Exploiting that $\mathrm{Cov}(|S_{N,j_1} - S_{N,k_1}|^a, |S_{N,j_2} - S_{N,k_2}|^a) = 0$ whenever $\{j_1,k_1\} \cap \{j_2,k_2\} = \emptyset$, we obtain

$$\mathrm{Var}(\tilde{T}_N) = \frac{1}{(N(N-1))^2} \sum_{1\le j_1 \ne k_1 \le N} \sum_{1\le j_2\ne k_2\le N} \mathrm{Cov}(|S_{N,j_1} - S_{N,k_1}|^a, |S_{N,j_2} - S_{N,k_2}|^a)$$
$$= \frac{1}{N(N-1)}\mathrm{Var}(|S_{N,1} - S_{N,2}|^a) + \frac{2(N-2)}{N(N-1)}\mathrm{Cov}(|S_{N,1} - S_{N,2}|^a, |S_{N,1} - S_{N,3}|^a).$$

The Cauchy-Schwarz inequality thus allows us to bound $\mathrm{Var}(\tilde{T}_N)$ from above:

$$\mathrm{Var}(\tilde{T}_N) \le \frac{2N-3}{N(N-1)}\mathrm{Var}(|S_{N,1} - S_{N,2}|^a).$$

To further bound $\mathrm{Var}(|S_{N,1} - S_{N,2}|^a)$ from above, we once more distinguish the cases $a \in (0,1)$ and $a \in [1,2)$. For $a \in [1,2)$, the c_r-inequality and the Rosenthal inequality imply

$$\mathrm{Var}(|S_{N,1} - S_{N,2}|^a) \le E\left(|S_{N,1} - S_{N,2}|^{2a}\right) \le 2^{2a}E\,|S_{N,1}|^{2a} = 2^{2a}E\left|\frac{1}{\sqrt{\tau_N}}\sum_{i=1}^{\tau_N}((X_i/\sigma)^2 - 1)\right|^{2a}$$
$$\le 2^{2a}\left[a_1(a)\tau_N^{1-a}E|(X_1/\sigma)^2 - 1|^{2a} + a_2(a)\gamma^{2a}\right].$$

In the other case, $a \in (0,1)$, the c_r-inequality and Jensen's inequality allow us to deduce

$$\mathrm{Var}(|S_{N,1} - S_{N,2}|^a) \le E\left(|S_{N,1} - S_{N,2}|^{2a}\right) \le \max(2, 2^{2a})E\left|\frac{1}{\sqrt{\tau_N}}\sum_{i=1}^{\tau_N}((X_i/\sigma)^2 - 1)\right|^{2a}$$
$$\le \max(2, 2^{2a})\left(E\left(\frac{1}{\sqrt{\tau_N}}\sum_{i=1}^{\tau_N}((X_i/\sigma)^2 - 1)\right)^2\right)^a = \max(2, 2^{2a})\gamma^{2a}.$$

Thus in both cases $\mathrm{Var}(\tilde{T}_N) \to 0$ is fulfilled, which proves the claim. □

Proof of Proposition 2.5. The proposition is basically a CLT for the U-statistic $\tilde{T}_N/\gamma^a$, where we center by its expectation $E(\tilde{T}_N/\gamma^a)$ instead of the corresponding limit $2^{\frac{a}{2}}E|Z|^a$, see Lemma 2.6. Defining $h(x,y) = |x-y|^a$ for $x, y \in \mathbb{R}$, we express $\tilde{T}_N/\gamma^a$ as

$$\tilde{T}_N/\gamma^a = \frac{1}{N(N-1)} \sum_{1 \le j \ne k \le N} h(S_{N,j}/\gamma, S_{N,k}/\gamma).$$

For any $N \in \mathbb{N}$ the Hoeffding decomposition of h is given by $\theta^{(N)} = Eh(S_{N,1}/\gamma, S_{N,2}/\gamma)$, where $g_1^{(N)}(x) = Eh(x, S_{N,1}/\gamma)$, $h_1^{(N)}(x) = g_1^{(N)}(x) - \theta^{(N)}$ and $h_2^{(N)}(x,y) = h(x,y) - \theta^{(N)} - h_1^{(N)}(x) - h_1^{(N)}(y)$. This allows us to express $\tilde{T}_N/\gamma^a$ as

$$\begin{aligned}
&\frac{1}{N(N-1)} \sum_{1 \le j \ne k \le N} h(S_{N,j}/\gamma, S_{N,k/\gamma}) \\
=&\theta^{(N)} + \frac{1}{N(N-1)} \sum_{1 \le j \ne k \le N} \left(h_1^{(N)}(S_{N,j}/\gamma) + h_2^{(N)}(S_{N,k}/\gamma) \right) + \frac{1}{N(N-1)} \sum_{1 \le j \ne k \le N} h_2^{(N)}(S_{N,j}/\gamma, S_{N,k}/\gamma) \\
=&\theta^{(N)} + \underbrace{\frac{2}{N} \sum_{1 \le j \le N} h_1^{(N)}(S_{N,j}/\gamma)}_{=a_N} + \underbrace{\frac{1}{N(N-1)} \sum_{1 \le j \ne k \le N} h_2^{(N)}(S_{N,j}/\gamma, S_{N,k}/\gamma)}_{=b_N}.
\end{aligned}$$

To prove the claim, we thus need to show $\sqrt{N}\frac{a_N}{2} \longrightarrow N(0, \frac{\tilde{\sigma}^2}{4})$ in distribution and $\sqrt{N} b_N \longrightarrow 0$ as $n \to \infty$.

Let us first establish the asymptotic normality of $\sqrt{N}\frac{a_N}{2}$ using the CLT of Lyapunov. Because of $\sqrt{N}\frac{a_N}{2} = \frac{1}{\sqrt{N}} \sum_{1 \le j \le N} h_1^{(N)}(S_{N,j}/\gamma)$ and $Eh_1^{(N)}(S_{N,1}/\gamma) = 0$, it is sufficient to prove that for some $\delta > 0$ and some constant $\sigma' > 0$ it holds

$$\sum_{1 \le j \le N} E\left| \frac{h_1^{(N)}(S_{N,j}/\gamma)}{\sqrt{N}} \right|^{2+\delta} \to 0 \text{ and } \sum_{1 \le j \le N} \mathrm{Var}\left(\frac{h_1^{(N)}(S_{N,j}/\gamma)}{\sqrt{N}} \right) \to \sigma'.$$

We start by investigating the first condition. Since for any fixed $N \in \mathbb{N}$ the $h_1^{(N)}(S_{N,j})$, $j = 1, \dots, N$, are identically distributed, the c_r-inequality allows us to deduce

$$\begin{aligned}
\sum_{1 \le j \le N} E\left| \frac{h_1^{(N)}(S_{N,j}/\gamma)}{\sqrt{N}} \right|^{2+\delta} &= \frac{1}{N^{1+\frac{\delta}{2}}} \sum_{1 \le j \le N} E\left| h_1^{(N)}(S_{N,j}/\gamma) \right|^{2+\delta} = \frac{1}{N^{\frac{\delta}{2}}} E\left| h_1^{(N)}(S_{N,1}/\gamma) \right|^{2+\delta} \\
&= \frac{1}{N^{\frac{\delta}{2}}} E\left| g_1^{(N)}(S_{N,1}/\gamma) - \theta^{(N)} \right|^{2+\delta} \le \frac{2^{1+\delta}}{N^{\frac{\delta}{2}}} \left(E\left| g_1^{(N)}(S_{N,1}/\gamma) \right|^{2+\delta} + \left(\theta^{(N)}\right)^{2+\delta} \right).
\end{aligned}$$

Denoting by F_N the distribution of $S_{N,1}/\gamma$, Jensen's inequality allows us to bound the relevant term from above:

$$\begin{aligned}
E\left| g_1^{(N)}(S_{N,1}/\gamma) \right|^{2+\delta} &= \int \left| \int |x-y|^a dF_N(y) \right|^{2+\delta} dF_N(x) \\
\le \int\int |x-y|^{a(2+\delta)} dF_N(y) dF_N(x) &= E|(S_{N,1} - S_{N,2})/\gamma|^{a(2+\delta)}.
\end{aligned}$$

Putting the pieces together, we thus get

$$\sum_{1 \le j \le N} E\left| \frac{h_1^{(N)}(S_{N,j}/\gamma)}{\sqrt{N}} \right|^{2+\delta} \le \frac{2^{1+\delta}}{N^{\frac{\delta}{2}}} \left(E|(S_{N,1} - S_{N,2})/\gamma|^{a(2+\delta)} + \left(\theta^{(N)}\right)^{2+\delta} \right). \quad (3)$$

Another application of the c_r-inequality time yields

$$E|(S_{N,1}-S_{N,2})/\gamma|^{a(2+\delta)} \leq \max(2, 2^{a(2+\delta)})E|S_{N,1}/\gamma|^{a(2+\delta)}.$$

In case of $a \in [1,2)$, it holds that $a(2+\delta) \geq 2$. The Rosenthal inequality then leads to

$$E|S_{N,1}/\gamma|^{a(2+\delta)} \leq a_1(a(2+\delta))\tau_N^{1-\frac{a}{2}(2+\delta)}E|((X_1/\sigma)^2-1)/\gamma|^{a(2+\delta)} + a_2(a(2+\delta)),$$

which is bounded by assumption. For the case $a \in (0,1)$ and a sufficiently small $\delta > 0$, we get $a(2+\delta) < 2$ and thus

$$E|S_{N,1}/\gamma|^{a(2+\delta)} = E|S_{N,1}/\gamma|^{2\frac{a}{2}(2+\delta)} \leq \left(E|S_{N,1}/\gamma|^2\right)^{\frac{a}{2}(2+\delta)} = 1$$

by Jensen's inequality. Thus $E|(S_{N,1}-S_{N,2})/\gamma|^{a(2+\delta)}$ is bounded for any $a \in (0,2)$. By Lemma 2.6, $\theta^{(N)}$ converges to a limit, due to $\frac{N}{\tau_n} \to 0$. Hence $\theta^{(N)}$ is bounded. Using (3) we therefore have verified the first condition of the CLT of Lyapunov:

$$\sum_{1\leq j\leq N} E\left|\frac{h_1^{(N)}(S_{N,j}/\gamma)}{\sqrt{N}}\right|^{2+\delta} \longrightarrow 0.$$

We now turn to the second CLT condition, so we need to check whether the sum $\sum_{1\leq j\leq N} \mathrm{Var}\left(h_1^{(N)}(S_{N,j}/\gamma)/\sqrt{N}\right)$ converges to some constant $\sigma' > 0$. Note that for any sequence $(x_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ with $x_n \to x$, the distribution of the random variables $S_{N,j}/\gamma - x_n$ converges to the normal distribution with mean $-x$ and unit variance. We can therefore apply the continuous mapping theorem for sequences of functions, yielding

$$\sum_{1\leq j\leq N} \mathrm{Var}\left(\frac{h_1^{(N)}(S_{N,j}/\gamma)}{\sqrt{N}}\right) = \mathrm{Var}\left(h_1^{(N)}(S_{N,1}/\gamma)\right) \longrightarrow \mathrm{Var}\left(h_1(Z)\right) > 0,$$

where $h_1(x) = E(h(x,Z'))$ and Z, Z' are i.i.d. standard Gaussian random variables. Thus we have established $\sqrt{N}\frac{a_N}{2} \longrightarrow N(0, \mathrm{Var}\left(h_1(Z)\right)$ in distribution using Lyapunov's CLT.
To prove the proposition, it remains to show $\sqrt{N}b_N \longrightarrow 0$. Since $E(b_N) = 0$, it suffices to consider the variance

$$\mathrm{Var}\left(\frac{\sqrt{N}}{N(N-1)}\sum_{1\leq j\neq k\leq N} h_2^{(N)}(S_{N,j}/\gamma, S_{N,k}/\gamma)\right) = \frac{1}{N-1}\mathrm{Var}\left(h_2^{(N)}(S_{N,1}/\gamma, S_{N,2}/\gamma)\right).$$

The above equality holds because the summands are uncorrelated by the construction of $h_2^{(N)}$ and are identically distributed. We further get

$$
\begin{aligned}
&\operatorname{Var}\left(h_2^{(N)}(S_{N,1}/\gamma, S_{N,2}/\gamma)\right) \le E\left(\left(h_2^{(N)}(S_{N,1}/\gamma, S_{N,2}/\gamma)\right)^2\right) \\
=&E\left(\left(h(S_{N,1}/\gamma, S_{N,2}/\gamma) - \theta^{(N)} - h_1^{(N)}(S_{N,1}/\gamma) - h_1^{(N)}(S_{N,2}/\gamma)\right)^2\right) \\
=&E\left(h(S_{N,1}/\gamma, S_{N,2}/\gamma)^2\right) - \left(\theta^{(N)}\right)^2 - 2E\left(h(S_{N,1}/\gamma, S_{N,2}/\gamma)h_1^{(N)}(S_{N,1}/\gamma)\right) - \left(\theta^{(N)}\right)^2 + \left(\theta^{(N)}\right)^2 \\
&+2\theta^{(N)}E\left(h_1^{(N)}(S_{N,1}/\gamma)\right) + 2\left[-E\left(h(S_{N,1}/\gamma, S_{N,2}/\gamma)h_1^{(N)}(S_{N,1}/\gamma)\right) + \theta^{(N)}E\left(h_1^{(N)}(S_{N,1}/\gamma)\right)\right. \\
&\quad\left. +E\left(\left(h_1^{(N)}(S_{N,1}/\gamma)\right)^2\right) + E\left(h_1^{(N)}(S_{N,1}/\gamma)\right)^2\right] \\
=&E\left(h(S_{N,1}/\gamma, S_{N,2}/\gamma)^2\right) - \left(\theta^{(N)}\right)^2 - 4E\left(h(S_{N,1}/\gamma, S_{N,2}/\gamma)h_1^{(N)}(S_{N,1}/\gamma)\right) + 2E\left(\left(h_1^{(N)}(S_{N,1}/\gamma)\right)^2\right) \\
=&E\left(h(S_{N,1}/\gamma, S_{N,2}/\gamma)^2\right) - \left(\theta^{(N)}\right)^2 - 2E\left(h(S_{N,1}/\gamma, S_{N,2}/\gamma)h_1^{(N)}(S_{N,1}/\gamma)\right) \\
\le&E\left(h(S_{N,1}/\gamma, S_{N,2}/\gamma)^2\right) = E\left(|(S_{N,1} - S_{N,2})/\gamma|^{2a}\right),
\end{aligned}
$$

where the last expression is bounded from above as shown in the proof of Proposition 2.4. Thus, $\operatorname{Var}\left(h_2^{(N)}(S_{N,1}/\gamma, S_{N,2}/\gamma)\right)$ is bounded and therefore $\sqrt{N}b_N$ converges to zero, which completes the proof.

□

Proof of Lemma 2.6. First, note that $E(\tilde{T}_N/\gamma^a) = E|(S_{N,1}-S_{N,2})/\gamma|^a$ and $2^{\frac{a}{2}}E|Z|^a = E|Z-Z'|^a$, where Z' is a standard Gaussian random variable independent of Z. We thus consider the standardized differences $\frac{1}{\gamma\sqrt{2}}(S_{N,1} - S_{N,2})$ and $\frac{1}{\sqrt{2}}(Z - Z')$ and define their distribution functions by

$$F_N(x) = P\left(\frac{1}{\gamma\sqrt{2}}(S_{N,1} - S_{N,2}) \le x\right) \quad \text{and} \quad \Phi(x) = P\left(\frac{1}{\sqrt{2}}(Z - Z') \le x\right).$$

By the Berry-Esseen theorem, we have $\Delta_N = \sup_{x\in\mathbb{R}} |F_N(x) - \Phi(x)| \le \frac{c_1}{\sqrt{\tau_N}}$ for some constant $c_1 \in \mathbb{R}^+$. By Theorem 11, Chapter V in Petrov (1975), this implies for $N \ge N_0$ sufficiently large the following non-uniform bound for the speed of convergence for any $x \in \mathbb{R}$ and appropriate constants $c_2, c_3 \in \mathbb{R}$:

$$|F_N(x) - \Phi(x)| \le c_2\frac{\Delta_N \log(\Delta_N^{-1})}{1 + x^2} \le c_3\frac{\tau_N^{-\frac{1}{2}} \log(\tau_N)}{1 + x^2}. \tag{4}$$

We now consider the standardized absolute differences $\frac{1}{\gamma\sqrt{2}}|S_{N,1} - S_{N,2}|$ and $\frac{1}{\sqrt{2}}|Z - Z'|$ and their distribution functions

$$\tilde{F}_N(x) = P\left(\frac{1}{\gamma\sqrt{2}}|S_{N,1} - S_{N,2}| \le x\right) \quad \text{and} \quad \tilde{\Phi}(x) = P\left(\frac{1}{\sqrt{2}}|Z - Z'| \le x\right).$$

From (4) we obtain

$$|\tilde{F}_N(x) - \tilde{\Phi}(x)| \le 2c_3\frac{\tau_N^{-\frac{1}{2}} \log(\tau_N)}{1 + x^2}. \tag{5}$$

Using integration by parts we have

$$E(\tilde{T}_N/\gamma^a) = E|(S_{N,1} - S_{N,2})/\gamma|^a = 2^{\frac{a}{2}} \int_0^\infty x^a d\tilde{F}_N(x) = a2^{\frac{a}{2}} \int_0^\infty (1 - \tilde{F}_N(x))x^{a-1}dx,$$

and similarly $2^{\frac{a}{2}} E|Z|^a = E|Z - Z'|^a = a2^{\frac{a}{2}} \int_0^\infty (1 - \tilde{\Phi}(x))x^{a-1}dx$. Finally, we get

$$\begin{aligned}\left|E(\tilde{T}_N/\gamma^a) - 2^{\frac{a}{2}} E|Z|^a\right| =& a2^{\frac{a}{2}} \left|\int_0^\infty (\tilde{F}_N(x) - \tilde{\Phi}(x))x^{a-1}dx\right| \\ \overset{(5)}{\leq}& a2^{\frac{a}{2}} c_3 \tau_N^{-\frac{1}{2}} \log \tau_N \int_0^\infty \frac{x^{a-1}}{1+x^2}dx = \mathcal{O}\left(\frac{\log \tau_N}{\sqrt{\tau_N}}\right)\end{aligned}$$

since the integral $\int_0^\infty \frac{x^{a-1}}{1+x^2}dx$ is finite for $a \in [1, 2)$. This implies the proposition.

□

Lemma 6.1 (Explicit representation of the limit in Theorem 2.1)**.** *Since*

$$\begin{aligned}E(|Z|^a) =& \int_{-\infty}^\infty \frac{|z|^a}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz = 2 \int_0^\infty \frac{z^a}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz \\ =& 2 \int_0^\infty \frac{\sqrt{2y}^{a-1}}{\sqrt{2\pi}} \exp(-y)\, dy = \frac{2^{\frac{a+1}{2}}}{\sqrt{2\pi}} \int_0^\infty y^{\frac{a-1}{2}} \exp(-y)\, dy \\ =& \frac{2^{\frac{a}{2}}}{\sqrt{\pi}} \int_0^\infty y^{\frac{a+1}{2}-1} \exp(-y)\, dy = \frac{2^{\frac{a}{2}}}{\sqrt{\pi}} \Gamma\left(\frac{a+1}{2}\right),\end{aligned}$$

we get $2^{\frac{a}{2}} E(|Z|^a) = \frac{2^a}{\sqrt{\pi}} \Gamma\left(\frac{a+1}{2}\right)$ *for the limit in Theorem 2.1.*

Lemma 6.2.

(i) For all $a \in (0, 1)$ *and* $x, y \in \mathbb{R}$ *we have*

$$||x|^a - |y|^a| \leq |x - y|^a. \tag{6}$$

(ii) For all $a \in (1, 2)$ $x, y \in \mathbb{R}$ *with* $|x|, |y| < 1$ *it holds*

$$||x|^a - |y|^a| \leq a|x - y|. \tag{7}$$

Note that, in the limit $a = 1$*, both inequalities give the same bound which holds by the triangle inequality for any* $x, y \in \mathbb{R}$*.*

Proof. (i) To show the proposition, it suffices to establish the nonnegativity of the function $f_x : [0, x] \to \mathbb{R}$ defined by $f_x(y) = (x - y)^a + y^a - x^a$ for any $x \in \mathbb{R}^+$. Note that the second derivative of f_x is $f_x''(y) = a(a - 1)(x - y)^{a-2} + a(a - 1)y^{a-2} \leq 0$ for all $0 \leq y \leq x$. Hence, the first derivative of f_x is nonincreasing. Combined with $f_x(0) = f_x(x) = 0$, this shows the nonnegativity of f_x and therefore inequality (6).

(ii) In a similar way, we prove the proposition showing the nonnegativity of the function $g_x : [0, x] \to \mathbb{R}$ defined by $g_x(y) = a(x - y) + y^a - x^a$ for any $x \in [0, 1)$. The first derivative of g_x is $g_x'(y) = -a + ay^{a-1}$ and thus is never zero on $(0, x)$ because of $|y| < 1$. Thus g_x does not attain any minimum on $(0, x)$. Together with $g_x(x) = 0 \geq 0$ and $g_x(0) = ax - x^a \geq 0$, one can thus deduce the nonnegativity of g_x and therefore show inequality (7).

□

References

J. Chen, A. K. Gupta (1997): Testing and Locating Variance Changepoints with Application to Stock Prices. *Journal of the American Statistical Association* 92 (438), 739–747.

C. S. J. Chu, K. Hornik, C. M. Kaun (1995): MOSUM tests for parameter constancy. *Biometrika* 82 (3), 603–617.

L. Davies, C. Höhenrieder, W. Krämer (2012): Recursive Computation of Piecewise Constant Volatilities. *Computational Statistics & Data Analysis* 56 (11), 3623–3631.

R. Fried (2012): On the Online Estimation of Local Constant Volatilities. *Computational Statistics & Data Analysis* 56 (11), 3080–3090.

C. Gerstenberger, D. Vogel (2015): On the efficiency of Gini's mean difference. *Statistical Methods & Applications* 24 (4), 569–596.

D. A. Hsu (1977): Tests for Variance Shift at an Unknown Time Point. *Journal of the Royal Statistical Society* 26 (3),279–284.

V. K. Jandhyala, S. B. Fotopoulos, D. M. Hawkins (2002): Detection and estimation of abrupt changes in the variability of a process. *Computational Statistics and Data Analysis* 40 (1), 1–19.

D. S. Matteson, N. A. James (2014): A Nonparametric Approach for Multiple Change Point Analysis of Multivariate Data. *Journal of the American Statistical Association* 109 (505), 334–345.

D. Mercurio, V. Spokoiny (2004): Statistical Inference for Time-inhomogeneous Volatility Models. *The Annals of Statistics* 32 (2), 577–602.

A. M. Mood (1954): On the Asymptotic Efficiency of Certain Nonparametric Two-Sample Tests. *The Annals of Mathematical Statistics* 25 (3), 514–522.

V. V. Petrov (1975): *Sums of Independent Random Variables.* Springer Verlag, Berlin.

R Core Team (2016): R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing*, Vienna, Austria, URL https://www.R-project.org/.

G. J. Ross (2013): Modelling Financial Volatility in the Presence of Abrupt Changes. *Physica A: Statistical Mechanics and its Applications* 392 (2), 350–360.

G. R. Shorack, J. A. Wellner (1986): *Empirical Processes with Applications to Statistics.* John Wiley & Sons, New York.

V. Spokoiny (2009): Multiscale Local Change Point Detection with Applications to Value-at-risk. *The Annals of Statistics* 37 (3), 1405–1436.

C. Stărică, C. Granger (2005): Nonstationarities in Stock Returns. *Review of Economics and Statistics* 87 (3), 503–522.

E. Vassiliou, I. C. Demetriou (2005): An Adaptive Algorithm for Least Squares Piecewise Monotonic Data Fitting. *Computational Statistics & Data Analysis* 49 (2), 591–609.

D. Wied, M. Arnold, N. Bissantz, D. Ziggel (2012): A New Fluctuation Test for Constant Variances with Applications to Finance. *Metrika* 75 (8), 1111–1127.

M. Wornowizki, R. Fried, S. G. Meintanis (2016): Fourier Methods for Analysing Piecewise Constant Volatilities. *SFB 823 Discussion Paper*, 43/2016, DOI 10.17877/DE290R-17225

A. Zeileis, F. Leisch, K. Hornik, C. Kleiber (2002): strucchange: An R Package for Testing for Structural Change in Linear Regression Models. *Journal of Statistical Software* 7 (2), 1–38, URL http://www.jstatsoft.org/v07/i02/.