

Fourier methods for analyzing piecewise constant volatilities

Max Wornowizki¹ · Roland Fried¹ ·
Simos G. Meintanis^{2,3}

Received: 30 September 2015 / Accepted: 24 January 2017 / Published online: 3 February 2017
© Springer-Verlag Berlin Heidelberg 2017

Abstract We develop procedures for testing whether a sequence of independent random variables has constant variance. If this is fulfilled, the modulus of a Fourier-type transformation of the volatility process is identically equal to one. Our approach takes advantage of this property considering a canonical estimator for the modulus under the assumption of piecewise identically distributed zero mean observations. Using block-wise variance estimation, we introduce several test statistics resulting from different weight functions. All of them are given by simple explicit formulae. We prove the consistency of the corresponding tests and compare them to alternative procedures on extensive Monte Carlo experiments. According to the results, our proposals offer fairly high power, particularly in the case of multiple structural breaks. They also allow for an adequate estimation of the change point positions. We apply our procedure to gold mining data and also briefly discuss how it can be modified to test for the stationarity of other distributional parameters.

Keywords Change point analysis · Variance · Piecewise identical distribution · Independence · Weight function

Simos G. Meintanis: On sabbatical leave from the University of Athens.

✉ Roland Fried
fried@statistik.tu-dortmund.de

¹ Department of Statistics, TU Dortmund University, Dortmund, Germany

² Department of Economics, National and Kapodistrian University of Athens, Athens, Greece

³ Unit for Business Mathematics and Informatics, North-West University, Potchefstroom, South Africa

1 Introduction

Let us consider a real-valued sequence of independent random variables X_t corresponding to the times $t = 1, \dots, n$ for some $n \in \mathbb{N}$. In addition, let $\sigma^2(t)$ denote the variance of X_t for each $t = 1, \dots, n$. We are interested in testing whether the random variables under study have a constant volatility as opposed to the alternative of one or several changes of the variance. This is expressed by the hypotheses

$$\mathbb{H}_0 : \forall t, t' = 1, \dots, n : \sigma^2(t) = \sigma^2(t') \text{ vs. } \mathbb{H}_1 : \exists t, t' = 1, \dots, n : \sigma^2(t) \neq \sigma^2(t').$$

Testing for changes in the volatility is a topic studied for the most part in the last forty years. First papers on this problem often rely on distributional assumptions in order to derive appropriate decision rules. For example, for the Gaussian case [Hsu \(1977\)](#) constructs a test using cumulative sums of χ^2 -type random variables, while [Jandhyala et al. \(2002\)](#) and others propose likelihood based procedures. To weaken the distributional assumptions later articles such as [Wied et al. \(2012\)](#) consider asymptotic CUSUM-type tests, while [Ross \(2013\)](#) and others introduce nonparametric alternatives. It should be noted that most of these methods split the sample into two parts and are therefore constructed to detect at most one change at a time. This can be quite problematic in cases where several change points lead to comparable variances for any split of the data in two parts.

The method proposed in the following tries to circumvent this problem by using a test statistic calculated from several nonoverlapping blocks of data. It performs quite well on data with several structural breaks, but is also competitive in case of only one change point. In addition, it does not make any assumptions on the distribution of the data and keeps the significance level for any sample size.

We adopt a framework within which marginal distributional features are locally stationary, but globally nonconstant. A blockwise constant volatility has drawn attention in [Mercurio and Spokoiny \(2004\)](#), [Davies et al. \(2012\)](#) and [Fried \(2012\)](#), among others, but there are also more general frameworks such as that of blockwise local stationarity suggested in [Guégan \(2015\)](#). The assumption of a blockwise constant variance can be formulated using some specified time points $0 = t_0 < t_1 < \dots < t_N = n$. They are understood as possible change point positions. If external knowledge allowing us to choose them appropriately is not available, one can select t_0, \dots, t_N equidistantly. The possible change point positions correspond to important events which may trigger an upward or downward change in the volatility. The values of the volatility process $\sigma^2(\cdot)$ are thus allowed to differ for some of the time blocks $B_j = \{t_{j-1} + 1, t_{j-1} + 2, \dots, t_j\}$, $j = 1, \dots, N$. Within each time block the volatility is assumed to be approximately constant.

In accordance with this model, we work under the assumption of independent zero mean random variables with identical distribution up to scale. In other words, we assume that the relation $P(X_t \leq x) = F\left(x/\sqrt{\sigma^2(t)}\right)$ holds for some unknown but fixed distribution function F for $t = 1, \dots, n$ and $x \in \mathbb{R}$. The centered Gaussian distribution is thus contained as a special case, but heavy tails are also included. The zero mean assumption is justifiable when dealing with returns or similar data obtained

from differences of consecutive observations. In other cases, it can be relaxed to blockwise constant means with known block structure, so that zero mean data results from preprocessing. For a detailed discussion of the other assumptions in the context of volatility, see for example [Spokoiny \(2009\)](#) and the references given therein.

The paper is structured as follows: in Sect. 2 a new statistic for testing changes in the volatility process is motivated. The corresponding test is introduced in Sect. 3, along with alternative procedures available in the literature. In Sect. 4, all methods presented before are compared in several simulation scenarios. Hereafter, the best methods are applied to quality control data concerning face sampling in South African gold mines. Section 5 summarizes the main results and provides an outlook on possible future work. A proof for the consistency of our test procedure is given in the Appendix.

2 Test statistics

In this section, we introduce a class of test statistics for testing the constancy of the volatility process. We then derive explicit representations for some members of this class.

2.1 Motivation

A reasonable first step in order to test \mathbb{H}_0 is the estimation of the local volatilities. Since all random variables are assumed to have zero mean, a natural estimator for the variance within the j -th block B_j , $j = 1, \dots, N$, is given by

$$\hat{\sigma}_j^2 = \frac{1}{\tau_j} \sum_{t \in B_j} X_t^2. \tag{1}$$

Thereby, $\tau_j = |B_j|$ denotes the number of random variables corresponding to the j -th block B_j , $j = 1, \dots, N$. These block lengths must be sufficiently large to ensure a reasonable accuracy of the variance estimations.

Instead of considering the estimated volatilities themselves, we rather work with their logarithms. This allows us to construct scale independent test procedures, as we demonstrate in the following. Under the null hypothesis of constant variance, the logarithmized volatility process $\log(\sigma^2(\cdot))$ constantly equals $\log(\sigma^2)$ for some unknown $\sigma^2 > 0$. In this case, the function $\varphi : \mathbb{R} \times \{1, \dots, n\} \rightarrow \mathbb{C}$ defined by $\varphi(u, t) = e^{iu \log(\sigma^2(t))}$, where $i = \sqrt{-1}$, does not depend on t . Hence, for any $t = 1, \dots, n$ it can be estimated by

$$\varphi_n(u) = \sum_{j=1}^N \frac{\tau_j}{n} e^{iu \log(\hat{\sigma}_j^2)} \tag{2}$$

in a straightforward way. Here, each $\tau_j/n = \tau_j/(\sum_{i=1}^N \tau_i)$ weights the corresponding summand according to the number of observations in the j -th block for $j = 1, \dots, N$.

If all blocks contain the same number of observations, all weights are equal to $\frac{1}{N}$. The applied transformation $f : \mathbb{R} \rightarrow \mathbb{C}$, $f(x) = e^{iux}$ is closely related to the Fourier transform and characteristic functions. In our situation, it has the following nice and intuitive behavior: under \mathbb{H}_0 , the blockwise estimations are closer to each other than under the alternative. Therefore, they are mapped to points on the unit circle close to each other. Consequently, the weighted mean $\varphi_n(u)$ lies relatively close to the unit circle for every $u \in \mathbb{R}$ and thus has a modulus near one. Under the alternative, the logarithmized blockwise estimations differ more than under \mathbb{H}_0 . They are thus mapped to comparatively distant points on the unit circle for most $u \in \mathbb{R}$. Hence, for most $u \in \mathbb{R}$, the weighted mean $\varphi_n(u)$ is closer to the origin than under the null hypothesis. Because of that, $\varphi_n(u)$ has a small modulus under the alternative. This behavior is related to the well-known property that the modulus of a characteristic function takes on its maximum value one, identically in u , if and only if the underlying random variable is degenerate; see Lo eve (1977, §14.1). In view of this fact, we propose to use statistics of the form

$$V_n = \int \left(1 - |\varphi_n(u)|^2\right) w(u) du$$

to test for global constant volatility, whereby here and henceforth integration is meant over the entire real line. V_n is nonnegative because of $|\varphi_n(u)|^2 \leq 1$. The weight function $w : \mathbb{R} \rightarrow \mathbb{R}_0^+$ is chosen such that a finite V_n is ensured. More details on the weight function along with explicit representations of V_n for several choices of w are given in the next subsection.

To the best of our knowledge, the idea of utilizing the property of a Fourier transform taking on its maximum modulus constantly over u is new in the literature. Nevertheless, similar transformations and characteristic functions have been used before for change point detection, and to a good effect; see for instance Huřkova and Meintanis (2006a, b), Hlavka et al. (2012), and Hlavka et al. (2015). Specifically in these papers, it is shown that approaches based on Fourier-type transformations are convenient from the computational point of view, lead to theoretically sound asymptotics, and are competitive compared to more classical procedures. We want to stress the fact that the aforementioned Fourier-type tests are constructed by comparing the two subsamples before and after each candidate change point using two-sample techniques (Meintanis 2005). This approach is based on the implicit assumption of only one or at least one dominant change point in the data. If the latter assumption does not hold, the corresponding methods may lose a considerable amount of power. As opposed to this, our test is designed to handle multiple structural changes by splitting the data set into several subsamples.

2.2 Explicit representation and weight functions

In this section, we derive a representation of V_n allowing to calculate it in a convenient way. To handle the integral figuring in V_n , we work with integrable weight functions w and set $W = \int w(u) du < \infty$. Then, V_n can be rewritten to

$$V_n = W - \int |\varphi_n(u)|^2 w(u) du.$$

Since the constant W does not depend on the data, it can be dropped for testing. Using the definition of φ in (2), the integral above reduces to

$$T_{\text{Four}} = \int |\varphi_n(u)|^2 w(u) du = \frac{1}{n^2} \sum_{j,k=1}^N \tau_j \tau_k I_w \left(\log \left(\widehat{\sigma}_j^2 \right) - \log \left(\widehat{\sigma}_k^2 \right) \right), \quad (3)$$

where

$$I_w(x) = \int \cos(ux) w(u) du.$$

Since small values of V support the null hypothesis, \mathbb{H}_0 is rejected for small values of T_{Four} . This test statistic depends on the data only via the terms $\log \left(\widehat{\sigma}_j^2 \right) - \log \left(\widehat{\sigma}_k^2 \right) = \log \left(\widehat{\sigma}_j^2 / \widehat{\sigma}_k^2 \right)$, $1 \leq j \neq k \leq N$. Therefore, thanks to taking the logarithm, any scale factor is canceled out and thus T_{Four} is scale invariant. For that reason, we propose to use the logarithmized estimated volatilities instead of the estimated volatilities themselves.

The function I_w can be expressed explicitly for several standard choices of w . These are the uniform, the Laplace and the Gaussian weighting with corresponding weight functions

$$w_U(u) = \mathbf{1}_{(-a,a)}(u), \quad w_L(u) = e^{-a|u|} \quad \text{and} \quad w_G(u) = e^{-au^2},$$

respectively. All of them depend on a parameter $a > 0$. Straightforward computations lead to

$$I_{w_U}(x) = \frac{2 \sin(ax)}{x}, \quad I_{w_L}(x) = \frac{2a}{a^2 + x^2}, \quad \text{and} \quad I_{w_G}(x) = \sqrt{\frac{\pi}{a}} \exp \left(\frac{-x^2}{4a} \right)$$

for the uniform, Laplace and Gaussian weighting, respectively (Hušková and Meintanis 2006a). Thereby, $I_{w_U}(0)$ is defined by its limit $\lim_{x \rightarrow 0} \frac{2 \sin(ax)}{x} = 2a$.

There also exist alternative choices for w . One example is the data adaptive weighting scheme proposed by Meintanis et al. (2014) for goodness-of-fit testing. Another weight function is studied in Matteson and James (2014) in the context of multivariate nonparametric detection of general distributional changes. Both weight functions were considered in our simulations. The corresponding results are not included in this paper, since, in accordance with our simulations in Sect. 4.1, the choice of the weight function does not seem to have a large impact on the performance of our tests. In particular, the data adaptive weighting leads to slightly worse and the weighting proposed by Matteson and James to essentially the same rejection rates as the three standard weight functions introduced above. These results match earlier work suggesting that the performance of similar Fourier-type procedures does not depend much on

the particular functional form of w ; see for instance [Hušková and Meintanis \(2006a\)](#) and [Hlávka et al. \(2012\)](#).

In general, the parameter a figuring in the weight function might also influence the method's results. For the weight functions presented before, a large value of a lets w decay more sharply while the opposite is true for smaller values of a . In the former case (resp. latter case), more emphasis is placed on $\widehat{\varphi}(u)$ for u near the origin (resp. away from the origin). Note that the behavior of a characteristic function near the origin reflects the tail behavior of the underlying distribution. Therefore, caution must be exercised in choosing a . If a is chosen too large, then we might overemphasize the tail of the underlying law, whereas in the opposite case of a being too small, these tails will be obscured; see for instance [Epps \(1993\)](#). Other than that, the whole issue of making an educated guess for a is typically highly technical. It has been theoretically investigated only under very stringent parametric assumptions both about the hypothesis being tested as well as about the direction of possible deviation from this hypothesis; see [Epps \(1999\)](#) and [Tenreiro \(2009\)](#). Fortunately, the simulations in [Sect. 4.1](#) illustrate that our test seems not to be influenced much by the choice of a , so that this problem is alleviated.

3 Testing for constant volatility

In this section, we first show how the hypothesis of global constant volatility can be tested using the statistics defined in [Sect. 2](#). Hereafter, we present a natural estimator of the structural break position in case of rejection is defined. The procedure allows for the location of several presumable structural break positions. The section closes by briefly introducing four alternative methods for the testing problem taken from or inspired by the literature.

3.1 Testing procedure

The distribution of our test statistic strongly depends on the distribution of the random variables X_1, \dots, X_n . Getting critical values without imposing distributional assumptions is thus not possible, at least for small sample sizes. In such situations, resampling strategies are often of great help. Since under the null hypothesis the X_1, \dots, X_n are independent and identically distributed, we propose to test the stationarity of the variance using the permutation principle introduced by [Fisher \(1935\)](#). Research in various fields shows that this approach can lead to quite powerful tests; see [Good \(2005\)](#) for a monograph treatment with applications to various fields and [Hušková and Meintanis \(2006b\)](#) in the context of characteristic functions. In our setting, the method is applied as follows: given the complete original sample, first generate p new samples by randomly permuting the observations p times. Then, determine the test statistic T_{Four} for each of the $p + 1$ samples assuming that the data was observed in the respective order. Thereby, the block lengths τ_1, \dots, τ_N as well as N , w and a are the same for each computation. Under the null hypothesis, the $p + 1$ test statistics are observations of identically distributed random variables. Thus, the permutation test rejects \mathbb{H}_0 at the predefined significance level α , if the test statistic determined on the original sample

falls below the empirical α -quantile of all $p + 1$ test statistics. A formal proof of the consistency of this testing procedure is given in the Appendix.

3.2 Estimation of structural break positions

If our test rejects \mathbb{H}_0 in favor of at least one variance change, we typically want to find out where the structural breaks occur. Therefore, our goal is to locate the presumable change point positions. A rough approximation for the position of the dominant change in variance is given by t_{j^*} , where $j^* = \operatorname{argmax} \left| \log \left(\widehat{\sigma}_j^2 \right) - \log \left(\widehat{\sigma}_{j+1}^2 \right) \right|$ and the maximization is performed over $j = 1, \dots, N - 1$. The resolution of this estimator is obviously limited by the block lengths $\tau_j, j = 1, \dots, N$. This is particularly problematic, if the block boundaries t_1, \dots, t_N are not determined by a priori knowledge. In order to alleviate this problem the presumable position of the dominant change in variance can be fine tuned as follows: since we expect a structural break near the rough estimate t_{j^*} , we consider the union of the two blocks around $t_{j^*}, B = \{t_{j^*-1}, t_{j^*-1} + 1, \dots, t_{j^*} - 1, t_{j^*}, t_{j^*} + 1, \dots, t_{j^*+1}\}$. For the moment, we focus solely on the observations with indices in B and exclude the ones from the remaining blocks from this part of the analysis. Our goal is to find the index $t^* \in B$ such that the empirical variance before $t^*, \widehat{\sigma}_1^2(t^*)$, differs the most from the empirical variance after $t^*, \widehat{\sigma}_2^2(t^*)$. Thereby, both $\widehat{\sigma}_1^2(t^*)$ and $\widehat{\sigma}_2^2(t^*)$ are computed analogously to (1) using observations from B only. The position of the presumable structural break is thus estimated by

$$t^* = \operatorname{argmax}_{t \in B} \left| \log \left(\widehat{\sigma}_1^2(t) \right) - \log \left(\widehat{\sigma}_2^2(t) \right) \right|.$$

To ensure meaningful estimations $\widehat{\sigma}_1^2(t)$ and $\widehat{\sigma}_2^2(t)$ we do not maximize over all $t \in B$ but leave out values of t near the boundaries of B .

Several structural break positions are located in a recursive manner in the spirit of [Vostrikova \(1981\)](#): after identifying the presumable position of the dominant change in variance as described above, the sample is split into two parts at that point. The test procedure is then conducted on each of the two subsets, as long as they are large enough to ensure reasonable estimations. In case of new rejections, the corresponding presumable change point positions are determined and the splitting continues. As soon as the test does not reject \mathbb{H}_0 on all current subsamples, the original sample appears to be split into homogeneous subsamples and the procedure stops. This approach indeed yields a level α test, since under \mathbb{H}_0 the permutation test conducted on the complete sample rejects in only α percent of the cases.

3.3 Alternative methods

The literature offers several approaches checking the stationarity of the variance for a series of random variables. One of them is the CUSUM procedure. It is a standard tool in the detection of structural breaks and hence a lot of work is available in this

context. We choose the method proposed by [Wied et al. \(2012\)](#) as a representative for this class of tests. It is based on the CUSUM statistic

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

where $\widehat{\sigma}_{1:l}^2$ denotes the empirical variance computed from the first l observations, $l = 1, \dots, n$. The normalizing scalar \widehat{D} is necessary to attain the asymptotic distribution. The CUSUM approach compares the discrepancies between the estimated variance on the whole sample to all estimated variances on proper subsamples. It then determines the maximal deviation signaling a possible structural break. The test is designed to detect at most one change in volatility. Critical values are derived from asymptotics. [Peña \(2005\)](#) also compares variances estimated on subsamples to a measure of volatility estimated on the complete sample. He proposes a test statistic built in a blockwise manner, namely

$$T_{\text{Log}} = n \log \left(\sum_{t=1}^n X_t^2 \right) - \sum_{j=1}^N \tau_j \log \left(\widehat{\sigma}_j^2 \right).$$

As for our test statistic, the distribution of T_{Log} under \mathbb{H}_0 heavily depends on the data. We therefore again apply the permutation principle to obtain a distribution-free test. Another approach checking the hypothesis of global constant volatility is given by [Ross \(2013\)](#). It is motivated by the classical distribution-free procedure proposed by [Mood \(1954\)](#). Instead of using X_1, \dots, X_n , their ranks in the complete sample, denoted by r_1, \dots, r_n , are determined. Hereafter, the sample is divided into two subsamples for each possible split position $t = 1, \dots, n$. For each of these splittings, the standardized statistic of the Mood test is calculated. Its expected value $\mu_t = t(n^2 - 1)/12$ and the standard deviation $\sigma_t = \sqrt{t(n-t)(n+1)(n^2-4)/180}$ hold under the null hypothesis. Taking the maximum over the possible split positions $t = 1, \dots, n$ results in

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

Since only the ranks of the observations contribute to the test statistic, the procedure is distribution-free. Appropriate critical values depend solely on the sample size n and can be derived by simulations. For several critical values and more details, we refer to [Ross \(2013\)](#).

In addition, we consider the monitoring procedure based on characteristic functions proposed by [Steland and Rafajłowicz \(2014\)](#). The authors develop a test statistic comparable to ours. It has the advantage that changes in the location process do not affect the monitoring of the volatility and vice versa. According to the authors,

$$S_j = \int \left[(\widehat{U}_j(u))^2 + (\widehat{V}_j(u))^2 \right] w(u) du$$

is an estimator in the context of characteristic functions which reflects the volatility in the j -th data block, $j = 1, \dots, N$. Thereby, w is a weight function and may be chosen as suggested in Sect. 2.2. The statistics \widehat{U}_j and \widehat{V}_j denote the standard estimators of the real and imaginary part of the characteristic function for the random variables in the j -th block, $j = 1, \dots, N$:

$$\widehat{U}_j(u) = \frac{1}{\tau_j} \sum_{t \in B_j} \cos(u \cdot X_t) \quad \text{and} \quad \widehat{V}_j(u) = \frac{1}{\tau_j} \sum_{t \in B_j} \sin(u \cdot X_t).$$

We transfer the monitoring procedure of [Steland and Rafajłowicz \(2014\)](#) to the retrospective case in the following way: since \mathbb{H}_0 should be rejected if the volatilities in two blocks are substantially different, we propose to use

$$T_{\text{cf}} = \max_{1 \leq j < k \leq N} |S_j - S_k|$$

to test the null hypothesis via the permutation principle. Note that for any $j = 1, \dots, N$ one can express S_j as

$$S_j = \frac{1}{\tau_j^2} \sum_{t, t' \in B_j} I_w(X_t - X_{t'}),$$

so that by (3) our statistic T_{Four} can be interpreted as a weighted version of S_j computed on the pseudo-observations $\log(\widehat{\sigma}_1^2), \dots, \log(\widehat{\sigma}_N^2)$. All four methods presented in this section reject the hypothesis of a global constant volatility for large values of the corresponding test statistic.

4 Evaluation of the methods

In this section, we compare the performance of our permutation tests based on T_{Four} to its competitors listed in Sect. 3.3. For this purpose, we compute the empirical power of each procedure for different data cases. We thereby address the choice of parameters and weighting functions. Hereafter, the best methods are applied to gold mining data.

4.1 Choice of settings

As a first step of the analysis, we assess the influence of the weight function w , its parameter a and the number of the blocks N on the tests based on the Fourier-type statistics T_{Four} and T_{cf} , respectively. Since both methods are constructed using the permutation principle, they attain a predefined significance level α under the null hypothesis of global constant volatility. We thus consider their empirical power under various alternatives as an adequate measure of performance.

The tests are evaluated on datasets consisting of 200 observations each. The first half of every sample is generated from the standard Gaussian distribution. The second

Table 1 Rejection rates given in rounded percent for data with one increase in the variance in the middle of the sample for the permutation tests based on T_{Four} and T_{cf} , respectively

	w_U			w_L			w_G		
	0.5	1	1.5	0.5	1	1.5	0.5	1	1.5
T_{Four}									
$N = 5$	80	80	81	81	80	80	80	80	81
$N = 10$	72	72	72	71	72	71	72	72	71
T_{cf}									
$N = 5$	68	69	65	60	65	68	66	68	69
$N = 10$	47	50	42	34	42	46	43	48	49

Table 2 Rejection rates given in rounded percent for data with one increase in the variance in the middle of the sample for the permutation tests based on T_{Four} and T_{cf} for different numbers of initial equidistant blocks $N = 2, \dots, 10$

	2	3	4	5	6	7	8	9	10
T_{Four}	96	82	93	79	85	71	79	66	74
T_{cf}	98	77	86	67	68	56	62	48	50

100 observations are sampled from the Gaussian distribution with increased standard deviation 1.5 and mean 0. In this manner, we proceed generating 1000 datasets and apply both tests to them. Following the discussion in Sect. 2.2, appropriate values for the weight parameter $a > 0$ are chosen from the literature on empirical characteristic functions, which are comparable to our quantity $\hat{\varphi}_N$; see for instance Jiménez-Gamero et al. (2009), Potgieter and Genton (2013) and Pardo-Fernández et al. (2015). In accordance with this prior experience, we consider the values $a = 0.5, 1, 1.5$ for the uniform (w_U), Laplacian (w_L) and Gaussian (w_G) weight functions. The number of equidistant blocks is set to $N = 5$ and $N = 10$ and 2000 permutations are conducted for both test procedures. The resulting rejection rates are shown in Table 1.

According to Table 1, the choice of the weight function w and its parameter a do not have a large influence on the performance of our test based on T_{Four} . The second method seems to be more strongly affected by these settings. Unsurprisingly, both procedures heavily depend on the splitting of the data into blocks. For the considered data scenario, the tests lead to lower rejection rates for $N = 10$ than for $N = 5$ although for $N = 5$ the true structural break lies in the middle of one block. We also observe that our newly proposed test considerably outperforms the procedure using T_{cf} in all cases. In the following, we restrict ourselves to the Gaussian weighting and $a = 1.5$ for both tests due to the results in the previous analysis. Repeating the simulation for $N = 2, \dots, 10$ yields the rejection rates in Table 2.

As expected, the choice of $N = 2$ blocks leads to the highest rejection rates, since this data splitting perfectly fits the true volatility clusters. For the same reason, even values of N attain better results than the corresponding odd neighbors $N - 1$ and $N + 1$. In general, the rejection rates decrease with N , since many small blocks lead to worse estimates of the volatility in the individual blocks. Also, for large N some blocks estimate the same volatility. The difference of the corresponding logarithmized

Table 3 Rejection rates given in rounded percent for data with one increase in the variance after 20% of the observations for the permutation tests based on T_{Four} and T_{cf} for different numbers of initial equidistant blocks $N = 2, \dots, 10$

	2	3	4	5	6	7	8	9	10
T_{Four}	25	34	48	69	53	44	42	49	50
T_{cf}	26	35	47	70	48	37	36	32	36

blockwise estimates figuring in T_{Four} is thus not equal to 0 solely due to sampling error. Therefore, more terms than necessary contribute to the mean T_{Four} , so that the impact of the terms reflecting actual changes of the volatility are downweighted. Lastly, we observe that, up to a certain degree, the correct position of the blocks is more relevant than their actual number. For example, even though $N = 2$ is optimal in this case, $N = 6$ is preferable to $N = 3$ and $N = 5$ for the test using T_{Four} .

To study the latter point in more detail, we repeat the simulation presented above with the structural change now taking place after 20% of the observations rather than in the middle of the sample. Table 3 provides the corresponding rejection rates.

According to the results, choosing a few oversized blocks can be more problematic than using many small blocks. In the scenario investigated, $N = 5$ blocks free of structural breaks attain the best results, although actually just two volatility regions exist. For $N = 2$, the correct number of blocks, the data is split at the wrong position leading to worse performance of both methods. As before, one can observe a general decrease in power after the optimal choice $N = 5$ and relatively good results for its multiple $N = 10$. These findings consistently apply for both tests. Thereby, T_{Four} leads to increasingly higher rejection rates compared to T_{cf} as N , the number of blocks increases. We will revisit the problem of block-choice in the discussion.

4.2 Method comparison

We now apply all tests introduced in Sect. 3 in six different data scenarios. To present the settings in a clear and compact way, let the notation $|n = n_1, \sigma = s_1 | n = n_2, \sigma = s_2|$ describe n_1 observations with standard deviation s_1 followed by n_2 observations with standard deviation s_2 . Generalizing this representation to the case of several data blocks allows us to represent the scenarios under study in the following way:

- (1) $|n = 200, \sigma = 1|$
- (2) $|n = 100, \sigma = 1 | n = 100, \sigma = 1.5|$
- (3) $|n = 100, \sigma = 1 | n = 100, \sigma = 1.5 | n = 100, \sigma = 1|$
- (4) $|n = 100, \sigma = 1 | n = 100, \sigma = 1.4 | n = 100, \sigma = 1 | n = 100, \sigma = 1.4 | n = 100, \sigma = 1|$
- (5) $|n = 100, \sigma = 1 | n = 50, \sigma = 1.6 | n = 100, \sigma = 1 | n = 150, \sigma = 1.2 | n = 100, \sigma = 1|$
- (6) Four random structural breaks in 500 observations

The first five of them correspond to the null hypothesis and the case of one, two, four and four nonequidistant changes in volatility, respectively. In scenario (6), we randomize

both the positions of the structural breaks as well as the variances within the blocks for every sample. The four change point positions are sampled uniformly from the values 40, 41, ..., 460 such that the resulting blocks contain at least 40 observations. The corresponding variances are randomly chosen from the values 0.6, 0.8, 1.0, 1.2, 1.4. Thereby, we make sure that subsequent blocks do not share the same variance. For each of these six scenarios, three distributions are considered. These are the standard Gaussian distribution (G), the t-distribution with 5 degrees of freedom (t5) and the exponential distribution with parameter $\lambda = 1$ (exp) shifted to have zero mean. In order to obtain the desired standard deviations, scaling is applied. The six data scenarios and the three distributions yield 18 data cases in total. For each of them, 10,000 replications are generated. All five tests introduced in Sect. 3 are applied to this data at a significance level of 5%. The permutation tests are executed with 2000 permutations. The procedures involving blocks are conducted using $N = 10$ equidistant data blocks on the original data and all subsamples. To reduce the computational burden, the tests based on T_{Four} and T_{cf} are carried out only for $a = 1.5$ and $w = w_G$ due to the analysis in the previous subsection. The resulting rejection rates are listed in Table 4.

Under the null hypothesis, all methods roughly keep the significance level of 5%. For non-Gaussian data, however, the asymptotic of the CUSUM test does not provide a good approximation so that its rejection rates are somewhat low under \mathbb{H}_0 . As expected, the CUSUM procedure leads to the best results for Gaussian data with one volatility change, but loses a considerable amount of power in presence of multiple structural breaks due to masking effects. We observe a similar loss of efficiency for the Mood-type test. Much like the CUSUM approach, the distribution-free procedure is based on a two-sample test. It therefore also implicitly anticipates one structural break at a time. Nevertheless, the procedure proposed by Ross (2013) clearly outperforms its competitors in the case of exponential data in all but the fifth data scenario. In the latter setting, its rejection rate is still fine, but the procedure does not dominate its competitors any longer. The tests based on T_{Four} and T_{Log} overall lead to competitive results. In particular, their performance often improves when multiple structural breaks occur, in contrast to the CUSUM and the Mood-type test. As a consequence, they clearly outperform the CUSUM test for all distributions under study and the test proposed by Ross for the symmetric distributions in case of more than one volatility change. Among the two, T_{Log} leads to slightly higher rejection rates. The test using T_{cf} performs similarly to T_{Four} and T_{Log} , but is inferior to both in all but two considered scenarios. This is in accordance with the results discussed in Sect. 4.1. In general, the rejection rates of all tests are lower for t-distributed data than for the corresponding Gaussian case. This indicates that heavier tails make the detection of the breaks more difficult. A good test for structural breaks should have high rejection rates under various alternatives. However, it also must adequately determine the causes of the heterogeneity causing the rejection. Otherwise, it connects the correct rejection with an irrelevant event leading to false conclusions. We therefore take a closer look at the estimated number and location of the structural breaks of the methods. Thereby, we focus on the results for the blockwise procedures based on T_{Four} and T_{Log} , which showed the best overall performance in our opinion. Both tests are conducted in the recursive manner as explained in Sect. 3.2. Their mean numbers of estimated change points are listed in Table 5 for the six data scenarios investigated above for t-distributed samples:

Table 4 Rejection rates given in percent for all tests introduced in Sect. 3 for the six different volatility scenarios on page 11 and for Gaussian (G), t-distributed (t5) and exponential data (exp)

	T_{Four}	T_{CUS}	T_{Mood}	T_{cf}	T_{Log}
\mathbb{H}_0					
G	5.2	4.9	5.3	5.0	5.4
t5	5.0	3.5	5.1	5.3	5.1
exp	5.0	3.0	4.8	5.2	5.1
1 break					
G	71.0	89.3	78.4	44.1	74.6
t5	47.2	51.1	65.8	35.1	47.4
exp	49.1	31.8	99.0	26.3	41.4
2 breaks					
G	78.5	2.6	38.1	55.2	82.7
t5	50.6	1.9	29.0	42.9	52.3
exp	44.3	1.4	83.6	31.0	41.1
4 breaks					
G	77.7	0.7	14.6	55.6	80.1
t5	70.1	0.7	20.4	64.8	69.7
exp	56.1	0.9	70.4	43.7	53.7
4 noneq. breaks					
G	88.9	0.3	14.6	80.8	92.0
t5	56.2	0.7	12.6	65.5	60.8
exp	44.7	0.8	42.6	43.3	46.4
4 random breaks					
G	93.7	45.4	85.9	91.5	93.0
t5	82.9	29.2	80.4	82.2	81.2
exp	75.6	21.2	95.6	68.5	72.9

Table 5 Mean number of detected structural breaks for the permutation tests using the blockwise statistics T_{Four} and T_{Log} on data generated by the t-distribution with five degrees of freedom for the six data cases introduced on page 11. In brackets, the mean number of detected structural breaks among the samples with rejection is given

	\mathbb{H}_0	1 break	2 breaks	4 breaks	4 noneq. breaks	4 random breaks
T_{Four}	0.07 (1.4)	0.70 (1.48)	1.05 (2.08)	1.66 (2.37)	2.16 (3.84)	1.73 (2.08)
T_{Log}	0.08 (1.56)	0.76 (1.60)	1.15 (2.20)	1.74 (2.48)	2.28 (3.75)	2.41 (2.96)

To interpret the results correctly keep in mind that each data case was chosen such that the rejection rates of all tests are below 100% so that the tests are comparable. Therefore, unsurprisingly, both tests do not detect all structural breaks, since the

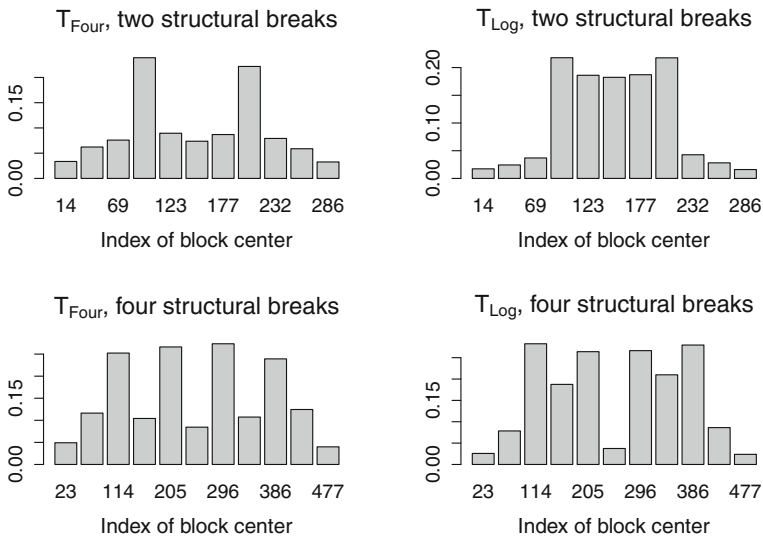


Fig. 1 Change point positions estimated by the permutation procedure for two (*top*) and four (*bottom*) present structural changes for the test based on T_{Four} (*left*) and T_{Log} (*right*)

volatility changes are simply not that clear by construction of the setting. Quite intuitively it seems harder to detect all structural changes the more of them occur. The test based on T_{Log} rejects more often and thus estimates more structural breaks on average. The differences are quite small though. We also provide the mean numbers of estimated structural breaks only considering the samples where the tests reject. As indicated by the values in brackets, both methods determine a reasonable number of presumable structural breaks if they reject at all. Their results are again quite similar. The estimated change point positions for the case of two and four equidistant structural breaks for t-distributed data are presented in Fig. 1 for both methods. Note that both tests estimate presumable change points in the same way in case of rejection, see Sect. 3.2. The different results are thus mainly a consequence of additional rejections of the method based on T_{Log} on subsamples. Figure 1 reveals that the test based on T_{Four} estimates the true change point positions much more precisely than the one using T_{Log} . Consequently, the estimated change points due to additional rejections of the latter method do not coincide with the true structural breaks for the most part. This suggests that tackling the detection of structural breaks via Fourier-type transformations is advantageous in comparison with a simpler blockwise approach. The price paid by a somewhat smaller rejection rate is outweighed by a more exact estimation of the change point position.

Next, we investigate how the structural break position influences the methods' performance. Therefore, we repeat the simulation for Gaussian data and one volatility change increasing from $\sigma = 1$ to $\sigma = 1.5$. This time we let the structural break occur after 10, 20, ..., 80%, or 90% of the 200 observations. We apply all tests presented, but leave out the permutation test based on T_{cf} due to its poor results in the previous sim-

Table 6 Rejection rates given in rounded percent for different volatility tests in case of one volatility change occurring after different proportions of the data

	10%	20%	30%	40%	50%	60%	70%	80%	90%
T_{Four}	28.6	50.3	65.3	67.4	75.1	70.1	58.7	45.3	20.3
T_{CUS}	9.3	25.3	57.1	80.1	88.5	88.5	79.6	55.7	11.8
T_{Mood}	16.2	47.7	66.9	71.2	78.5	75.0	68.9	59.7	38.2
T_{Log}	25.3	47.3	63.1	68.2	76.8	75.4	65.8	57.6	34.2

ulations. The corresponding rejection rates based on 1000 replications are presented in Table 6.

The best results for all methods are attained when the structural break is in the middle of a sample. The rejection rates decrease the more the volatility clusters differ in size. The best procedure for a break in the middle of the sample, the CUSUM test, is most severely affected by this fact. Its rejection rates drop the most when the data is almost exclusively generated from one model. In general, better results are achieved for all tests if the majority of the data has a standard deviation of 1 rather than 1.5. The overall variance of the data in this case is smaller leading to a more precise decision process.

The simulation presented next allows to assess the importance of our zero mean assumption. As before, we induce one or two structural breaks on 1000 Gaussian samples of size 200, respectively, and apply our test based on T_{Four} to the data. Hereafter, the data sets are centered by their arithmetic means and again passed to the permutation test. In case of one volatility change, we get rejection rates of 70.3% for the uncentered data and 70.2% for the centered equivalents. For two structural breaks, the results are 76.0 and 75.2%. We thus conclude that demeaning has a negligible effect on the performance of our test even for comparatively small numbers of observations. Instead of zero mean data, it is thus possible to assume an unknown constant mean. Finally, we compare the tests on artificial data with linearly increasing variance. For this purpose, we simulate 500 datasets containing 200 observations each. The first 100 observations for each of them are generated from the standard Gaussian distribution. The remaining ones are drawn from Gaussian distributions with mean zero and a variance increasing by 0.025, that is, $\sigma^2(101) = 1.025$, $\sigma^2(102) = 1.05$, ..., $\sigma^2(200) = 3.5$. The corresponding rejection rates for the test under study are 83% for T_{Four} , 94% for T_{CUS} , 89% for T_{Mood} and 88% for T_{Log} . We repeat the simulation now increasing the variance by 0.1 each time. This time all tests reject the null hypotheses of constant variance in all cases. Kernel density estimates of the estimated change point positions for each test are given in Fig. 2. Hereby, we use the Gaussian kernel and the default bandwidth selection for the *kernel* routine in R.

The test based on T_{Four} does not detect the structural break position well in case of a comparatively small linear increase in the variance. In contrast, the other tests correctly estimate potential change points almost exclusively after the 100th observation. In case of increased variance growth, our procedure performs much better than before relatively to its competitors. It is more sensitive to the first change of the variance and does not get distracted by the linear increase after the 100th observation.

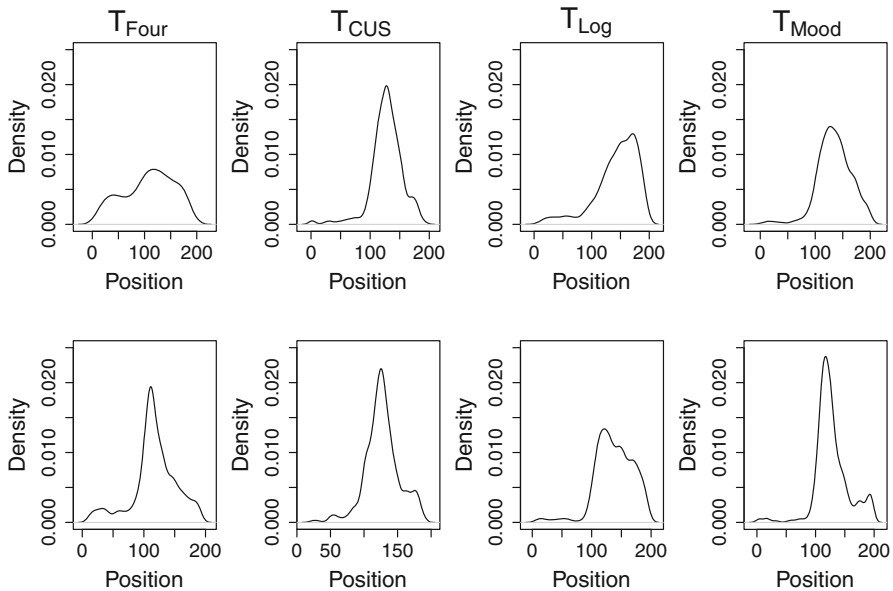


Fig. 2 Kernel density estimates of the estimated positions of potential structural changes for different tests in case of a linear increase in the variance by 0.025 (*upper row*) and 0.1 (*lower row*) after 100 observations

4.3 Application

For further illustration of the methods, we consider a data set related to gold mining in South Africa introduced in [Rowland and Sichel \(1960\)](#). The extraction and processing of gold are quite expensive. Therefore, ore samples are collected in mines and checked for their gold content by chemical examinations to discover promising cultivating regions. However, taking representative samples is quite complicated due to the highly irregular gold concentration. This leads to unreliable results in particular for new samplers. Thus, experienced supervisors resample part of the work. The data considered in the following and presented in [Fig. 3](#) are taken from [Jandhyala et al. \(2002\)](#). It consists of 157 logarithmised ratios of gold contents in samples collected by a junior sampler to samples mined by a supervisor at the same locations. The data are arranged in chronological order. Values with a large modulus therefore indicate a high disagreement between the two corresponding samples. Such a result is interpreted as nonrepresentative sampling by the junior sampler. We follow the arguments of [Jandhyala et al. \(2002\)](#) stating that unreliable sampling does not induce a bias in the data, but rather leads to more unstable results. In accordance with these authors, we thus assume zero mean observations. To evaluate the work of the junior sampler, we check the constancy of the variance of the data applying the tests performing best on artificial data.

Using $N = 5$ almost equidistant blocks, the Gaussian weighting and $a = 1$ the test based on T_{Four} detects a variance change at observation 81. This seems plausible in view of the data and is in accordance with the results in [Jandhyala et al. \(2002\)](#). Note

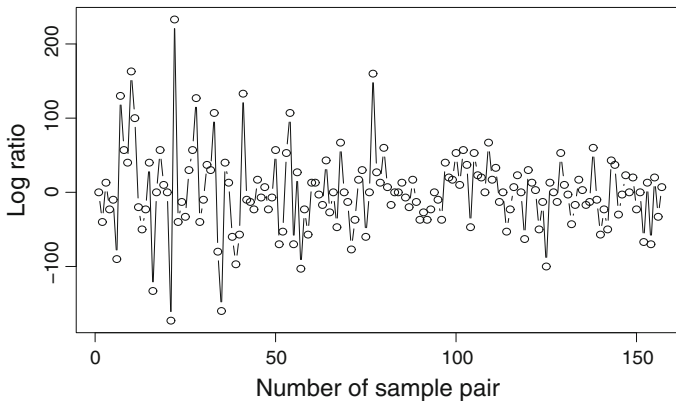


Fig. 3 Logarithmised ratios of gold contents in samples collected by a junior sampler to samples mined by a supervisor at the same 157 locations (Jandhyala et al. 2002)

that last-named authors propose a method relying on Gaussianity, while our procedure is valid regardless of this assumption. The test based on T_{Log} detects only one change point at observation 42. It thereby neglects both the region of comparatively high variance around observation 55 as well as the high peak at observation 77. Therefore, the procedure based on T_{Four} appears to be more favorable in this application.

5 Conclusion and outlook

In this work, we consider the problem of testing whether a sequence of random variables has constant volatility over time. In accordance with Spokoiny (2009), Davies et al. (2012) and other authors we assume that the volatility is approximately piecewise constant. We thus propose a new test based on blockwise volatility estimates and a Fourier-type transformation. According to our extensive simulations, it performs competitively in comparison with alternative procedures from the literature in case of data from symmetric distributions. In particular, our test is recommendable if several structural changes may occur. In case of rejection, it determines the structural break positions adequately. An application to gold mining data also leads to meaningful results.

Our concept can easily be adapted for testing the constancy of other distributional properties such as skewness, kurtosis or tail behavior. This can be achieved in a straightforward way: one simply substitutes the estimator of the local variance presented in (1) by another measure reflecting the quantity of choice. Robust estimators of scale can also be incorporated in this way, if outliers are an issue. Simulation results for the case of kurtosis not reported here show that the corresponding method performs similar to the volatility case.

In forthcoming work, we would like to further explore the asymptotic behavior of our test statistic. First results in this regard indicate that under certain additional assumptions, the asymptotic distribution of T_{Four} for a particular weight function is Gaussian under the null hypothesis. These investigations however are the subject of ongoing

work. Also, we are interested in extending our methods to dependent observations considering a more general framework; see for instance Brooks et al. (2005); Feunou and Tédongap (2012); Harvey and Siddique (1999). In this connection, appropriate resampling techniques might be of great help. Finally, our procedure might benefit from a more refined choice of data blocks. We have considered two initial approaches for this purpose. The first one pre-estimates potential structural breaks by applying a fast heuristic to the data. For any possible split position, the heuristic estimates the variance both shortly before as well as shortly after the split position. It then returns the positions with the largest discrepancies between the two corresponding estimated local variances. To avoid representing the same structural break several times, we thereby exclude positions close to ones with even greater local variance changes. The second approach determines the best of several given data partitions by maximizing a properly standardized version of our test statistic. The choice of $N = 10$ fixed equidistant blocks performed almost equal or better than both alternative strategies for the simulation settings presented in this manuscript and in particular for the case of four randomized structural breaks. Thus, for the moment, we recommend to use somewhat more than twice as many equidistant blocks as the expected number of structural breaks. In this way, the majority of the blocks are free of structural breaks so that the test's power should be decent. Nevertheless, there is certainly room for further improvement.

Acknowledgements The work was supported in part by the Collaborative Research Center 823, Project C3 (Analysis of Structural Change in Dynamic Processes), of the German Research Foundation and by Grant No.11699 of the Special Account for Research Grants (ELKE) of the National and Kapodistrian University of Athens. We also thank the anonymous referees for their valuable remarks which helped us to improve this work.

Appendix

In this subsection, we prove the consistency of our permutation test based on V_n against a change in volatility alternative, as specified in \mathbb{H}_1 .

Let us first recall our definitions

$$\varphi_n(u) = \sum_{j=1}^N \frac{\tau_j}{n} e^{iu \log(\hat{\sigma}_j^2)} \quad \text{and} \quad V_n = \int \left(1 - |\varphi_n(u)|^2\right) w(u) du,$$

whereby $\hat{\sigma}_j^2 = \frac{1}{\tau_j} \sum_{k \in B_j} X_k^2$ for $j = 1, \dots, N$. The X_1, \dots, X_n are assumed to be independent zero mean random variables. The corresponding variances remain constant within each block B_j , i.e., $\text{Var}(X_t) = \sigma_j^2, \forall t \in B_j, j = 1, \dots, N$. We consider weight functions w which are integrable and positive, except possibly over a set of measure zero.

We prove consistency of our test in scenarios with an arbitrary but fixed (in n) number of change points, where the number of observations inbetween the change points grows linearly with the same rate (infill asymptotics). Assume that N equidistant blocks are chosen, with $N = N_n$ growing slower than linearly in n , such that the block

length grows in n , too. Then the fraction of the blocks corresponding to the same true variance asymptotically equals the fraction κ_j , $j = 1, \dots, \nu$ of observations with this variance. Thereby, ν denotes the number of different true variances. We then have

$$\varphi_n(u) \rightarrow \varphi(u) = \sum_{j=1}^{\nu} \kappa_j e^{iu \log \sigma_j^2}.$$

φ is the characteristic function of a random variable, say Y , which satisfies

$$\mathbb{P}(Y = \log \sigma_j^2) = \kappa_j \quad \text{for } j = 1, \dots, \nu.$$

By taking into account that $1 - |\varphi_n(u)|^2 \leq 1$ and that w is integrable, the Lebesgue theorem of dominated convergence entails

$$V_n = \int_{-\infty}^{\infty} (1 - |\varphi_n(u)|^2) w(u) du \rightarrow \int_{-\infty}^{\infty} (1 - |\varphi(u)|^2) w(u) du = V,$$

almost surely. Clearly $|\varphi(u)|^2 \leq 1$ holds. Along with $w > 0$ this implies that V is positive unless $|\varphi(u)|^2 = 1$, identically in u . However, the latter would mean that φ is the characteristic function of a degenerate random variable, i.e., $\mathbb{P}(Y = c) = 1$ for some $c \in \mathbb{R}$; see Loève (1977, §14.1). Since this does not hold under \mathbb{H}_1 , our test statistic V_n converges toward $V > 0$ for any alternative.

Let us now consider a random sample $\tilde{X}_1, \dots, \tilde{X}_n$ generated from the original sample via permutation. Define \tilde{V}_n in analogy V_n , but computed on the permuted rather than the original sample. For any $n \in \mathbb{N}$, the $\tilde{X}_1, \dots, \tilde{X}_n$ are strictly stationary. Thus we get $\tilde{V}_n \rightarrow 0$ using similar arguments as above. Hence, the permutation test rejecting \mathbb{H}_0 for large values of V_n is strongly consistent against a general change-point alternative such as \mathbb{H}_1 . An analogous result holds for the permutation test based on T_{Four} , since T_{Four} and V_n are identical up to constants.

References

- Brooks, C., Burke, S.P., Heravi, S., Persaud, G.: Autoregressive conditional kurtosis. *J. Financ. Econom.* **3**(3), 399–421 (2005)
- Davies, L., Hönenrieder, C., Krämer, W.: Recursive computation of piecewise constant volatilities. *Comput. Stat. Data Anal.* **56**(11), 3623–3631 (2012)
- Epps, T.W.: Characteristic functions and their empirical counterparts: geometrical interpretations and applications to statistical inference. *Am. Stat.* **47**(1), 33–38 (1993)
- Epps, T.W.: Limiting behavior of the ICF test for normality under Gram–Charlier alternatives. *Stat. Prob. Lett.* **42**(2), 175–184 (1999)
- Feunou, B., Tédongap, R.: A stochastic volatility model with conditional skewness. *J. Bus. Econ. Stat.* **30**(4), 576–591 (2012)
- Fisher, R.A.: *The design of experiments*. Oliver & Boyd, Oxford (1935)
- Fried, R.: On the online estimation of local constant volatilities. *Comput. Stat. Data Anal.* **56**(11), 3080–3090 (2012)
- Good, P.: *Multiple tests. Permutation, parametric and bootstrap tests of hypotheses*. Springer, New York (2005)

- Guégan, D.: Non-stationary samples and meta-distribution. In: Basu, A., Samanta, T., Sen Gupta, A. (eds.) *Statistical Paradigms Recent Advances and Reconciliations*, vol. 14. World Scientific, New Jersey (2015)
- Harvey, C.R., Siddique, A.: Autoregressive conditional skewness. *J. Financ. Quant. Anal.* **34**(4), 465–487 (1999)
- Hlávka, Z., Hušková, M., Kirch, C., Meintanis, S.G.: Monitoring changes in the error distribution of autoregressive models based on Fourier methods. *Test* **21**(4), 605–634 (2012)
- Hlávka, Z., Hušková, M., Kirch, C., Meintanis, S.G.: Fourier-type tests involving martingale difference processes. *Econ. Rev.* (2015). doi:[10.1080/07474938.2014.977074](https://doi.org/10.1080/07474938.2014.977074)
- Hsu, D.A.: Tests for variance shift at an unknown time point. *J. R. Stat. Soc.* **26**(3), 279–284 (1977)
- Hušková, M., Meintanis, S.G.: Change point analysis based on empirical characteristic functions. *Metrika* **63**(2), 145–168 (2006a)
- Hušková, M., Meintanis, S.G.: Change-point analysis based on empirical characteristic functions of ranks. *Seq. Anal.* **25**(4), 421–436 (2006b)
- Jandhyala, V.K., Fotopoulos, S.B., Hawkins, D.M.: Detection and estimation of abrupt changes in the variability of a process. *Comput. Stat. Data Anal.* **40**(1), 1–19 (2002)
- Jiménez-Gamero, M.D., Alba-Fernández, V., Muñoz-García, J., Chalco-Cano, Y.: Goodness-of-fit tests based on empirical characteristic functions. *Comput. Stat. Data Anal.* **53**(12), 3957–3971 (2009)
- Loève, M.: *Probability Theory I*. Springer, New York (1977)
- Matteson, D.S., James, N.A.: A nonparametric approach for multiple change point analysis of multivariate data. *J. Am. Stat. Assoc.* **109**(505), 334–345 (2014)
- Meintanis, S.G.: Permutation tests for homogeneity based on the empirical characteristic function. *Non-parametr. Stat.* **17**(5), 583–592 (2005)
- Meintanis, S.G., Swanepoel, J., Allison, J.: The probability weighted characteristic function and goodness-of-fit testing. *J. Stat. Plan. Inference* **146**, 122–132 (2014)
- Mercurio, D., Spokoiny, V.: Statistical inference for time-inhomogeneous volatility models. *Ann. Stat.* **32**(2), 577–602 (2004)
- Mood, A.M.: On the asymptotic efficiency of certain nonparametric two-sample tests. *Ann. Math. Stat.* **25**(3), 514–522 (1954)
- Pardo-Fernández, J.C., Jiménez-Gamero, M.D., El Gouch, A.: A nonparametric ANOVA-type test for regression curves based on characteristic functions. *Scand. J. Stat.* **42**, 197–213 (2015)
- Peña, D.: *Análisis de Series Temporales*, vol. 319. Alianza Editorial, Madrid (2005)
- Potgieter, C.J., Genton, M.G.: Characteristic function-based semiparametric inference for skew-symmetric models. *Scand. J. Stat.* **40**(3), 471–490 (2013)
- Ross, G.J.: Modelling financial volatility in the presence of abrupt changes. *Phys. A Stat. Mech. Appl.* **392**(2), 350–360 (2013)
- Rowland, R.S.J., Sichel, H.: Statistical quality control of routine underground sampling. *J. S. Afr. Inst. Min. Metal* **60**, 251–284 (1960)
- Spokoiny, V.: Multiscale local change point detection with applications to value-at-risk. *Ann. Stat.* **37**(3), 1405–1436 (2009)
- Steland, A., Rafajłowicz, E.: Decoupling change-point detection based on characteristic functions: methodology, asymptotics, subsampling and application. *J. Stat. Plan. Inference* **145**, 49–73 (2014)
- Tenreiro, C.: On the choice of the smoothing parameter for the BHEP goodness-of-fit test. *Comput. Stat. Data Anal.* **53**, 1038–1053 (2009)
- Vostrikova, L.J.: Detecting disorder in multidimensional random process. *Sov. Math. Dokl.* **24**, 55–59 (1981)
- Wied, D., Arnold, M., Bissantz, N., Ziggel, D.: A new fluctuation test for constant variances with applications to finance. *Metrika* **75**(8), 1111–1127 (2012)