The Annals of Statistics
2017, Vol. 45, No. 4, 1542–1578
DOI: 10.1214/16-AOS1499

NONPARAMETRIC CHANGE-POINT ANALYSIS OF VOLATILITY

BY MARKUS BIBINGER[1], MORITZ JIRAK[1] AND MATHIAS VETTER

Philipps-Universität Marburg, Technische Universität Braunschweig and Christian-Albrechts-Universität zu Kiel

In this work, we develop change-point methods for statistics of high-frequency data. The main interest is in the volatility of an Itô semimartingale, the latter being discretely observed over a fixed time horizon. We construct a minimax-optimal test to discriminate continuous paths from paths with volatility jumps, and it is shown that the test can be embedded into a more general theory to infer the smoothness of volatilities. In a high-frequency setting, we prove weak convergence of the test statistic under the hypothesis to an extreme value distribution. Moreover, we develop methods to infer changes in the Hurst parameters of fractional volatility processes. A simulation study is conducted to demonstrate the performance of our methods in finite-sample applications.

1. Introduction. Change-point theory traditionally focused on detecting one or several structural breaks in the trend of time series. Statistical methods to infer change-points have a long and rich history, dating back to the pioneering work of Page (1955). Prominent approaches—here, we mention Hinkley (1971), Pettitt (1980), Andrews (1993), Bai and Perron (1998) among many others—provide statistical tests for the hypothesis of no change-point against the alternative that changes occur. Moreover, they allow for localization of change-points (estimation) and confidence intervals. Change-point methods usually rely on the order statistics and exploit limit theorems from extreme value theory; see Csörgő and Horváth (1997) for an overview. Less focus, however, has been laid on discriminating jumps from continuous motion in a nonparametric framework. Important exceptions are Müller (1992), Müller and Stadtmüller (1999), Spokoiny (1998) and Wu and Zhao (2007) in the context of nonparametric regression analysis. The latter serves as an important point of orientation for this work.

Statistics of high-frequency data is concerned with discretizations of continuous-time stochastic processes, most generally Itô semimartingales. The continu-

Received February 2016; revised June 2016.

[1]Financial support from the Deutsche Forschungsgemeinschaft via SFB 649 *Ökonomisches Risiko* and FOR 1735 *Structural Inference in Statistics: Adaptation and Efficiency* is gratefully acknowledged.

MSC2010 subject classifications. Primary 62M10; secondary 62G10.

Key words and phrases. High-frequency data, nonparametric change-point test, minimax-optimal test, stochastic volatility, volatility jumps.

ous part of an Itô semimartingale is of the form

$$X_t = X_0 + \int_0^t a_s\,ds + \int_0^t \sigma_s\,dW_s, \tag{1}$$

defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ with a standard $(\mathcal{F}_t)$-Brownian motion W and adapted drift and volatility processes a and σ. One key topic is statistical inference on the volatility under high-frequency asymptotics when the mesh $\Delta_n = n^{-1}$ of a discretization on the fixed time horizon [0, 1] tends to zero. There is a vast body of works related to this problem and its economic implications; see, for example, Andersen and Bollerslev (1998), Mykland and Zhang (2009) and Jacod and Rosenbaum (2013), among many others. As highlighted by Mykland (2012), statistics for discretized continuous-time martingales is closely related to Gaussian calculus. This observation turns out to be central for our analysis as well.

A natural question in high-frequency statistics is whether or not jumps are present in the Itô semimartingale modeling the log-price of a financial asset, and a statistical test was developed in Aït-Sahalia and Jacod (2009). More involved, but of key interest for economics and finance, too, is to infer the smoothness of the underlying stochastic volatility process and to check whether volatility jumps occur. In particular, inference on volatility jumps allows us to investigate the impact of certain news arrivals on financial risk. A first empirical study by Todorov and Tauchen (2011) indicates that volatility jumps indeed occur, but it has been based on direct observations of the VIX, the most prominent available volatility index. So far, there has been a lack of appropriate statistical methods when working with the price processes only. Further contributions consider joint price-volatility jumps. Jacod and Todorov (2010) have designed a test to decide from high-frequency observations whether common jumps of an Itô semimartingale and its volatility process have taken place at least once over some fixed time interval. These methods do not generalize directly to test for volatility jumps on the considered time interval, as one has to restrict to a finite set of large changes in the price first. One main profit from our change-point analysis of high-frequency data is a general test for volatility jumps. Moreover, results on estimation of the time of a volatility jump are provided.

As an example, we illustrate in Figure 1 the evolution of the log-prices of two blue-chip stocks, 3M and GE, over the NASDAQ intra-day trading period (6.5 h rescaled to the unit interval) on March 18, 2009. We consider one minute returns from executed trades[2] to ensure that the semimartingale model is adequate and to limit a manipulation by market microstructure frictions. It seems as if a common source of news drives both price dynamics at the end of that day, but standard tests do not identify jumps in both price processes with high significance. In particular,

[2]Reconstructed from the order book using LOBSTER, https://lobster.wiwi.hu-berlin.de/.

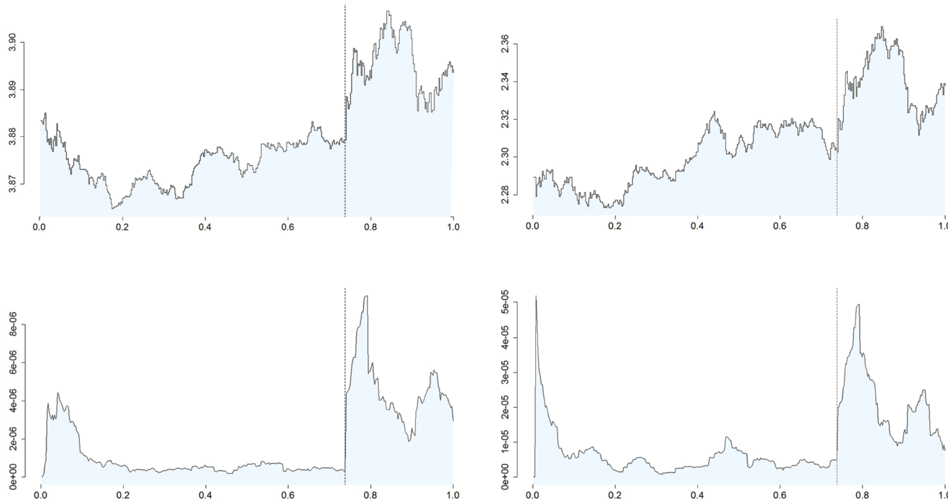

FIG. 1. *Log-price intra-day evolutions* (*top*) *and estimated spot squared volatilities* (*bottom*) *for* 3*M* (*left*) *and GE* (*right*) *on March* 18, 2009.

the test by Jacod and Todorov (2010) cannot be applied. The picture becomes much clearer, however, when focusing on the estimated spot squared volatilities in Figure 1, for which we average at each time point the previous 20 rescaled squared returns. The volatilities of both assets shoot up at exactly the same time, which suggests that at least volatility dynamics vary over time. This common volatility jump coincides with a press release at 02:15 p.m. EST subsequent to a meeting of the Federal Open Market Committee. The time is marked in Figure 1 by the dashed lines. In light of increasing economic slack, the FOMC announced "to employ all available tools to promote economic recovery and to preserve price stability,"[3] including a guarantee for an exceptionally low level of the federal funds rate for an extended period and a considerable increase of the size of the Federal Reserve's balance sheet. The statistical concepts developed in this work provide a novel device to infer volatility dynamics and jumps.

Change-point methods for volatility in a time-series environment, which is quite different to our high-frequency semimartingale setting, have been discussed by Spokoiny (2009). Quasi-likelihood estimation of a change-point in a diffusion parameter in a high-frequency setting has been considered by Iacus and Yoshida (2012), pointing out already one very useful bridge between change-point theory and high-frequency statistics. Our main focus is on testing for the presence of changes in a general setup. Beyond the analysis of possible jumps of the volatility, there is great interest in the smoothness regularity of volatilities, not least because

[3] www.federalreserve.gov/monetarypolicy/fomcminutes20090318.htm.

of its crucial role for setting up volatility models; see, for example, Gatheral, Jaisson and Rosenbaum (2014) for a recent work.

We focus on volatilities which are almost surely locally bounded and strictly positive adapted processes. For our testing problem, we consider classes of squared volatilities

$$\Sigma(\mathfrak{a}, L_n) = \Big\{(\sigma_t^2)_{t\in[0,1]} \,\big|\, \sup_{s,t\in[0,1],|s-t|<\delta} |\sigma_t^2 - \sigma_s^2| \le L_n\delta^{\mathfrak{a}}\Big\}, \tag{2}$$

for an appropriate sequence L_n converging to infinity; cf. Assumption 3.1 for a precise statement about the conditions under the null hypothesis. The regularity exponent $\mathfrak{a} > 0$ is the key parameter to describe the null hypothesis H_0. We may now more formally ask the following questions:

(i) Is there a jump in the volatility, that is, $\Delta\sigma_\theta^2 = (\sigma_\theta^2 - \lim_{s\uparrow\theta}\sigma_s^2) > 0$ for some $\theta \in (0, 1)$?

(ii) Does volatility get rougher in the sense that the regularity exponent drops to $\mathfrak{a}' < \mathfrak{a}$ on $(\theta, 1]$?

Question (i) poses a *local* problem, whereas question (ii) refers to a *local* or a *global* problem. Here, we call a change-point problem *global* if the regularity parameter changes from $\mathfrak{a}$ to $\mathfrak{a}' < \mathfrak{a}$ for some $\theta \in (0, 1)$.[4] This means that for $t \in [0, \theta)$ we have $\sigma_t^2 \in \Sigma(\mathfrak{a}, L_n)$, but for $t \in [\theta, 1]$ we only have $\sigma_t^2 \in \Sigma(\mathfrak{a}', L_n)$ *and* $\sigma_t^2 \notin \Sigma(\mathfrak{a}, L_n)$. In other words, starting from θ the volatility σ_t^2 fluctuates considerably more, and this stronger fluctuation persists for the remaining time period. As a key example, the global problem covers a change in the Hurst parameter in a fractional stochastic volatility model.

Contrarily, we speak of a *local* change-point problem if the assumption $\sigma_t^2 \in \Sigma(\mathfrak{a}, L_n)$ is at least violated once for a very short time period. This may even happen only at a single point $\theta \in (0, 1)$, for instance, if the volatility σ_t^2 exhibits a jump at θ, that is, $\Delta\sigma_\theta^2 > 0$. Related, but of different nature, are abrupt yet continuous adjustments in the volatility over a short time period, which are also covered by our theory. Note that local changes may happen multiple times, and they may even happen jointly with a global change on top, like presumably in Figure 1.

In this work, we present methods to test local and global alternatives, both relying on different foundations. A desirable property for a global approach is robustness with respect to local changes. This is crucial to distinguish the two problems, as it means that a test statistic to decide the question of global changes should not be affected by a fixed number of volatility jumps, which is nested in the hypothesis of no global change. On the other hand, an approach to test for local changes may very well be affected by a global change as well, for instance, at the very time of change or by an exceptionally large fluctuation. In practice, it is attractive to interconnect both methods which complement each other.

[4]One may as well consider the symmetric case $\mathfrak{a}' > \mathfrak{a}$.

We consider minimax-optimal testing and estimation in both situations, covering broad classes of volatility processes. For a conceptual introduction to minimax-optimal tests, let us focus first on the discrimination of smooth volatilities in the sense of (2) from volatilities with at least one jump. From a statistical perspective, the key question is which sizes of volatility jumps can be detected. For example, it is clear that we cannot detect jumps of arbitrarily small size. Loosely speaking, if we say that "no jump" is our null hypothesis H_0 and "there is a jump" is our alternative H_1, then we face the problem of *distinguishability between H_0 and H_1*. The minimum size b_n of a jump $\Delta\sigma_\theta^2$, such that we are still able to uniformly control the type I and type II errors, is called the *detection boundary*. If we are interested to test for the presence of jumps, we are thus led to consider for $\theta \in (0, 1)$ alternatives of the form

$$\mathcal{S}_\theta^J(\mathfrak{a}, b_n, L_n) = \{(\sigma_t^2)_{t\in[0,1]} | (\sigma_t^2 - \Delta\sigma_t^2)_{t\in[0,1]} \in \Sigma(\mathfrak{a}, L_n); |\Delta\sigma_\theta^2| \geq b_n\} \tag{3}$$

with a decreasing sequence b_n. We then address the testing problem

$$\begin{aligned} &H_0 : (\sigma_t^2(\omega))_{t\in[0,1]} \in \Sigma(\mathfrak{a}, L_n) \qquad \text{vs.} \\ &H_1 : \exists\theta \in (0, 1) \qquad \text{with } (\sigma_t^2(\omega))_{t\in[0,1]} \in \mathcal{S}_\theta^J(\mathfrak{a}, b_n, L_n). \end{aligned} \tag{4}$$

In this context, θ is commonly referred to as a *change-point*. Under the alternative at least one jump occurs, but we do not exclude multiple jumps. The dependence on ω in (4) is natural in the definition of the hypotheses, as different realizations might lead to different paths on [0, 1].

For the testing problem (4), we establish the minimax-optimal rate of convergence under high-frequency asymptotics. We follow the notion of minimax-optimality of statistical tests from the seminal contribution of Ingster (1993): For tests ψ that map a sample $\mathbf{X}_n$ to zero or one, where ψ accepts the null hypothesis H_0 if $\psi = 0$ and rejects it if $\psi = 1$, we consider the maximal type I error $\alpha_\psi(\mathfrak{a}) = \sup_{\sigma^2\in\Sigma(\mathfrak{a},L_n)} \mathbb{P}_\sigma(\psi = 1)$ and the maximal type II error $\beta_\psi(\mathfrak{a}, b_n) = \sup_{\theta\in(0,1)} \sup_{\sigma^2\in\mathcal{S}_\theta^J(\mathfrak{a},b_n,L_n)} \mathbb{P}_\sigma(\psi = 0)$ and define the global testing error as

$$\gamma_\psi(\mathfrak{a}, b_n) = \alpha_\psi(\mathfrak{a}) + \beta_\psi(\mathfrak{a}, b_n). \tag{5}$$

The primary interest now is to find tests that minimize $\gamma_\psi(\mathfrak{a}, b_n)$, given the boundary b_n. We aim to find sequences of tests ψ_n and boundaries b_n with the property that

$$\gamma_{\psi_n}(\mathfrak{a}, b_n) \to 0 \qquad \text{as } n \to \infty.$$

The smaller $b_n > 0$, the harder it is for a test to control the global testing error, that is, to distinguish between H_0 and H_1. It is thus natural to ask: given $\mathfrak{a}$, what is the minimal size of $b_n > 0$ such that

$$\lim_{n\to\infty} \inf_\psi \gamma_\psi(\mathfrak{a}, b_n) = 0 \tag{6}$$

holds? The optimal b_n^{opt} is called the *minimax distinguishable boundary*, and a sequence of tests ψ_n that satisfies (6) for all $b_n \geq b_n^{\mathrm{opt}}$ is called *minimax-optimal*. If $L_n = L$ in (2) is constant, we prove that $b_n \propto (n/\log(n))^{\frac{-\mathfrak{a}}{2\mathfrak{a}+1}}$ constitutes the minimax distinguishable boundary for testing (4) and that the constructed test attains minimax-optimality.[5] If L_n is indeed a sequence, the rate only slightly changes; see Section 4.1 for precise results.

For the proof of the lower bound, we simplify the problem by information-theoretic reductions passing to more informative subclasses of the parameter space. The lower bound established for the subclass then serves a fortiori as a lower bound in the more general and less informative model. After gradually transforming the problem by showing strong Le Cam equivalences of the considered subexperiment to more common situations with i.i.d. chi-square and Gaussian variables, the lower bound is proved by classical arguments based on the theory in Ingster and Suslina (2003).

The paper is organized as follows: Section 2 serves as an illustration for the benefit of cusum-based statistics in the simple but important model of a continuous Itô semimartingale with constant volatility. More involved, but also more important in practice, is the case where the volatility is both time-varying and random. Section 3 is devoted to this nonparametric local problem. As the volatility process is latent, which requires estimation based on smoothed squared increments of the semimartingale, this poses a complex statistical problem which to the best of the authors' knowledge had not been addressed so far. We establish a consistent test and derive a limit theorem under the hypothesis. The asymptotic analysis utilizes nonparametric change-point theory, stochastic calculus and bounds on the approximation error in the invariance principle. Our test allows us to distinguish paths with jumps from continuous paths under remarkably general smoothness assumptions on the hypothesis. In Section 3.2, we discuss the situation in which the underlying Itô semimartingale might have jumps as well. Section 4.1 provides the theory on minimax-optimality by discussing the lower bound, while Section 4.2 deals with the estimation of the location of the change in the volatility under the alternative. Finally, a minimax-optimal nonparametric test for the global problem is established in Section 5. A simulation study which investigates the finite-sample performance of the proposed methods and discusses some practical issues can be found in Section 6. Proofs are postponed to the Appendix and the supplementary material [Bibinger, Jirak and Vetter (2016b)].

2. Change-points in a parametric volatility model. The simplest model of a continuous-time Itô diffusion X is the case of no drift and a constant volatility:

$$X_t = X_0 + \int_0^t \sigma \, dW_s, \qquad t \in [0, 1], \tag{7}$$

[5]The notation $\propto$ means *proportional to*, that is, an identity if the right-hand side is multiplied with some constant.

where W denotes a standard Brownian motion. Throughout this work, the underlying process X is recorded at the discrete regular times $i\Delta_n$ with a mesh $\Delta_n \to 0$. To keep the notation simple, we assume to be on the fixed time interval [0, 1] and set $n = \Delta_n^{-1} \in \mathbb{N}$, so that we have observations $X_{i\Delta_n}$, $i = 0, \dots, n$.

Inference on the squared volatility σ^2 is usually based on the squares of the increments $\Delta_i^n X = X_{i\Delta_n} - X_{(i-1)\Delta_n}$. In case one is interested in changes in the volatility, a natural quantity to discuss is the cusum statistic, that is,

$$S_{n,m} = \frac{1}{\sqrt{n}} \sum_{i=1}^{m} \left(n(\Delta_i^n X)^2 - \sum_{j=1}^{n} (\Delta_j^n X)^2 \right), \qquad m \in \{1, \dots, n\}. \tag{8}$$

In order to derive the asymptotics of the cusum statistic, recall the functional (stable) central limit theorem for the realized volatility from observations of a continuous Itô semimartingale (1) by Jacod (1997). Under mild assumptions, we have

$$\sqrt{n}\left(\sum_{i=1}^{\lfloor nt \rfloor} (\Delta_i^n X)^2 - \int_0^t \sigma_s^2 \, ds \right) \to \int_0^t \sqrt{2}\sigma_s^2 \, dB_s, \qquad t \in [0, 1], \tag{9}$$

as $n \to \infty$, weakly in the Skorokhod space with a standard Brownian motion B, independent of W. In particular, if $\sigma_s = \sigma$ is constant, this result directly implies

$$S_{n,\lfloor nt \rfloor} \to \gamma(B_t - tB_1), \tag{10}$$

with $\gamma^2 = \lim_{n\to\infty} n \operatorname{Var}(\sum_{i=1}^n (\Delta_i^n X)^2) = 2\sigma^4$, which coincides with a standard cusum limit theorem in the vein of Phillips (1987). The quarticity estimator by Barndorff-Nielsen and Shephard (2002),

$$\hat{\gamma}^2 = (2n/3) \sum_{i=1}^{n} (\Delta_i^n X)^4,$$

may be used to obtain a self-normalizing version

$$(\hat{\gamma}^2)^{-1/2} S_{n,\lfloor nt \rfloor} \to B_t - tB_1, \tag{11}$$

where the limiting process is a standard Brownian bridge. A test for jumps (resp., structural breaks, change-points) in the volatility is then based on the test statistic

$$T_n = \sup_{m=1,\dots,n} \left| (\hat{\gamma}^2)^{-1/2} S_{n,m} \right|, \tag{12}$$

which (under the null hypothesis of a constant volatility) tends to a Kolmogorov–Smirnov law as $n \to \infty$; see Marsaglia et al. (2003). Under the alternative, T_n diverges almost surely.

Figure 2 shows an example in which we observe $n = 10{,}000$ values of a standard Brownian motion under the hypothesis, while under the alternative the volatility jumps at $t = 1/2$ from 1 to 1.1. Out of 10,000 Monte Carlo iterations under the

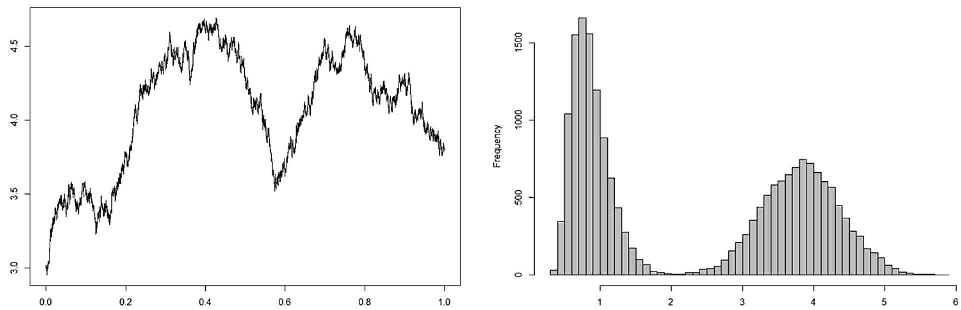

FIG. 2. *Left*: *One realized path of* X *with a structural break in the volatility at* $t = 1/2$. *Right*: *Empirical results for the test statistic from* 10,000 *iterations under the alternative and the hypothesis*.

hypothesis and the alternative, only 21 realizations of (12) under the hypothesis are larger than the minimum under the alternative. On the other hand, in 11 iterations the values under the alternative fall below the maximum of the generated values from the hypothesis. The cusum approach hence clearly allows us to separate hypothesis and alternative here, even for the relatively small volatility jump which is not readily identifiable from the path of X in Figure 2.

This observation is illustrated on the right-hand side of Figure 2. The left part of the histogram stems from the realizations under the hypothesis which closely track the limiting Kolmogorov–Smirnov law. The right part is due to realizations under the alternative instead. For larger volatility jumps, the right part moves further to the right such that the two distributions separate even more clearly. Thus, the Kolmogorov–Smirnov type test based on (12) allows us to test the hypothesis of a constant volatility against structural breaks in an efficient way.

Beyond this bridging of classical change-point analysis and structural breaks in a parametric volatility model, our main focus in the sequel is nonparametric: to distinguish volatility jumps from a continuous motion of volatility or to identify changes in the regularity exponent.

3. A nonparametric change-point test for the volatility against jumps.

3.1. *Construction of the test and its limit behavior under the hypothesis*. Suppose we observe a continuous Itô semimartingale (1) at the regular times $i\Delta_n$, $i = 0, \dots, n$. In this setting, we want to construct a test for (4). With the volatility process being time-varying, it becomes apparent from (9) that the test statistic (12) is not suitable to test H_0 against H_1. Our core idea is to utilize local two-sample t-tests over asymptotically small time blocks instead. As a first test statistic, we consider

$$V_n = \max_{i=0,\dots,\lfloor n/k_n\rfloor-2} |RV_{n,i}/RV_{n,i+1} - 1|, \tag{13}$$

where $k_n \to \infty$ is an auxiliary sequence of integers depending on n and

$$RV_{n,i} = \frac{n}{k_n} \sum_{j=1}^{k_n} (\Delta_{ik_n+j}^n X)^2, \qquad i = 0, \dots, \lfloor n/k_n \rfloor - 1, \tag{14}$$

is a rescaled local version of realized volatility over blocks of the partition $[ik_n\Delta_n, (i+1)k_n\Delta_n]$. The $RV_{n,i}$ estimate a blockwise constant proxy of the spot volatility $\sigma_{ik_n\Delta_n}$ on the respective blocks. Asymptotic properties of $RV_{n,i}$ were derived, for example, in Alvarez et al. (2012). By construction, a large distance between $RV_{n,i}$ and $RV_{n,i+1}$ suggests the presence of a jump, or of unsmooth breaks, in the volatility close to time $ik_n\Delta_n$. In order to obtain normalized statistics we work with ratios instead of differences. Thus, V_n appears to be a reasonable test statistic for our purpose.

The second test statistic is of the same nature as (13), but it takes all overlapping blocks of k_n increments into account:

$$V_n^* = \max_{i=k_n,\dots,n-k_n} \left| \frac{\frac{n}{k_n} \sum_{j=i-k_n+1}^{i} (\Delta_j^n X)^2}{\frac{n}{k_n} \sum_{j=i+1}^{i+k_n} (\Delta_j^n X)^2} - 1 \right|. \tag{15}$$

In comparison to nonparametric change-point approaches, like the one by Wu and Zhao (2007), both statistics (13) and (15) are based on ratios rather than differences. This makes sense intuitively, since we are not dealing with the typical additive error structure of time series models. In our setting, we have, for example, $n(\Delta_i^n X)^2 \approx \sigma_{i\Delta_n}^2 \xi_i^2$, $i = 1, \dots, n$, with i.i.d. χ_1^2-distributed random variables ξ_i^2, so that the volatility σ plays the role of a multiplicative error. Therefore, by computing ratios first, we basically deal with a maximum of identically distributed variables in the asymptotics. This is of key importance in order to obtain a distribution-free limit under the hypothesis.

In order to discuss the asymptotics of V_n and V_n^* under the null hypothesis, we need a couple of additional assumptions, all of which are rather mild and are covered by a variety of stochastic volatility models.

ASSUMPTION 3.1. *The following conditions on a and σ are satisfied*:

(1) *a and σ are locally bounded processes.*
(2) *σ is almost surely strictly positive, that is*, $\inf_{t\in[0,1]} \sigma_t^2 \geq \sigma_-^2 > 0$.
(3) *On the hypothesis*, $\Omega^c \subset \Omega$, *the modulus of continuity*

$$w_\delta(\sigma)_t = \sup_{s,r\leq t} \{|\sigma_s - \sigma_r| : |s-r| < \delta\}$$

is locally bounded in the sense that there exists $\mathfrak{a} > 0$ and a sequence of stopping times $T_n \to \infty$ such that $w_\delta(\sigma)_{(T_n\wedge 1)} \leq L_n \delta^{\mathfrak{a}}$, for some $\mathfrak{a} > 0$ and some (*a.s. finite*) *random variables L_n.*

In particular, Assumption 3.1 implies that $\sigma_t^2 \in \Sigma(\mathfrak{a}, L_n)$ for some n almost surely on Ω^c. Considering sequences L_n becomes important when developing lower bounds; see the first paragraph of Section 4.1 for a detailed explanation.

We choose the sequence $k_n \to \infty$, as $n \to \infty$, such that the following growth condition holds:

$$k_n^{-1}\Delta_n^{-\epsilon} + \sqrt{k_n}(k_n\Delta_n)^{\mathfrak{a}}\sqrt{\log(n)} \to 0, \tag{16}$$

for some $\epsilon > 0$ and with $\mathfrak{a} > 0$ from Assumption 3.1(3). Condition (16) consists of two assumptions which are reciprocal by nature. First, $k_n \to \infty$ faster than some power of n which gives a mild lower bound on the growth of k_n and ensures consistency of the estimates (14). The second condition gives an upper bound related to the continuity of σ. Naturally, the smaller $\mathfrak{a}$ (and the less smooth σ) is, the smaller we have to choose the size of the blocks over which we estimate σ.

THEOREM 3.2. *Set $m_n = \lfloor n/k_n \rfloor$ and $\gamma_{m_n} = [4\log(m_n) - 2\log(\log(m_n))]^{1/2}$. If Assumption 3.1 holds and k_n satisfies condition (16), then we have on Ω^c (under H_0)*

$$\sqrt{\log(m_n)}\big((k_n^{1/2}/\sqrt{2})V_n - \gamma_{m_n}\big) \xrightarrow{w} V, \tag{17}$$

$$\sqrt{\log(m_n)}(k_n^{1/2}/\sqrt{2})V_n^* - 2\log(m_n) - \frac{1}{2}\log\log(m_n) - \log(3) \xrightarrow{w} V, \tag{18}$$

where V follows an extreme value distribution with distribution function

$$\mathbb{P}(V \le x) = \exp\big(-\pi^{-1/2}\exp(-x)\big). \tag{19}$$

REMARK 3.3. It is remarkable that Theorem 3.2, in combination with condition (16), already allows us to distinguish between volatility paths with and without jumps, where we only require some smoothness $\mathfrak{a} > 0$ in Assumption 3.1(3). Note that less smooth paths require smaller block lengths k_n by (16) which reduces the rate in Theorem 3.2 and the power of the test. Most importantly, we can cope with standard models for σ. For a continuous semimartingale volatility, we have $\mathfrak{a} \approx 1/2$. In this case, we take $k_n \propto n^{1/2-\epsilon}$ for $\epsilon > 0$ and ϵ small in order to preserve the highest possible power. Similarly, for a Lipschitz volatility, that is, $\mathfrak{a} = 1$, one might choose $k_n \propto n^{2/3-\epsilon}$. Thus, the choice of the block length is close to the optimal window size for spot volatility estimation.

As we show in Theorem 4.3 that V_n^* and V_n diverge almost surely under the alternative, Theorem 3.2 provides a consistent test with asymptotic power 1, if we use the critical values from the limit law under the hypothesis.

3.2. *A test in the presence of jumps in the observed process.* In order to provide an approach which is feasible in various economic applications, an important aim is to account for possible jumps in the process X as well. We consider a general Itô semimartingale

$$X_t = X_0 + \int_0^t a_s\,ds + \int_0^t \sigma_s\,dW_s + \int_0^t \int_{\mathbb{R}} \kappa\big(\delta(s,x)\big)(\mu-\nu)(ds,dx) + \int_0^t \int_{\mathbb{R}} \bar{\kappa}\big(\delta(s,x)\big)\mu(ds,dx), \tag{20}$$

where a truncation function κ, $\bar{\kappa}(x) = x - \kappa(x)$, separates large jumps from compensated small jumps. The compensating intensity measure ν of the Poisson random measure μ admits the form $\nu(ds,dx) = ds \otimes \lambda(dx)$ for a σ-finite measure λ. Our notation follows Jacod (2008).

ASSUMPTION 3.4. *Suppose Assumption 3.1 for the continuous part of X holds. Assume that $\sup_{\omega,x} |\delta(s,x)|/\gamma(x)$ is locally bounded for some deterministic nonnegative function γ which satisfies for some $r < 2$:*

$$\int_{\mathbb{R}} \big(1 \wedge \gamma^r(x)\big)\lambda(dx) < \infty. \tag{21}$$

In condition (21), r is a jump activity index which bounds the pathwise generalized Blumenthal–Getoor index from above. Imposing $r < 1$ restricts to jumps of finite variation and $r = 0$ to finite jump activity. We develop test statistics which are robust against jumps by using the truncation principle which was originally introduced for the estimation of integrated volatility by Mancini (2009). The analogue of (13) with truncated squared increments is

$$V_{n,u_n} = \max_{i=0,\dots,\lfloor n/k_n\rfloor-2} |TRV_{n,u_n,i}/TRV_{n,u_n,i+1} - 1| \qquad \text{where} \tag{22}$$

$$TRV_{n,u_n,i} = \frac{n}{k_n}\sum_{j=1}^{k_n} (\Delta_{ik_n+j}^n X)^2 \mathbb{1}_{\{|\Delta_{ik_n+j}^n X| \le u_n\}}, \qquad i = 0,\dots,\lfloor n/k_n\rfloor - 1. \tag{23}$$

The truncation sequence, $u_n \propto n^{-\tau}$, $\tau \in (0, 1/2)$, is used to exclude large squared increments which are ascribed to jumps. In the same way, we can generalize statistic (15) with overlapping blocks:

$$V_{n,u_n}^* = \max_{i=k_n,\dots,n-k_n} \left| \frac{\frac{n}{k_n}\sum_{j=i-k_n+1}^{i} (\Delta_j^n X)^2 \mathbb{1}_{\{|\Delta_j^n X| \le u_n\}}}{\frac{n}{k_n}\sum_{j=i+1}^{i+k_n} (\Delta_j^n X)^2 \mathbb{1}_{\{|\Delta_j^n X| \le u_n\}}} - 1 \right|. \tag{24}$$

We prove that truncation is an appropriate concept to asymptotically eliminate the influence by jumps, at least under certain restrictions on the jump activity, on k_n

and on τ. In particular, under the hypothesis we obtain the same limit behavior of the test statistics as in Theorem 3.2.

PROPOSITION 3.5. *Suppose $k_n \propto n^\beta$ for $0 < \beta < 1$, such that condition (16) is satisfied. Furthermore, grant Assumption 3.4 for some*

$$r < \min\big(2(2 - \tau^{-1}(1 - \beta/2)), \tau^{-1} \min(1/2, 1 - \beta)\big) \tag{25}$$

as well. Then, with $m_n = \lfloor n/k_n \rfloor$ and $\gamma_{m_n} = [4 \log(m_n) - 2 \log(\log(m_n))]^{1/2}$ as before, and if either $r = 0$ or the jump process is a time-inhomogeneous Lévy process, we have on Ω^c (under H_0)

$$\sqrt{\log(m_n)}\big((k_n^{1/2}/\sqrt{2})V_{n,u_n} - \gamma_{m_n}\big) \xrightarrow{w} V, \tag{26}$$

$$\sqrt{\log(m_n)}(k_n^{1/2}/\sqrt{2})V^*_{n,u_n} - 2\log(m_n) - \frac{1}{2}\log\log(m_n) - \log(3) \xrightarrow{w} V, \tag{27}$$

where V is distributed according to (19).

REMARK 3.6. A simple computation shows that a necessary condition in order for (25) to hold is $r < 1$, but in general further conditions on the interplay between r, β and τ have to be taken into account. Thus, this restriction on the jump activity is stronger than the usual $r < 1$ for truncated realized volatility as in Jacod (2008). This is because we use a maximum in (24) instead of linear estimators.

REMARK 3.7. It appears to be most relevant from an applied perspective that the test based on (24) copes with finite activity jumps. In this case, (25) reads as $\tau > 1/2 - \beta/4$, and the only requirement is that u_n is not chosen too large. Beyond the finite activity case, the choice of the tuning parameters becomes more complex and depends on the statistician's interest. Ideally, one would choose β large to secure a high power, but it should not become too large as (25) then is more restrictive, and interesting models on the jumps might be ruled out. As an equilibrium choice $\beta \approx 1/2$ is recommended in which case $\tau \approx 1/2$ is optimal, leading to the condition $r < 1$. Choosing τ close to $1/2$ and u_n small improves the precision of the localized truncated realized volatilities (23). As an exact choice for u_n one typically picks it sufficiently large to just not interfere with the continuous component of X_t. Based on extreme value theory for Gaussian sequences,

$$u_n = C\sqrt{2\log(n)}n^{-1/2} \tag{28}$$

with constant C is a suitable choice. Once it is guaranteed that $C > \sup_{t\in[0,1]} \sigma_t^2$, almost surely no increments of the continuous part of (20) are truncated. In practice, some suitable upper bound C for the volatility can be obtained from historical data.

4. Asymptotic minimax-optimality results for the local change problem.

4.1. *Consistency and minimax-optimal rate of convergence.* In this section, it becomes important that the *stochastic* squared volatility processes lie in $\Sigma(\mathfrak{a}, L_n)$, defined in (2), under H_0, where we take into account strictly positive increasing sequences L_n. This is crucial as we cannot assume the random processes to be members of a fixed Hölder class in general. For instance, if σ_t^2 satisfies

$$\mathbb{E}\big[|\sigma_t^2 - \sigma_s^2|^{\mathfrak{b}}\big] \leq C|t-s|^{\gamma+\mathfrak{b}\mathfrak{a}} \qquad \text{for some } \mathfrak{b}, C > 0 \text{ and } \gamma > 1,$$

then the Kolmogorov–Čentsov theorem implies $\lim_{n\to\infty} \mathbb{P}((\sigma_t^2)_{0\leq t\leq 1} \in \Sigma(\mathfrak{a}, L_n)) = 1$, provided $L_n \to \infty$ arbitrarily slowly. Hence, up to a negligible set, $\Sigma(\mathfrak{a}, L_n)$ contains the *paths* generated by a huge number of popular volatility models when considering $L_n \to \infty$. On the other hand, if $L_n = L$ is fixed, we are in the familiar framework of Hölder classes.

At this stage, we integrate alternatives where the volatility is less smooth than under the hypothesis, but which do not necessarily include jumps. The statistical devices developed above may be applied to discriminate H_0 from alternatives without jumps where, until some change-point $\theta \in [0, 1)$, the process $(\sigma_{t\wedge\theta}^2)$ behaves as a process in $\Sigma(\mathfrak{a}, L_n)$. After θ, the regularity exponent drops to some $0 < \mathfrak{a}' < \mathfrak{a}$. Since $\Sigma(\mathfrak{a}, L_n) \subset \Sigma(\mathfrak{a}', L_n)$, we have to ensure that the processes "exploit their roughness" (close to θ), such that in particular $(\sigma_t^2)_{t\in[0,1]} \notin \Sigma(\mathfrak{a}, L_n)$. To describe the alternative sets formally, define

$$\Delta_h^{\mathfrak{a}'} f_t = \frac{f_{t+h} - f_t}{h^{\mathfrak{a}'}}, \qquad t \in [0, 1], h \in [0, 1-t].$$

Then the set of possible alternatives is given by

$$\mathcal{S}_\theta^{\mathrm{R}}(\mathfrak{a}, \mathfrak{a}', b_n, L_n) = \Big\{ \big(\sigma_{t\wedge\theta}^2\big)_{t\in[0,1]} \in \Sigma(\mathfrak{a}, L_n) \,\big|\, \inf_{h\in[0,2k_n\Delta_n)} \Delta_h^{\mathfrak{a}'} \sigma_\theta^2 \geq b_n \\ \text{or} \sup_{h\in[0,2k_n\Delta_n)} \Delta_h^{\mathfrak{a}'} \sigma_\theta^2 \leq -b_n \Big\},$$

and we consider the testing problem

$$\begin{aligned} &H_0 : \big(\sigma_t^2(\omega)\big)_{t\in[0,1]} \in \Sigma(\mathfrak{a}, L_n) \qquad \text{vs.} \\ &H_1^{\mathrm{R}} : \exists \theta \in [0, 1) | \big(\sigma_t^2(\omega)\big)_{t\in[0,1]} \in \mathcal{S}_\theta^{\mathrm{R}}(\mathfrak{a}, \mathfrak{a}', b_n, L_n). \end{aligned} \tag{29}$$

Since k_n will be chosen to depend on L_n and $\mathfrak{a}$, the dependence of $\mathcal{S}_\theta^{\mathrm{R}}$ on k_n above is not indicated in our notation.

Let us elaborate on the specific form of the alternative sets. In general, it is impossible to test $\Sigma(\mathfrak{a}, L_n)$ against $\Sigma(\mathfrak{a}', L_n)$ for $\mathfrak{a} > \mathfrak{a}'$, and it is necessary to consider special subsets of $\Sigma(\mathfrak{a}', L_n)$. Intuitively, it is clear that one needs at least to remove $\Sigma(\mathfrak{a}, L_n)$ from $\Sigma(\mathfrak{a}', L_n)$, but this is not sufficient. In fact, one needs to focus on those functions which exploit their roughness in a certain sense; cf. Hoffmann and Nickl (2011) for a detailed discussion in a related context. Geometrically, this means that the functions of interest are those with discontinuities or with rough behavior as characterized in $\mathcal{S}_\theta^{\mathrm{R}}$ (or which fluctuate considerably more, like the ones considered in Section 5). However, as the sample size n grows, we only require the difference quotient to exceed a level b_n that becomes smaller and smaller. Assuming that the exceedance period of the difference quotient in $\mathcal{S}_\theta^{\mathrm{R}}$ is at least two block lengths finally ensures that our blockwise comparison in the test statistics (13) and (15) is able to detect the roughness, also for $\theta = 0$.

For the testing problems (4) and (29), we first present a negative result which also serves as a minimax lower bound for the problem described in (6).

THEOREM 4.1. *Assume that* $\mathfrak{a} > \mathfrak{a}' > 0$ *and* $\inf_t \sigma_t^2 \geq \sigma_-^2 > 0$. *Consider either set of hypotheses* $\{H_0, H_1\}$ *or* $\{H_0, H_1^{\mathrm{R}}\}$. *Then for*

$$b_n \leq \bigl(n/\log(m_n)\bigr)^{-\frac{\mathfrak{a}-\mathfrak{a}'}{2\mathfrak{a}+1}} (L_n)^{\frac{2\mathfrak{a}'+1}{2\mathfrak{a}+1}} \sigma_-^2, \tag{30}$$

with $\mathfrak{a}' = 0$ *for* H_1, *we have in both cases* $\lim_{n\to\infty} \inf_\psi \gamma_\psi(\mathfrak{a}, b_n) = 1$.

REMARK 4.2. In the proof of Theorem 4.1, we show that the local change-point problem is asymptotically equivalent (in strong Le Cam sense) to a high-dimensional location problem which is of independent interest. We then use this result to deduce the lower bound.

Theorem 4.1 reveals that it is impossible to construct a minimax-optimal test in the sense of (6) if b_n is bounded as in (30). In Theorem 4.3, we shall establish a corresponding upper bound up to a constant, and thus the right-hand side of (30) already gives the optimal rate for the minimax distinguishable boundary. Observe that based on V_n^* from (15), we can obtain the following test $\psi^\diamond$:

$$\psi^\diamond\bigl((X_{i\Delta_n})_{0\leq i\leq n}\bigr) = 1 \qquad \text{if } V_n^* \geq 2C^\diamond \sqrt{2\log(m_n^\diamond)/k_n^\diamond}, \tag{31}$$

$$\text{where } C^\diamond > 2 \text{ and } k_n^\diamond = \Bigl(\sqrt{\log(m_n^\diamond)} n^{\mathfrak{a}}/L_n\Bigr)^{\frac{2}{2\mathfrak{a}+1}}, \qquad m_n^\diamond = \lfloor n/k_n^\diamond \rfloor. \tag{32}$$

Alternatively, one might base a test on V_n from (13).

To simplify the discussion, we restrict ourselves to positive volatility jumps, $\inf_t \Delta\sigma_t \geq 0$, which appears natural from an economic point of view. We point out that an analogous result can be shown for negative, or positive and negative jumps, which, however, requires a further technical condition—stating that successive jumps do not cancel in case of multiple jumps close to each other.

THEOREM 4.3. *Consider* (4) *with* $\inf_t \Delta\sigma_t \geq 0$, *or* (29) *with* $0 < \mathfrak{a}' < \mathfrak{a} \leq 1$ *and* $L_n = \mathcal{O}((n/k_n^\diamond)^{\mathfrak{a}-\mathfrak{a}'})$. *If*

$$b_n^\diamond > \Bigl(4C^\diamond\sqrt{2} \sup_{t\in[0,1]} \sigma_t^2 + 2\Bigr)\bigl(n/\log(m_n)\bigr)^{-\frac{\mathfrak{a}-\mathfrak{a}'}{2\mathfrak{a}+1}} (L_n)^{\frac{2\mathfrak{a}'+1}{2\mathfrak{a}+1}}, \tag{33}$$

where $k_n^\diamond$, $m_n^\diamond$ *and* $C^\diamond$ *are as in* (32), *then* $\lim_{n\to\infty}\gamma_{\psi^\diamond}(\mathfrak{a}, b_n^\diamond)=0$. *This implies*

$$b_n^{\mathrm{opt}} \propto \big(n/\log(m_n)\big)^{-\frac{\mathfrak{a}-\mathfrak{a}'}{2\mathfrak{a}+1}}(L_n)^{\frac{2\mathfrak{a}'+1}{2\mathfrak{a}+1}}.$$

REMARK 4.4. If $L_n = L$ defined in (2) is a deterministic constant, we get the minimax distinguishable boundary $b_n \propto (n/\log(n))^{\frac{-(\mathfrak{a}-\mathfrak{a}')}{2\mathfrak{a}+1}}$.

REMARK 4.5. Volatility paths in $\mathcal{S}_\theta^{\mathrm{R}}$ locally have a rough—but still continuous—increase or decrease and cannot be distinguished from volatility paths with jumps at or below the boundary b_n stated in (30). Still, even though both alternatives H_1 and H_1^{R} are in this sense intimately connected, we cannot include $\mathcal{S}_\theta^J$ in $\mathcal{S}_\theta^{\mathrm{R}}$ by setting $\mathfrak{a}'=0$, since for $\mathcal{S}_\theta^{\mathrm{R}}$ we require the roughness to persist over an (asymptotically small) interval.

Let us also point out that in this testing problem, the union of hypothesis and alternative does not cover the set of all possible volatility paths. One situation of interest in which $\sigma_t^2 \notin \{\mathcal{S}_\theta^{\mathrm{R}} \cup \mathcal{S}_\theta^J \cup \Sigma(\mathfrak{a}, L_n)\}$ is the case of fractional processes with a Hurst parameter $\mathfrak{a}' < \mathfrak{a}$. This different situation is addressed in Section 5.

4.2. *Estimating the change-point.* Once one has opted to reject the null hypothesis of no change, the actual locations of jumps become of interest for further inference. Such location problems have been extensively discussed in the literature in different frameworks; see, for instance, Csörgő and Horváth (1997) and Müller (1992).

4.2.1. *One change-point alternative.* First, we restrict ourselves to the "one change-point alternative" involving a jump in the volatility, that is, we specify the alternative hypothesis H_1^* as

$$H_1^*: \big|\sigma_\theta^2 - \sigma_{\theta-}^2\big| =: \delta_n > 0 \qquad \text{for a unique } \theta \in (0,1).$$

The jump size δ_n may be fixed or we consider a decreasing sequence (δ_n). To assess the possible time of change, we use slightly modified versions of the building blocks of the test statistic V_n^* from (15), namely

$$V_{n,i}^\diamond = \frac{1}{\sqrt{k_n}}\Bigg|\sum_{j=i-k_n+1}^{i} n\big(\Delta_j^n X\big)^2 - \sum_{j=i+1}^{i+k_n} n\big(\Delta_j^n X\big)^2\Bigg|,$$

for $i = k_n, \dots, n-k_n$, and $V_{n,i}^\diamond = 0$ else. The possible time of the change is then estimated via

$$n\widehat{\theta}_n = \operatorname*{argmax}_{i=k_n,\dots,n-k_n} V_{n,i}^\diamond. \tag{34}$$

In contrast to the construction of V_n^* we use a simpler unweighted version here, but one could also consider the rescaled versions as in V_n^*, and we conjecture that the following theoretical results of these estimators coincide. Switching from ratios to differences, however, simplifies the analysis and yet allows us to obtain the following properties.

PROPOSITION 4.6. *Assume that the assumptions of Theorem 3.2 hold and that H_1^* is valid. Then, for $\delta_n \geq 2k_n^{-1/2}\sqrt{\log(n)}\sup_{t\in[0,1]}\sigma_t^2$, we have that*

$$|\widehat{\theta}_n - \theta| = \mathcal{O}_{\mathbb{P}}\left(\frac{\sqrt{k_n \log(n)}}{n\delta_n}\right). \tag{35}$$

If δ_n does not tend to zero, the condition on δ_n in the proposition is always satisfied.

REMARK 4.7. The estimator extends to more general situations:

(i) In case of jumps in the process X, we use truncation as in (24), and Proposition 4.6 then applies to the generalized estimator under the assumptions of Proposition 3.5.

(ii) In the setup of continuous breaks under alternative $\mathcal{S}_\theta^{\mathrm{R}}$ the same estimator is consistent only when the volatility is $\mathfrak{a}$-regular except on a small interval around θ. When the maximal length of this interval is $\sqrt{k_n \log(n)}/(n\delta_n)$, (35) applies when we replace δ_n by $\delta_n(k_n\Delta_n)^{\mathfrak{a}'}$. Clearly, when the volatility violates $\mathfrak{a}$-regularity over longer time horizons such a result is not available.

Obviously, the quality of the estimator $\widehat{\theta}_n$ depends on the bandwidth k_n, and the smaller, the better. This is the complete opposite case compared to the test based on statistic V_n^*, where a larger choice of k_n increases the power. This is no contradiction, since both problems have a different, essentially reciprocal nature. Also note that k_n cannot be chosen arbitrarily small; see condition (16).

While classical estimators as the argmax of statistic (8) attain a standard $\sqrt{n}$-rate, corresponding to $k_n \approx n$, our nonparametric localization approach allows for improved convergence rates as known for state-of-the-art change-point estimators like, for example, in Aue et al. (2009). The following proposition sheds light on optimal convergence rates for the estimation problem.

PROPOSITION 4.8. *If the assumptions of Proposition 4.6 hold, then a consistent estimator for θ does not exist in the case that $\delta_n = \mathcal{O}(\sqrt{\log(n)}k_n^{-1/2})$.*

4.2.2. *Multiple change-point alternatives.* In the sequel, we demonstrate how the previous theory can be extended to multiple change-points. To keep this exposition at a reasonable length, we focus on the alternative where the volatility exhibits jumps. From a general perspective, multiple change-point detection is typically a much more challenging multiple testing problem than the one change-point detection problem. The main probabilistic difficulty usually lies in controlling the overall stochastic error. Fortunately, in the present context we have already dealt with the overall stochastic error successfully; see Theorems 3.2 and 4.3. Thus, treating the multiple change-point problem only requires small adjustments. For $N \in \mathbb{N}$, let

$$0 < \theta_1 < \cdots < \theta_N < 1, \qquad \Theta_N = \{\theta_1, \ldots, \theta_N\}. \tag{36}$$

We then consider the alternative

$$H_1^* : \left|\sigma_{\theta_i}^2 - \sigma_{\theta_i -}^2\right| =: \delta_{n,i} > 0 \qquad \text{for } 1 \leq i \leq N.$$

The number of changes N is unknown to the experimenter, and the goal is to provide uniformly consistent estimates for the multiple change-points. To this end, given an index set $\mathcal{I} \subseteq \{k_n, \ldots, n - k_n\}$, we define in analogy to (34)

$$n\widehat{\theta}_n(\mathcal{I}) = \operatorname*{argmax}_{i \in \mathcal{I}} V_{n,i}^{\diamond}. \tag{37}$$

Based on the test $\psi^{\diamond}$ introduced in (31), we propose the following algorithm for the multiple change-point detection.

ALGORITHM 4.9.
Initialize *Set* $\hat{\mathcal{I}} = \{k_n, \ldots, n - k_n\}$, $\hat{\Theta} = \varnothing$ *and select* $r_n = \mathcal{O}(n)$ *such that* $k_n = \mathcal{O}(r_n)$, $k_n \to \infty$.

(i) *If* $\psi^{\diamond}((X_{i\Delta_n})_{i \in \hat{\mathcal{I}}}) = 0$, *stop and return* $\hat{\mathcal{I}}$ *and* $\hat{\Theta}$. *Otherwise go to step* (**ii**).

(ii) *Estimate one time of change* θ *using* $\widehat{\theta}_n(\hat{\mathcal{I}})$ *from* (37).

(iii) *Set* $\hat{\mathcal{I}} = \hat{\mathcal{I}} \setminus \{\lfloor \widehat{\theta}_n(\hat{\mathcal{I}})n \rfloor - r_n, \ldots, \lceil \widehat{\theta}_n(\hat{\mathcal{I}})n \rceil + r_n\}$, $\hat{\Theta} = \hat{\Theta} \cup \{\widehat{\theta}_n(\hat{\mathcal{I}})\}$, *and go to step* (**i**).

Algorithm 4.9 is a sequential top-down algorithm, similar in spirit to the well-known bisection methods. Observe that it not only returns an estimate for the set of change-points Θ_N, but also the set $\hat{\mathcal{I}}$ of noncontaminated indices, which can be used for further inference. The following result provides consistency of the proposed set estimators.

PROPOSITION 4.10. *On the assumptions of Theorem* 3.2 *and under* H_1^*, *if*:

(i) *for some* $N' = \mathcal{O}(n/r_n)$, *it holds that* $\inf_{1 \leq i \leq N-1} |\theta_{i+1} - \theta_i| \geq (N')^{-1}$,

(ii) $\inf_{1 \leq i \leq N} \delta_{n,i} \geq 2k_n^{-1/2}\sqrt{\log(n)} \sup_{t \in [0,1]} \sigma_t^2$,

then we have consistency of $\hat{\Theta}$: $\mathbb{P}(|\hat{\Theta}| = N) \to 1$, *and* $\sup_{n=1,\ldots,N} |\widehat{\theta}_n - \theta_n| \xrightarrow{\mathbb{P}} 0$.

Note that we can allow for increasing $N = \mathcal{O}(N')$ as the sample size n increases. The bound in condition (ii) is optimal if k_n is selected in the optimal way $k_n \propto (\sqrt{\log(n)}n^{\mathfrak{a}})^{\frac{2}{2\mathfrak{a}+1}}$.

5. Change-point test and asymptotic results for the global change problem. In Section 3, we present methods to test the hypothesis of $\mathfrak{a}$-regular volatilities against local alternatives of less regular volatilities. An important example have been volatility jumps which violate the hypothesis for any $\mathfrak{a} > 0$. If the hypothesis is rejected for a pre-specified $\mathfrak{a}$, however, the test does not reveal if this is due to a volatility jump or due to a change of the regularity where the volatility is $\mathfrak{a}$-regular on $[0, \theta)$ and $\mathfrak{a}'$-regular on $(\theta, 1]$ with $\mathfrak{a}' < \mathfrak{a}$. In case of global changes the test (under certain conditions) rejects as well. Moreover, the alternatives in Section 4 do not cover a change in the Hurst parameter of a fractional volatility process.

The latter constitutes a different testing problem of great interest which is addressed in this section. We develop a new test which discriminates between volatilities which are $\mathfrak{a}$-regular on $[0, 1]$, except for a finite number of discontinuities, and volatilities where at $\theta \in (0, 1)$ the regularity exponent drops to $\mathfrak{a}' < \mathfrak{a}$ such that $\mathfrak{a}$-regularity is permanently violated on $[\theta, 1]$. We consider processes which satisfy the following regularity assumptions.

ASSUMPTION 5.1. (i) *Drift and volatility process in* (1) *are càdlàg with* $\inf_{t\in[0,1]} \sigma_t^2 > 0$.

(ii) *Only for a finite set* $\mathcal{T} = \{\tau_1, \dots, \tau_N\}$ *of stopping times we have* $\sigma_{\tau_j} \neq \lim_{t<\tau_j, t\to\tau_j} \sigma_t$, *and we assume* $\sup_{j=1,\dots,N} |\sigma_{\tau_j} - \lim_{t<\tau_j,t\to\tau_j} \sigma_t| \le K$ *with a constant* $K < \infty$.

(iii) $\sigma_t^2 = \nu_t + \varrho_t$ *with the* σ*-algebra* $\sigma(\varrho_s, 0 \le s \le 1)$ *being independent of* $\sigma(W_s, \nu_s, 0 \le s \le 1)$. *For any* $0 \le s, \tau \le 1$ *with* $[s, \tau] \cap \mathcal{T} = \varnothing$, *we have with a constant* K *and some* $\epsilon > 0$:

$$\big(\mathbb{E}\big[|\nu_\tau - \nu_s|^8\big]\big)^{1/8} \le K|\tau - s|^{(1/2+\epsilon)}. \tag{38}$$

(iv) *For* $\Delta_i^n \varrho = n \int_{(i-1)\Delta_n}^{i\Delta_n} (\varrho_s - \varrho_{s-\Delta_n})\, ds$, *it holds that*

$$\max_{2\le m\le n} \left| \sum_{i=2}^{m} \big((\Delta_i^n \varrho)^2 - \mathbb{E}\big[(\Delta_i^n \varrho)^2\big]\big) \right| = \mathcal{O}_{\mathbb{P}}(\sqrt{n}). \tag{39}$$

TESTING PROBLEM 5.2. *Hypothesis*: The process $(\varrho_t)_{t\in[0,1]}$ is $\mathfrak{a}$-regular in the following sense:

$$\mathbb{E}\big[(\Delta_i^n \varrho)^2\big] = \vartheta_n^2 + \mathcal{O}\big(n^{-1/2}\big) \qquad \text{for all } i, \vartheta_n^2 \le K n^{-2\mathfrak{a}}, \tag{40}$$

$$\big(\mathbb{E}\big[(\Delta_i^n \varrho)^8\big]\big)^{1/8} \le K n^{-\mathfrak{a}} \qquad \text{for all } i, \tag{41}$$

for some constant K and $\mathfrak{a} > 0$. Thus, on the hypothesis $(\sigma_t^2)_{t\in[0,1]}$ is $\min(1/2, \mathfrak{a})$-regular.

Alternative: For $\theta \in (0, 1)$, $(\varrho_{t\wedge\theta})_{t\ge 0}$ satisfies (40) and (41) on $[0, \theta)$. For some $\mathfrak{a}' < \min(1/2, \mathfrak{a})$, $b_n' > 0$ and for all $i\Delta_n \ge \theta$, we have

$$\mathbb{E}\big[(\Delta_i^n \varrho)^2\big] \ge b_n' n^{-2\mathfrak{a}'}. \tag{42}$$

Contrary to the setup of Section 3, the hypothesis allows for a finite number of discontinuities in $(\nu_t)_{t\in[0,1]}$. On the other hand, under the alternative the volatility *permanently "exploits its roughness"* in the sense of violating (40) permanently over $[\theta, 1]$. Condition (38) can be extended also to the case where ν is an Itô semimartingale with jumps of finite activity. Assumption 5.1(iv) ensures that $\mathbb{E}[(\varrho_t - \varrho_s)^2]$ does not vary too much over time. This condition is obsolete for $\mathfrak{a} > 1/4$.

Our setup covers many stochastic volatility models, and in particular it applies to discriminate fractional volatility models with different Hurst parameters. While most fractional stochastic volatility models include independence of log-price and volatility, possible dependence (leverage) is usually allowed in the literature when one uses a semimartingale volatility. In this light, the decomposition in Assumption 5.1 appears natural, where $(\varrho_t)_{t\geq 0}$ is independent of $(W_t)_{t\geq 0}$ and $(\nu_t)_{t\geq 0}$, whereas $(\nu_t)_{t\geq 0}$ comprises leverage. The following simple example reveals the interplay of Assumption 5.1(iv) and (40) and the Hurst parameter.

EXAMPLE 5.3 (Fractional Brownian motion). Suppose that (ϱ_t) is a fractional Brownian motion with a Hurst parameter H. Let $\xi_k = n^H(\varrho_{k\Delta_n} - \varrho_{(k-1)\Delta_n})$, $k \geq 1$. Then, see, for example, Embrechts and Maejima (2000):

(i) $\varrho_r - \varrho_s \overset{d}{=} \varrho_{r-s} \overset{d}{=} (r-s)^H \xi_1$, $r \geq s$,
(ii) $|\mathbb{E}[\xi_0\xi_k]| \leq K(k+1)^{2H-2}$, $k \geq 1$, for some constant K.

The scaling property (i) implies (40) and (41) with $H = \mathfrak{a}$. If $1/4 \leq H \leq 1$, then (i) suffices also to guarantee the validity of Assumption 5.1(iv). If $0 < H < 1/2$, we have

$$\mathbb{E}\left[\max_{2\leq m\leq n}\left|\sum_{i=2}^{m}\left((\Delta_i^n \varrho)^2 - \mathbb{E}[(\Delta_i^n \varrho)^2]\right)\right|\right]$$
$$\leq n^{-2H}\int_{[0,1]^2}\mathbb{E}\left[\max_{2\leq m\leq n}\left|\sum_{k=1}^{m-1}\left(\xi_{k+r}\xi_{k+s} - \mathbb{E}[\xi_{k+r}\xi_{k+s}]\right)\right|\right]dr\,ds.$$

Then (ii) together with the joint Gaussianity of (ξ_{k+r}, ξ_{k+s}) implies that $(\xi_{k+r}\xi_{k+s})$ is a short memory sequence. In particular, using the results in Arcones (1994) and Móricz, Serfling and Stout (1982) we obtain

$$n^{-2H}\sup_{0\leq r,s\leq 1}\mathbb{E}\left[\max_{2\leq m\leq n}\left|\sum_{k=1}^{m-1}\left(\xi_{k+r}\xi_{k+s} - \mathbb{E}[\xi_{k+r}\xi_{k+s}]\right)\right|\right] = \mathcal{O}\left(n^{-2H+1/2}\right),$$

and hence the validity of Assumption 5.1(iv).

More generally, our setup includes prominent realistic volatility models, such as fractional Ornstein–Uhlenbeck processes discussed in Comte and Renault (1998) which will be considered in Section 6.

Because of the different permanent nature of the change we derive a statistical device to address Testing problem 5.2 which differs from the methods in Section 3. In particular, we propose a global cusum-type test statistic instead of localized ones. Define for $i = 2, \dots, n$,

$$Q_{n,i} = n^2 \Big(\big((\Delta_i^n X)^2 - (\Delta_{i-1}^n X)^2 \big)^2 - \frac{2}{3} \big((\Delta_i^n X)^4 + (\Delta_{i-1}^n X)^4 \big) \Big). \tag{43}$$

Our cusum-type test statistic based on the statistics (43) is

$$V_n^\dagger = \frac{1}{\sqrt{n-1}} \max_{m=2,\dots,n} \left| \sum_{i=2}^m \left(Q_{n,i} - \frac{\sum_{i=2}^n Q_{n,i}}{n-1} \right) \right|. \tag{44}$$

Intuitively, statistics $Q_{n,i}$ are small if $|\sigma_{i\Delta_n}^2 - \sigma_{(i-1)\Delta_n}^2|$ is small and become larger the larger $|\sigma_{i\Delta_n}^2 - \sigma_{(i-1)\Delta_n}^2|$. The regularity $\mathfrak{a}$ thus directly influences the average behavior of the $Q_{n,i}$, and a change at time θ can be detected by (44).

THEOREM 5.4. *Suppose that we are under the hypothesis of Testing problem 5.2 and that Assumption 5.1 holds. Then the cusum-process associated with statistic (44) satisfies the functional convergence*

$$\frac{\sqrt{3/80}}{\sqrt{n-1}} \left(\sum_{i=2}^{\lfloor nt \rfloor} \left(Q_{n,i} - \frac{\sum_{i=2}^n Q_{n,i}}{n-1} \right) \right) \xrightarrow{\omega-(st)} \left(\int_0^t \sigma_s^4 \, dB_s - t \int_0^1 \sigma_s^4 \, dB_s \right), \tag{45}$$

stable with respect to $\mathcal{F}$, weakly in the Skorokhod space, where (B_s) denotes a Brownian motion independent of $\mathcal{F}$.

As an immediate consequence of Theorem 5.4, we obtain

$$\sqrt{\frac{3}{80}} V_n^\dagger \xrightarrow{\omega-(st)} V^\dagger, \qquad V^\dagger = \sup_{0 \le t \le 1} \left| \int_0^t \sigma_s^4 dB_s - t \int_0^1 \sigma_s^4 dB_s \right|. \tag{46}$$

In order to construct a test $\psi^\dagger$ based on $V_n^\dagger$, the key object are the (conditional) quantiles

$$q_\alpha(V^\dagger | \mathcal{F}) = \inf \{ x \ge 0 : \mathbb{P}(V^\dagger \le x | \mathcal{F}) \ge \alpha \}. \tag{47}$$

The latter depend on the unknown volatility $(\sigma_t)_{t \in [0,1]}$ and are therefore a priori not available. One way to circumvent this problem is to estimate $(\sigma_t)_{t \in [0,1]}$ locally and to apply an appropriate bootstrap procedure to approximate $q_\alpha(V^\dagger | \mathcal{F})$. Alternatively, as seen in Proposition 5.6, a standardized version of (45) converges to a Kolmogorov–Smirnov limit law under an additional condition, and hence allows for the construction of a test as well. Consider for $K_n \to \infty$, $K_n / n \to 0$,

$$\hat{\sigma}_{i\Delta_n}^4 = \frac{n^2}{3K_n} \sum_{j=i-K_n}^{i} (\Delta_j^n X)^4, \qquad K_n + 1 \le i \le n. \tag{48}$$

For a sequence of i.i.d. standard normals $\{Z_i\}_{1\leq i\leq \lfloor nK_n^{-1}\rfloor}$, let

$$\hat{V}_n^{\dagger} = \sup_{0\leq t\leq 1} |\hat{S}_{nt} - t\hat{S}_n|, \qquad \hat{S}_{nt} = \left(\frac{K_n}{n}\right)^{1/2} \sum_{i=1}^{\lfloor nt/K_n\rfloor} \hat{\sigma}^4_{(iK_n+1)\Delta_n} Z_i. \tag{49}$$

Based on $\hat{V}_n^{\dagger}$, we construct the approximate (conditional) quantiles

$$\hat{q}_\alpha(\hat{V}_n^{\dagger}|\mathcal{F}) = \inf\{x\geq 0 : \mathbb{P}(\hat{V}_n^{\dagger}\leq x|\mathcal{F})\geq \alpha\}. \tag{50}$$

We can compute $\hat{q}_\alpha(\hat{V}_n^{\dagger}|\mathcal{F})$ as accurate as we want by using Monte Carlo approximations. Testing problem 5.2 is now addressed with the test

$$\psi_\alpha^{\dagger}((X_{i\Delta_n})_{0\leq i\leq n}) = \begin{cases} 1, & \text{if } \sqrt{\dfrac{3}{80}} V_n^{\dagger} > \hat{q}_{1-\alpha}(\hat{V}_n^{\dagger}|\mathcal{F}), \\ 0, & \text{otherwise.}\end{cases} \tag{51}$$

Observe that this test does not require any pre-specification of $\mathfrak{a}, \mathfrak{a}'$. It reacts to the change under the alternative for any $\mathfrak{a}, \mathfrak{a}'$. From a statistical perspective, it is important that:

(A) $\psi_\alpha^{\dagger}$ (asymptotically) correctly controls the type I error under H_0,
(B) $\psi_\alpha^{\dagger}$ provides optimal power (in the minimax sense).

Both properties are established in Theorem 5.5 and Theorem 5.7, respectively.

THEOREM 5.5. *Suppose that Assumption 5.1 holds. Then, for any fixed $\alpha > 0$,*

$$|\mathbb{P}(V_n^{\dagger} \leq \hat{q}_\alpha(\hat{V}_n^{\dagger}|\mathcal{F})) - \alpha| \to 0.$$

Before discussing property (B) of the test, let us first touch on a second approach that avoids a bootstrap. Slightly modified assumptions allow for the following standardized version of (44).

PROPOSITION 5.6. *Suppose that Assumption 5.1 is satisfied and that the hypothesis of Testing problem 5.2 holds with $\mathfrak{a} > 1/4$. Then (45) holds true. Moreover, $(3/80)^{1/2}\bar{V}_n^{\dagger}$ with*

$$\bar{V}_n^{\dagger} = n^{-1/2} \max_{K_n+1\leq m\leq n} \Bigg| \sum_{i=K_n+1}^{m} (\hat{\sigma}^4_{(i-2)\Delta_n})^{-1} Q_{n,i} - (m/n) \sum_{i=K_n+1}^{n} (\hat{\sigma}^4_{(i-2)\Delta_n})^{-1} Q_{n,i} \Bigg|$$

weakly converges to a Kolmogorov–Smirnov law and the associated limiting process in (45) becomes a standard Brownian bridge.

THEOREM 5.7. *For this testing problem, the minimax distinguishable boundary satisfies*

$$b_n \propto \left(n^{-1/2+2\mathfrak{a}'} + n^{-2(\mathfrak{a}-\mathfrak{a}')}\right). \tag{52}$$

In particular, for $b_n = \mathcal{O}(n^{-1/2+2\mathfrak{a}'} + n^{-2(\mathfrak{a}-\mathfrak{a}')})$ *a consistent test does not exist*:

$$\lim_{n\to\infty} \inf_{\psi} \gamma_{\psi}(\mathfrak{a}, b_n) = 1.$$

REMARK 5.8. The lower bound in Theorem 5.7 reveals that detecting an alternative with "too smooth" volatility, $\mathfrak{a}' > 1/4$, is not possible in a high-frequency setting.

The boundary b_n in Theorem 5.7 is of slightly different nature than the one in Theorem 4.1. Roughly speaking, the testing problem in Section 4.1 can be associated with a high-dimensional statistical experiment, whereas the one here is connected to a univariate statistical experiment. In this case, an optimal test $\psi = \psi_{\alpha}$ can only reach the lower bound up to a pre-specified nominal level $1-\alpha$, $0 < \alpha < 1$, see Ingster and Suslina (2003). Equivalently, we call a sequence of tests ψ_n minimax-optimal if for any b_n' with $n^{-1/2+2\mathfrak{a}'} + n^{-2(\mathfrak{a}-\mathfrak{a}')} = \mathcal{O}(b_n')$

$$\lim_{n\to\infty} \gamma_{\psi_n}\left(\mathfrak{a}, b_n'\right) = 0. \tag{53}$$

PROPOSITION 5.9. *The test* $\psi_{\alpha}^{\dagger}$ *in* (51) *is minimax-optimal. A corresponding test based on Proposition* 5.6 *is also minimax-optimal if* $\mathfrak{a} > 1/4$. *In particular, under the alternative of Testing problem* 5.2 *we have* $V_n^{\dagger} \xrightarrow{\mathbb{P}} \infty$, *when*

$$n^{-1/2+2\mathfrak{a}'} + n^{-2(\mathfrak{a}-\mathfrak{a}')} = \mathcal{O}\left(b_n'\right). \tag{54}$$

REMARK 5.10. Truncation in (44), analogous to Section 3.2, gives a method robust to jumps of (X_t). Also, the time of change θ can be estimated using the argmax. Precise results on the latter aspects are left for future research.

We have established a minimax-optimal test for global changes. The methods from Section 3 react under some conditions also to global changes, but do not attain optimality. Combining both approaches provides the statistician with suitable devices to analyze volatility dynamics.

6. Simulations. We examine the finite-sample properties of the proposed methods in a simulation study. Complementary simulations including a concise discussion of data applications are given in Bibinger, Jirak and Vetter (2016a). Here, we consider a moderate sample size $n = 500$ which is realistic when analyzing high-frequency data from only one single day.[6] In Bibinger, Jirak and Vetter

[6]Cf. our discussion in Bibinger, Jirak and Vetter (2016a).

(2016a), we additionally consider larger and even smaller sample sizes. We simulate observations at regular times of (1) with a stochastic semimartingale volatility model:

$$\sigma_t = \left(\int_0^t c \cdot \rho \, dW_s + \int_0^t \sqrt{1-\rho^2} \cdot c \, dW_s^{\perp} + 1\right) \cdot v_t \tag{55}$$

which fluctuates around a deterministic seasonality function

$$v_t = 1 - 0.2 \sin\left(\tfrac{3}{4}\pi t\right), \qquad t \in [0, 1], \tag{56}$$

with $c = 0.1$ and $\rho = 0.5$, where $W^{\perp}$ is a standard Brownian motion independent of W. The start value is $X_0 = 4$, and the drift is constant, $a = 0.1$. (56) mimics a realistic volatility shape with a strong decrease after opening and a slight increase before closing, and the model poses an intricate setup to discriminate jumps from continuous motion based on the $n = 500$ discrete recordings of X.

Under the local alternative, we add one jump of size 0.2 at the fixed time $t = 2/3$ to σ_t, which equals the range of the continuous movement in (56) and shifts the volatility back to its maximum start value. This is in line with effects evoked from macroeconomic news in the financial context. Changing the time of the volatility jump does not affect the results substantially, though. One jump of X at a uniformly drawn arrival time is implemented for the hypothesis and the alternative as well, and under the alternative X additionally exhibits a common jump of X and σ at $t = 2/3$. All these jumps are $N(0.5, 0.1)$ distributed.

Since X contains jumps, we apply the test statistic (24). We focus on V^*_{n,u_n} with overlapping blocks as it significantly outperforms the test with nonoverlapping blocks. For the truncation sequence, we set $u_n = \sqrt{2\log(n)}n^{-1/2} \approx 3.53 \; n^{-1/2}$; see Remark 3.7. In all cases, we iterate 10,000 Monte Carlo runs.

Since the approximation of the limit law is often imprecise for limit theorems with extreme value distributions in finite-sample applications, it is common practice to apply bootstrap procedures; see, for example, Hall (1991). We apply a wild bootstrap-type procedure based on Monte Carlo simulations. For example, one can use the statistics (23) to pre-estimate the (in practice) unknown volatility first. Afterwards, a linear filter with equal weights and k_n lags can be used to derive an estimated volatility shape. Then the statistics (15) are iteratively simulated, with X being a discretized Itô process without jumps and drift and with the pre-estimated volatility, to obtain critical values from the bootstrapped distribution. It is also possible to simply perform this procedure without pre-estimated volatility by just using discretized Brownian motion which is close to the procedure of Wu and Zhao (2007), or to use any convex combination of both statistics.

Let us briefly touch on the motivation and the consistency of both bootstrap methods. Looking at the proof of Theorem 3.2, one finds that both methods mimic specific stages of the proof, where it is shown that the underlying model can be subsequently reduced to (nested) simpler models. In particular, consistency of the proposed bootstrap methods is implied. In our simulations, both methods work reasonably well. The simpler method without pre-estimation tends to underestimate quantiles of the test statistic, especially in more complex configurations, whereas the version with pre-estimation slightly overestimates the quantiles. We fix a convex combination with weights 0.8 and 0.2 in favor of the version with pre-estimation across all our experiments.

In our simulations, we work with $k_{500} = 125$, but minor modifications of k_n do not change the results substantially. In general, k_n is chosen according to the optimal theoretical value

$$k_n = c_k \big(\log(n)\big)^{\frac{1}{2\mathfrak{a}+1}} n^{\frac{2\mathfrak{a}}{2\mathfrak{a}+1}} L_n^{-\frac{2}{2\mathfrak{a}+1}}, \tag{57}$$

with some proportionality constant c_k, and for $\mathfrak{a} = 1/2$ we obtain the best performance for $c_k \in [1.5, 3.5]$.[7] As illustrated in Figure 3, null and alternative are reasonably well distinguished, and even the fit by the asymptotic distribution attains a high accuracy here. In some other experiments given in Bibinger, Jirak and Vetter (2016a), the bootstrap does more clearly a better job than relying on the asymptotic results. Nevertheless, the latter are remarkably accurate across all settings as well. In light of the intricate setup, Figure 3 confirms a good finite-sample performance. In Bibinger, Jirak and Vetter (2016a), we further show that having $n = 1000$ observations or, as in Figure 1, a larger jump increases the finite-sample accuracy considerably. The density curve of the bootstrapped law in Figure 3 is obtained from a kernel density estimate with R's standard bandwidth selection using Silverman's rule of thumb.

In the financial literature, many stochastic volatility models rely on fractional nonsemimartingale processes, and in particular the interest in volatilities with small regularity has increased recently; see Gatheral, Jaisson and Rosenbaum (2014). To see how our methods perform in such models, we modify our setup using the prominent fractional log-volatility model by Comte and Renault (1998), that is, we replace the semimartingale above by a fractional OU-process

$$d\big(\log(\tilde{\sigma}_t)\big) = -0.1 \log(\tilde{\sigma}_t)\, dt + 0.1\, dB_t^H, \qquad \sigma_t = \tilde{\sigma}_t \cdot v_t, \tag{58}$$

with a fractional Brownian motion $(B_t^H)_{0 \le t \le 1}$ and a Hurst parameter H. The fractional process is implemented following Choleski's method with a code similar to the one in Appendix A.3 of Coeurjolly (2000).

The upper part of Figure 4 presents the finite-sample precision of the test for volatility jumps when using a small Hurst parameter $H = 0.2$ to model the volatility. The outcomes are almost as accurate as for the semimartingale volatility and broadly give a similar picture. Therefore, there is even finite-sample precision for detecting jumps in a fractional volatility process with a small Hurst parameter.

[7] We conjecture that the asymptotic sharp minimax constant lies in this interval. We further address robustness against different choices of k_n in Bibinger, Jirak and Vetter (2016a).

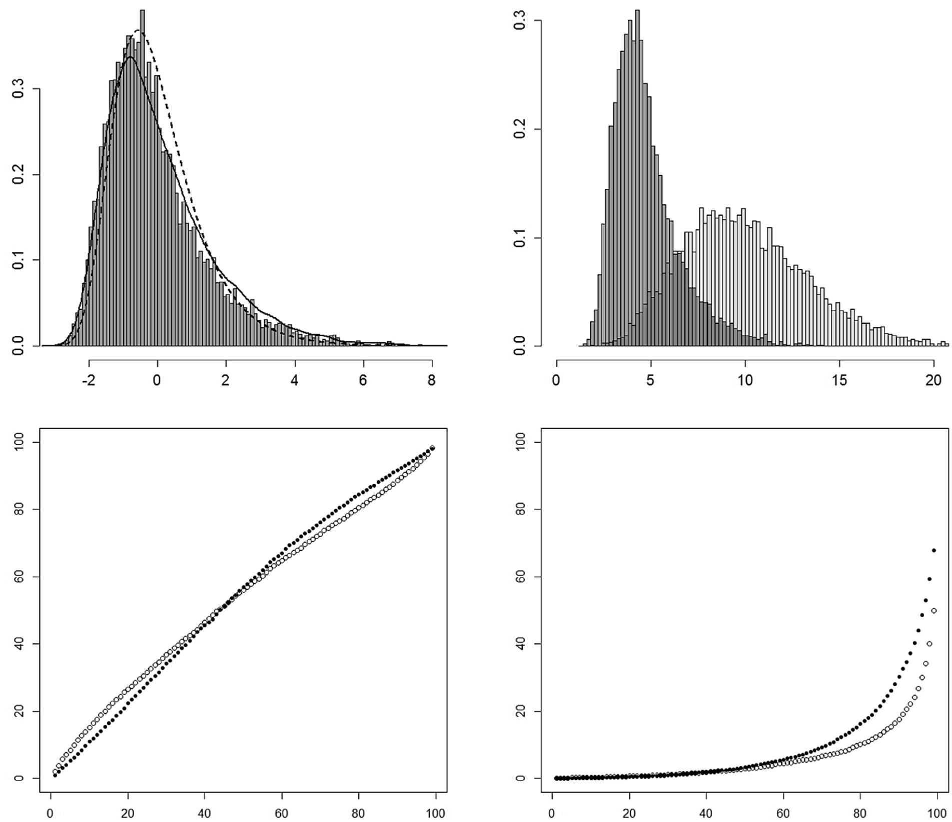

FIG. 3. *Top*: *Histograms of* (24) *for* $k_{500} = 125$ *under the hypothesis and the alternative* (*right*) *and a rescaled version comparing the left hand side with the limit law of* (27) *and the bootstrapped law* (*left*); *limit law density marked by dotted and bootstrapped by solid line*. *Bottom*: *Empirical size* (*left*) *and power* (*right*) *of the test by comparing the empirical percentiles to the ones of the limit law under* H_0 (*light points*) *and to the bootstrapped percentiles* (*dark points*).

Coming back to our introductory data example from Figure 2 for intra-day prices on March 18, 2009, the test rejects the null for both 3M and GE with p-values very close to zero. The point in time where the difference of adjacent statistics is maximized estimates the time of the structural change under the alternative. In both examples, we find grid point 285, corresponding to 02:15 p.m. EST, as the estimated change-point.

Finally, we examine the test for global changes based on (44) in a simulation. Consider hypothesis (55) for the volatility against the alternative that $(\sigma_t)_{t\geq\theta}$ follows (58) with a Hurst parameter $H = 0.15$, where under the alternative the change in smoothness happens at $\theta = 0.5$. The lower part of Figure 4 confirms a remarkable finite sample performance of this test. Small modifications of H under the alternative do not affect the results substantially. Therefore, we expect that the methodology in Section 5 opens up valuable new ways for inference on the reg-

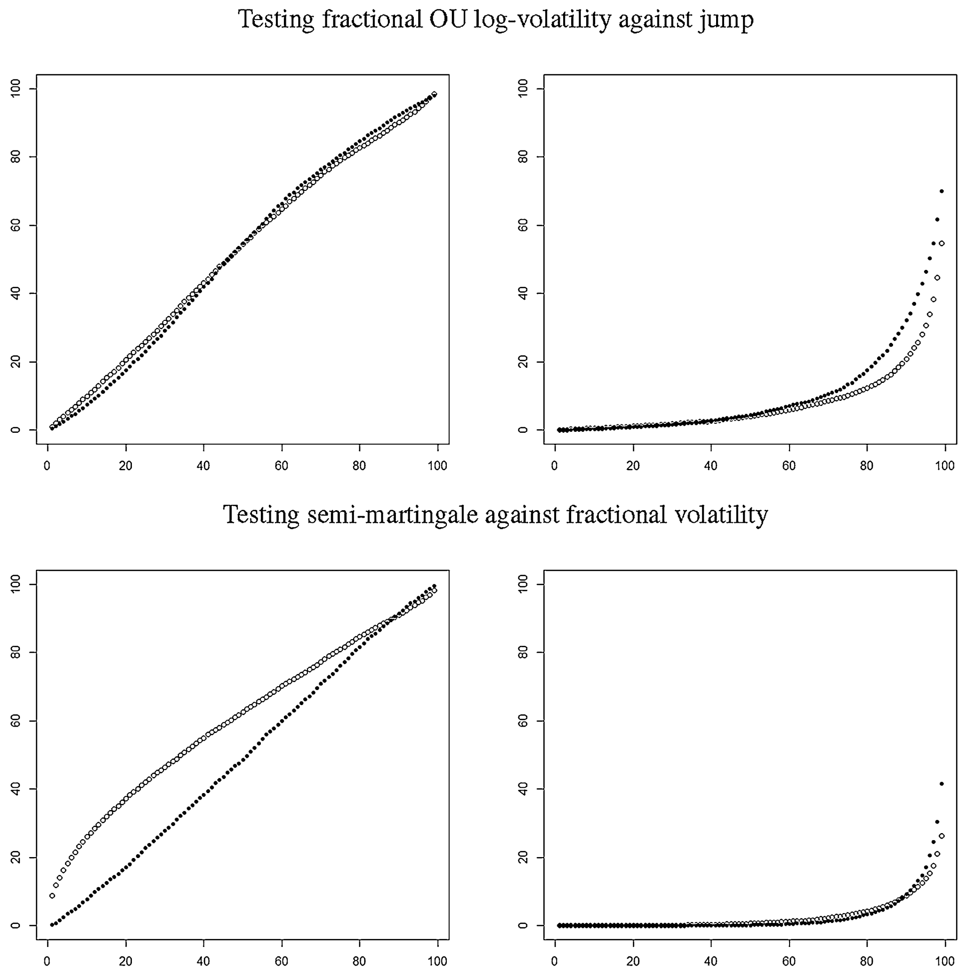

FIG. 4. *Empirical size* (*left*) *and power* (*right*) *of the tests by comparing the empirical percentiles to ones of the limit law under* H_0 (*light points*) *and to the bootstrapped ones* (*dark points*). *Top*: *Test* (24) *with* $k_{500} = 125$; *bottom*: *Test* (44).

ularity of volatlity, which is useful for studies as the one in Gatheral, Jaisson and Rosenbaum (2014).

APPENDIX A: PROOF OF THEOREM 3.2

First, we reduce the proof of Theorem 3.2 to Propositions A.1–A.5. The main part in the analysis of V_n from (13) is to replace it by the statistic

$$U_n = \max_{i=0,\dots,\lfloor n/k_n \rfloor - 2} |Y_{n,i}/Y_{n,i+1} - 1|, \tag{59}$$

in which the original statistics $RV_{n,i}$ from (14) are approximated by

$$Y_{n,i} = \frac{n}{k_n} \sum_{j=1}^{k_n} \sigma^2_{ik_n\Delta_n} \left(\Delta^n_{ik_n+j} W\right)^2. \tag{60}$$

Up to different (random) factors in front, the maximum in U_n is constructed from functionals of the i.i.d. increments of Brownian motion, which helps a lot in the derivation of its asymptotic behavior. We start with a result on the approximation error due to replacing V_n by U_n.

PROPOSITION A.1. *Suppose that we are under the null. If Assumption 3.1 and (16) hold, then we have*

$$\sqrt{\log(n) k_n}(V_n - U_n) \xrightarrow{\mathbb{P}} 0.$$

Recall that the variables $Y_{n,i}$ are not only computed over different intervals, but come with different volatilities in front as well. In order to obtain a statistic which is independent of σ let us define

$$\widetilde{Y}_{n,i} = \frac{n}{k_n} \sum_{j=1}^{k_n} \sigma^2_{(i-1)k_n\Delta_n} \left(\Delta^n_{ik_n+j} W\right)^2, \tag{61}$$

where the volatility factor is shifted in time now. Set then

$$\widetilde{U}_n = \max_{i=0,\ldots,\lfloor n/k_n \rfloor - 2} \left|Y_{n,i}/\widetilde{Y}_{n,i+1} - 1\right|. \tag{62}$$

PROPOSITION A.2. *Suppose that we are under the null. If Assumption 3.1 and (16) hold, then we have*

$$\sqrt{\log(n) k_n}(U_n - \widetilde{U}_n) \xrightarrow{\mathbb{P}} 0.$$

In the final step, we replace $\widetilde{Y}_{n,i+1}$ in the denominator by its limit $\sigma^2_{ik_n\Delta_n}$. Set

$$\widetilde{V}_n = \max_{i=0,\ldots,\lfloor n/k_n \rfloor - 2} \left|\frac{Y_{n,i} - \widetilde{Y}_{n,i+1}}{\sigma^2_{ik_n\Delta_n}}\right|. \tag{63}$$

PROPOSITION A.3. *Suppose that we are under the null. If condition (16) is satisfied, then we have*

$$\sqrt{\log(n) k_n}(\widetilde{U}_n - \widetilde{V}_n) \xrightarrow{\mathbb{P}} 0.$$

From Propositions A.1 to A.3, we have $\sqrt{\log(n)k_n}(V_n - \widetilde{V}_n) \xrightarrow{\mathbb{P}} 0$, while

$$\widetilde{V}_n = \max_{i=0,\dots,\lfloor n/k_n\rfloor-2}\Bigg|\frac{1}{k_n}\sum_{j=1}^{k_n}\big(\sqrt{n}\Delta^n_{ik_n+j}W\big)^2 - \frac{1}{k_n}\sum_{j=1}^{k_n}\big(\sqrt{n}\Delta^n_{(i+1)k_n+j}W\big)^2\Bigg|. \tag{64}$$

This statistic corresponds to the statistic D_n given in (13) of Wu and Zhao (2007); see as well Proposition A.5. Precisely, after subtracting the mean on both sides above, their $(X_k)_{1\le k\le n}$ correspond to $((\sqrt{n}\Delta^n_k W)^2) - 1)_{1\le k\le n}$, which forms an i.i.d. sequence of shifted χ^2_1-variables.

In the same fashion, we can prove that the asymptotics of V_n^* in (15) can be traced back to the statistics D_n^* in (12) of Wu and Zhao (2007); see Proposition A.5.

PROPOSITION A.4. *We have that* $\sqrt{\log(n)k_n}(V_n^* - \tilde{V}_n^*) \xrightarrow{\mathbb{P}} 0$, *with*

$$\widetilde{V}_n^* = \max_{i=k_n,\dots,n-k_n}\Bigg|\frac{1}{k_n}\sum_{j=i+1}^{i+k_n}\big((\sqrt{n}\Delta^n_j W)^2 - (\sqrt{n}\Delta^n_{j-k_n}W)^2\big)\Bigg|. \tag{65}$$

Theorem 1 of Wu and Zhao (2007) establishes limit theorems of the form (17) and (18) under more restrictive assertions on k_n than (16), as they consider the behavior for a class of weakly dependent random sequences $(X_k)_{k\ge1}$. The next proposition provides a more specific limit theorem tailored to the asymptotic analysis of the statistics (64) and (65). In particular, instead of using the strong approximation theory under weak dependence from Wu (2007), we rely on classical bounds for the approximation error in the invariance principle for i.i.d. variables with existing moments. This is applicable in a more general setup with much smaller block lengths k_n.

PROPOSITION A.5. *Consider a sequence* $(X_k)_{k\in\mathbb{N}}$ *of i.i.d. random variables with* $\operatorname{Var}[X_k] = \varsigma^2$ *and* $\mathbb{E}[|X_k|^p] < \infty$ *for some* $p \ge 4$. *If*

$$k_n^{-p/2} n = \mathcal{O}\big((\log(n))^{-p/2}\big), \tag{66}$$

then with $m_n = \lfloor n/k_n\rfloor$ *the statistic*

$$D_n^* = \frac{1}{k_n}\max_{k_n\le i\le n-k_n}\Bigg|\sum_{j=i+1}^{k_n+i} X_j - \sum_{j=i-k_n+1}^{i} X_j\Bigg|.$$

obeys the weak convergence

$$\sqrt{\log(m_n)}\big(k_n^{1/2}\varsigma^{-1}\big)D_n^* - 2\log(m_n) - \frac{1}{2}\log\log(m_n) - \log 3 \xrightarrow{w} V,$$

where V is distributed according to (19). *The statistic*

$$D_n = \max_{1 \le i \le \lfloor n/k_n \rfloor - 2} \left| \sum_{j=1}^{k_n} X_{ik_n+j} - X_{(i+1)k_n+j} \right|$$

using nonoverlapping blocks satisfies under the same assumptions

$$\sqrt{\log(m_n)}\big((k_n^{1/2}\varsigma^{-1})D_n - [4\log(m_n) - 2\log(\log(m_n))]^{1/2}\big) \xrightarrow{w} V.$$

As all moments of the χ_1^2 distribution exist and k_n is at least of polynomial growth in n, Proposition A.5 applied to (64) and (65) implies Theorem 3.2. We provide proofs of Propositions A.1–A.5 in the supplement Bibinger, Jirak and Vetter (2016b).

APPENDIX B: LOWER BOUND FOR THE LOCAL PROBLEM

PROOF OF THEOREM 4.1. The proof is based on equivalences of statistical experiments in the strong Le Cam sense. After information-theoretic reductions, we subsequently move to statistical experiments that allow a simpler treatment; see (69) below. Our final experiment $\mathcal{E}_4$ is a special high-dimensional signal detection problem, from which we will deduce the lower bound by classical arguments.

First, consider alternatives with a jump as in (4). Throughout this proof, we set

$$k_n = c_k\big(\sqrt{\log(m_n)}n^{\mathfrak{a}}/L_n\big)^{\frac{2}{2\mathfrak{a}+1}}, \tag{67}$$

with a constant $c_k > 0$. In the preliminary step, we first grant the experimenter additional knowledge. We restrict to a subclass of $\mathcal{S}_\theta^J(\mathfrak{a}, b_n, L_n)$, where we have one jump at time $\theta \in (0, 1)$ in the volatility, $|\sigma_\theta^2 - \sigma_{\theta-}^2| \ge b_n$. Then we assume that $\theta n k_n^{-1} \in \{1, 2, \dots, \lfloor n/k_n \rfloor - 1\}$, such that the jump time is in the set of observation grid points which are multiples of k_n. Furthermore, we can stick to $X_0 = 0$ and $a_s = 0$, $s \in [0, 1]$. From an information-theoretic view, obtaining this additional knowledge can only decrease the lower boundary on minimax distinguishability. Consequently, a lower bound derived for the submodel carries over to the less informative general situation.

To ease the exposition, we first set $\sigma_-^2 = 1$ and $L_n = 1$ and generalize the result at the end of this proof. Next, denote with $[a]_b = a \mod b$ and let

$$\sigma_{j\Delta_n}^2 = \begin{cases} 1 + \big(k_n - [j]_{k_n}\big)^{\mathfrak{a}} n^{-\mathfrak{a}}, & \theta n \le j < \theta n + k_n, \\ 1, & \text{else.} \end{cases} \tag{68}$$

The discretized squared volatility exhibits a jump (resp., change-point) of order b_n at θ and then decays on the window $[\theta, \theta + k_n\Delta_n]$ smoothly with regularity $\mathfrak{a}$ and is constant elsewhere. It suffices to consider the subclass $\Sigma_\theta \subset \mathcal{S}_\theta^J(\mathfrak{a}, b_n, L_n)$

of squared discretized volatility processes of the above form for which it remains unknown on which window the jump occurs.

Introduce a sequence r_n with $r_n \to \infty$ such that $r_n k_n^{-1} \to 0$ as $n \to \infty$. We specify the following stepwise approximation of $(\sigma^2_{j\Delta_n})_{0\le j\le n} \in \Sigma_\theta$:

$$\widetilde{\sigma}^2_{j\Delta_n} = \begin{cases} 1 + (k_n - ir_n)^{\mathfrak{a}} n^{-\mathfrak{a}}, & \theta n + (i-1)r_n \le j \le \theta n + ir_n, 1 \le i \le \frac{k_n}{r_n}, \\ 1, & \text{else.} \end{cases}$$

Denote the observations by $\eta_j = \sigma_{(j-1)\Delta_n}(W_{j\Delta_n} - W_{(j-1)\Delta_n})$ and $\widetilde{\eta}_j = \widetilde{\sigma}_{(j-1)\Delta_n}(W_{j\Delta_n} - W_{(j-1)\Delta_n})$, $j = 1, \dots, n$, respectively, with W the Wiener process in X_t. In the sequel, we distinguish the two cases where $\mathfrak{a} > 1/2$ and $\mathfrak{a} \le 1/2$.

Case $\mathfrak{a} > 1/2$: As alluded to above, we relate different experiments:

$\mathcal{E}_1$: Observe $(\eta_j)_{1\le j\le n}$ and information $\theta n k_n^{-1} \in \{1, 2, \dots, \lfloor n/k_n\rfloor - 1\}$ is provided.

$\mathcal{E}_2$: Observe $(\widetilde{\eta}_j)_{1\le j\le n}$ and information $\theta n k_n^{-1} \in \{1, 2, \dots, \lfloor n/k_n\rfloor - 1\}$ is provided.

$\mathcal{E}_3$: Observe $\boldsymbol{\chi} = ((\widetilde{\sigma}^2_{ik_n\Delta_n}\chi_i)_{i\in\mathcal{I}_1}, (\widetilde{\sigma}^2_{\theta+(i-1)r_n\Delta_n}\tilde{\chi}_i)_{i\in\mathcal{I}_2})$, where indices $(ik_n, i \in \mathcal{I}_1)$ expand over all multiples of k_n, except the one where the jump is located, that is, $\mathcal{I}_1 = \{1, \dots, \theta n k_n^{-1} - 1, \theta n k_n^{-1} + 1, \dots, \lfloor n/k_n\rfloor - 1\}$, and $(\theta n + (i-1)r_n, i \in \mathcal{I}_2)$ over all multiples of r_n in the window of length $k_n\Delta_n$ where (σ_j^2) is nonconstant, that is, $\mathcal{I}_2 = \{1, 2, \dots, k_n r_n^{-1}\}$. $(\chi_i)_{i\in\mathcal{I}_1}$ and $(\tilde{\chi}_i)_{i\in\mathcal{I}_2}$ are i.i.d. random variables having chi-square distribution with degrees of freedom k_n for $i \in \mathcal{I}_1$ and r_n for $i \in \mathcal{I}_2$. Moreover, information $\theta n k_n^{-1} \in \{1, 2, \dots, \lfloor n/k_n\rfloor - 1\}$ is provided.

$\mathcal{E}_4$: We observe $\boldsymbol{\xi} = ((k_n^{-1/2}\xi_i\widetilde{\sigma}^2_{ik_n\Delta_n} + \widetilde{\sigma}^2_{ik_n\Delta_n})_{i\in\mathcal{I}_1}, (r_n^{-1/2}\tilde{\xi}_i\widetilde{\sigma}^2_{\theta+(i-1)r_n\Delta_n} + \widetilde{\sigma}^2_{\theta+(i-1)r_n\Delta_n})_{i\in\mathcal{I}_2})$ where $(\xi_i, \tilde{\xi}_i)$ are i.i.d. standard normal random variables. Moreover, information $\theta n k_n^{-1} \in \{1, 2, \dots, \lfloor n/k_n\rfloor - 1\}$ is provided.

When considering the above experiments, we always have $(\sigma^2_{j\Delta_n}) \in \Sigma_\theta$ [or $(\widetilde{\sigma}^2_{j\Delta_n}) \in \Sigma_\theta$] as unknown parameter that index a family of probability measures $\{\mathbb{P}_{(\sigma^2_{j\Delta_n})}\}$. For the sake of readability, we move this formalism to the background and omit subscripts indicating the parameter space. We show the following relations for the experiments, where $\sim$ marks strong Le Cam equivalence and $\approx$ asymptotic equivalence:

$$\mathcal{E}_1 \approx \mathcal{E}_2 \sim \mathcal{E}_3 \approx \mathcal{E}_4. \tag{69}$$

Finally, we shall derive the lower bound in $\mathcal{E}_4$ which carries over to $\mathcal{E}_1$ by the above relations and thus also to our general model. The proof is now divided into four main steps.

Step 1 $\mathcal{E}_1 \approx \mathcal{E}_2$: For random variables U, V and their laws $\mathbb{P}_U, \mathbb{P}_V$, we denote the Kullback–Leibler divergence $\mathbf{D}(U\|V) = \mathbf{D}(\mathbb{P}_U\|\mathbb{P}_V) = \int \log(d\mathbb{P}_U/d\mathbb{P}_V)\, d\mathbb{P}_U$.

For normal families with unknown variance $\mathbb{P}_\theta = N(0,\theta)$, it is known that

$$\mathbf{D}(\mathbb{P}_\theta \| \mathbb{P}_{\theta'}) = \mathbb{E}_\theta\left[\log\left(\frac{d\mathbb{P}_\theta}{d\mathbb{P}_{\theta'}}\right)\right] = -\frac{1}{2}\left(\log\left(\frac{\theta}{\theta'}\right) + 1 - \frac{\theta}{\theta'}\right),$$

such that for $\theta = \theta' + \delta$ and considering asymptotics where $\delta \to 0$, we obtain

$$\mathbf{D}(\mathbb{P}_{\theta'+\delta} \| \mathbb{P}_{\theta'}) = -\frac{1}{2}\left(\log\left(1 + \frac{\delta}{\theta'}\right) - \frac{\delta}{\theta'}\right) = \frac{\delta^2}{4(\theta')^2} + \mathcal{O}(\delta^3). \tag{70}$$

As $\mathcal{E}_1$ and $\mathcal{E}_2$ share a common space on which the considered random variables are accommodated, asymptotic equivalence holds if $\|\mathbb{P}_{(\eta_j)} - \mathbb{P}_{(\tilde\eta_j)}\|_{TV} \to 0$ as $n \to \infty$ where $\|\cdot\|_{TV}$ denotes the total variation distance and $\mathbb{P}_{(\eta_j)}$ the law of observations (η_j). We exploit Pinsker's inequality

$$\|\mathbb{P}_{(\eta_j)} - \mathbb{P}_{(\tilde\eta_j)}\|_{TV}^2 \le \frac{1}{2}\mathbf{D}\big((\eta_j)\|(\tilde\eta_j)\big). \tag{71}$$

By Gaussianity and independence of Brownian increments, implying additivity of the Kullback-Leibler divergences, it follows with (70) for a piecewise constant approximation of a function with regularity $\mathfrak{a}$ on $k_n r_n^{-1}$ intervals of length $r_n\Delta_n$:

$$\mathbf{D}\big((\eta_j)\|(\tilde\eta_j)\big) = \mathcal{O}(1)\sum_{i=1}^{k_n r_n^{-1}}\sum_{j=1}^{r_n}(j\Delta_n)^{2\mathfrak{a}} = \mathcal{O}\big(n^{-2\mathfrak{a}}k_n r_n^{2\mathfrak{a}}\big),$$

which tends to zero for $r_n k_n^{-1} = \mathcal{O}(n^{-\epsilon})$ for some $\epsilon > 0$.

Step 2 $\mathcal{E}_2 \sim \mathcal{E}_3$: The vector of averages

$$\left(\left(k_n^{-1}\sum_{j=1}^{k_n}\tilde\eta^2_{ik_n+j-1}\right)_{i\in\mathcal{I}_1}, \left(r_n^{-1}\sum_{j=1}^{r_n}\tilde\eta^2_{\theta n+(i-1)r_n+j-1}\right)_{i\in\mathcal{I}_2}\right)$$

forms a sufficient statistic for $(\tilde\sigma^2_{\frac{j}{n}-1})_{1\le j\le n}$. Thereby, we conclude [see, e.g., Lemma 3.2 of Brown and Low (1996)] the strong Le Cam equivalence.

Step 3 $\mathcal{E}_3 \approx \mathcal{E}_4$: Let $\boldsymbol{\chi}^\diamond = (k_n^{-1/2}(\tilde\sigma^2_{ik_n\Delta_n}(\chi_i - k_n))_{i\in\mathcal{I}_1}, r_n^{-1/2}(\tilde\sigma^2_{\theta+(i-1)r_n\Delta_n}(\tilde\chi_i - r_n))_{i\in\mathcal{I}_2})$ and $\boldsymbol{\xi}^\diamond = ((\xi_i\tilde\sigma^2_{ik_n\Delta_n})_{i\in\mathcal{I}_1}, (\tilde\xi_i\tilde\sigma^2_{\theta+(i-1)r_n\Delta_n})_{i\in\mathcal{I}_2})$. In both experiments, random variables are accommodated on the same space. Rescaling and a location shift yield with Pinsker's inequality

$$\|\mathbb{P}_{\boldsymbol{\chi}} - \mathbb{P}_{\boldsymbol{\xi}}\|_{TV}^2 = \|\mathbb{P}_{\boldsymbol{\chi}^\diamond} - \mathbb{P}_{\boldsymbol{\xi}^\diamond}\|_{TV}^2 \le \frac{1}{2}\mathbf{D}\big(\boldsymbol{\chi}^\diamond\|\boldsymbol{\xi}^\diamond\big).$$

By independence, it follows that

$$\begin{aligned}\mathbf{D}\big(\boldsymbol{\chi}^\diamond\|\boldsymbol{\xi}^\diamond\big) \le{}& \sum_{i\in\mathcal{I}_1}\mathbf{D}\big(k_n^{-1/2}\tilde\sigma^2_{\frac{ik_n}{n}}(\chi_i - k_n)\|\xi_i\tilde\sigma^2_{\frac{ik_n}{n}}\big)\\ &+ \sum_{i\in\mathcal{I}_2}\mathbf{D}\big(r_n^{-1/2}\tilde\sigma^2_{\theta+\frac{(i-1)r_n}{n}}(\tilde\chi_i - r_n)\|\tilde\xi_i\tilde\sigma^2_{\theta+\frac{(i-1)r_n}{n}}\big).\end{aligned}$$

An application of Theorem 1.1 in Bobkov, Chistyakov and Götze (2013) yields

$$\sum_{i\in\mathcal{I}_1} \mathbf{D}\big(k_n^{-1/2}\tilde{\sigma}^2_{ik_n\Delta_n}(\chi_i - k_n)\|\xi_i\tilde{\sigma}^2_{ik_n\Delta_n}\big) = \mathcal{O}\big(nk_n^{-2}\big),$$

$$\sum_{i\in\mathcal{I}_2} \mathbf{D}\big(r_n^{-1/2}\tilde{\sigma}^2_{\theta+(i-1)r_n\Delta_n}(\tilde{\chi}_i - r_n)\|\tilde{\xi}_i\tilde{\sigma}^2_{\theta+(i-1)r_n\Delta_n}\big) = \mathcal{O}\big(k_n r_n^{-2}\big).$$

For $\mathfrak{a} > 1/2$, we have $nk_n^{-2} = \scriptstyle\mathcal{O}(1)$. Choosing r_n sufficiently large such that $k_n r_n^{-2} = \scriptstyle\mathcal{O}(1)$, it follows that

$$\|\mathbb{P}_{\boldsymbol{\chi}} - \mathbb{P}_{\boldsymbol{\xi}}\|_{TV} = \scriptstyle\mathcal{O}(1), \tag{72}$$

what ensures the claimed asymptotic equivalence.

Step 4: By the previous steps, it suffices to establish a lower bound for the distinguishability in experiment $\mathcal{E}_4$. Adding an additional drift, which gives clearly an equivalent experiment, we consider observations $\boldsymbol{\xi} = ((k_n^{-1/2}\xi_i\tilde{\sigma}^2_{ik_n\Delta_n} + \tilde{\sigma}^2_{ik_n\Delta_n} - 1)_{i\in\mathcal{I}_1}, (r_n^{-1/2}\tilde{\xi}_i\tilde{\sigma}^2_{\theta+(i-1)r_n\Delta_n} + \tilde{\sigma}^2_{\theta+(i-1)r_n\Delta_n} - 1)_{i\in\mathcal{I}_2})$. Then the testing problem can be interpreted as a high dimensional location signal detection problem in the sup-norm. More precisely, we test the hypothesis:

$$H_0 : \sup_j\big(\tilde{\sigma}_j^2 - 1\big) = 0 \quad \text{against the alternative} \quad H_1 : \sup_j\big(\tilde{\sigma}_j^2 - 1\big) \geq b_n, \tag{73}$$

and we are interested in the maximal value $b_n \to 0$ such that the hypothesis H_0 and H_1 are nondistinguishable in the minimax sense. Nondistinguishability in the minimax sense is formulated as

$$\lim_{n\to\infty} \inf_{\psi} \gamma_\psi(\mathfrak{a}, b_n) = 1, \tag{74}$$

and the detection boundary here is $b_n \propto (k_n\Delta_n)^{\mathfrak{a}} \propto n^{-\frac{\mathfrak{a}}{2\mathfrak{a}+1}}$. In order to show (74), we proceed in the fashion of Section 3.3.7 of Ingster and Suslina (2003). Let $\mathbb{P}_{\boldsymbol{\xi}}$ be the law of the observations. We consider the probability measures

$$\mathbb{P}_0 = \mathbb{P}_{\boldsymbol{\xi}} \times \mathbb{P}_{\theta_0} \quad \text{and} \quad \mathbb{P}_1 = \mathbb{P}_{\boldsymbol{\xi}} \times \mathbb{P}_{\theta_1},$$

where $\mathbb{P}_{\theta_0}$ means the hypothesis of the test applies (no jump) and $\mathbb{P}_{\theta_1}$ draws a jump-time θ with $\theta nk_n^{-1} \in \{1, \dots, \lfloor n/k_n\rfloor - 1\}$ uniformly from this set. Therefore, $\mathbb{P}_0$ represents the probability measure without signal, and $\mathbb{P}_1$ the measure where a signal is present. It then follows that

$$\begin{aligned}\inf_{\psi} \gamma_\psi(\mathfrak{a}, b_n) &\geq 1 - \frac{1}{2}\|\mathbb{P}_1 - \mathbb{P}_0\|_{TV} \\ &\geq 1 - \frac{1}{2}\big|\mathbb{E}_{P_0}\big[L^2_{0,1} - 1\big]\big|^{1/2},\end{aligned}$$

with $L_{0,1} = d\mathbb{P}_1/d\mathbb{P}_0$ the likelihood ratio of the measures $\mathbb{P}_1$ and $\mathbb{P}_0$. For the validity of (74), it thus suffices to establish

$$\mathbb{E}_{\mathbb{P}_0}\big[L_{0,1}^2\big] \to 1 \qquad \text{as } n \to \infty. \tag{75}$$

To this end, for given θ we denote with $u_i^{\theta n k_n^{-1}} = \widetilde{\sigma}^4_{\theta+(i-1)r_n\Delta_n}$ for $i \in \mathcal{I}_2$, $v_i^{\theta n k_n^{-1}} = (u_i^{1/2} - 1)r_n^{1/2}$. We first perform some preliminary computations. Denote with $\varphi_Y(x)$ the density function of a Gaussian random variable Y, not necessarily standard normal, and for $a, b \in \{1, \dots, \lfloor n/k_n \rfloor - 1\}$

$$I_{a,b}(x, y) := \prod_{i\in\mathcal{I}_2} \frac{\varphi_{\tilde{\xi}_i (u_i^a)^{1/2}+v_i^a}(x_i)}{\varphi_{\tilde{\xi}_i}(x_i)} \prod_{i\in\mathcal{I}_2} \frac{\varphi_{\tilde{\xi}_i (u_i^b)^{1/2}+v_i^b}(y_i)}{\varphi_{\tilde{\xi}_i}(y_i)}.$$

Then we have that $I_{a,b} := \int I_{a,b}(x, y) \prod_{i\in\mathcal{I}_2} \varphi_{\tilde{\xi}_i}(x_i)\, dx_i \prod_{i\in\mathcal{I}_2} \varphi_{\tilde{\xi}_i}(y_i)\, dy_i = 1$. Next, for $a \in \{1, \dots, \lfloor n/k_n \rfloor - 1\}$, consider

$$II_a(x) := \prod_{i\in\mathcal{I}_2} \left(\frac{\varphi_{\tilde{\xi}_i (u_i^a)^{1/2}+v_i^a}(x_i)}{\varphi_{\tilde{\xi}_i}(x_i)}\right)^2.$$

Observe that for a standard Gaussian random variable Z and $s, t \in \mathbb{R}$, $|s| < 1/2$:

$$\mathbb{E}\big[\exp\big(sZ^2 + tZ\big)\big] = (1 - 2s)^{-1/2} \exp\left(\frac{t^2}{2 - 4s}\right). \tag{76}$$

This, together with the inequality

$$C_0 k_n (k_n\Delta_n)^{2\mathfrak{a}} \le r_n \sum_{i=0}^{k_n/r_n - 1} \big((k_n - i r_n)\Delta_n\big)^{2\mathfrak{a}} \le k_n (k_n\Delta_n)^{2\mathfrak{a}}$$

for some constant $C_0 > 0$ and routine calculations, yields for some $C_0 \le C_1 \le 1$

$$II_a := \int II_a(x) \prod_{i\in\mathcal{I}_2} \varphi_{\tilde{\xi}_i}(x_i)\, dx_i \le e^{C_1 k_n (k_n\Delta_n)^{2\mathfrak{a}}}\big(1 + \mathcal{O}(1)\big).$$

With all the preliminary calculations, we are now ready to derive a bound for

$$\begin{aligned}\mathbb{E}_{\mathbb{P}_0}\big[L_{0,1}^2\big] - 1 &= \sum_{a,b=1, a\neq b}^{\lfloor n/k_n \rfloor - 1} \mathbb{P}\big(\theta n k_n^{-1} = a\big)\mathbb{P}\big(\theta n k_n^{-1} = b\big)(I_{a,b} - 1) \\ &\quad + \sum_{a=1}^{\lfloor n/k_n \rfloor - 1} \mathbb{P}\big(\theta n k_n^{-1} = a\big)^2 (II_a - 1),\end{aligned}$$

where the first sum vanishes. For an appropriate choice of $c_k > 0$ in (67), we have that $k_n(k_n\Delta_n)^{2\mathfrak{a}} = C_2 \log(n/k_n)$ for some $C_2 < C_1^{-1}$. Since $\mathbb{P}(\theta n k_n^{-1} = a) =$

$k_n\Delta_n$, we thus obtain

$$
\begin{aligned}
\left|\mathbb{E}_{P_0}[L_{0,1}^2]-1\right| &\le \sum_{a=1}^{\lfloor n/k_n\rfloor} \mathbb{P}(\theta n k_n^{-1}=a)^2\left(e^{C_1 k_n (k_n\Delta_n)^{2\mathfrak{a}}}-1\right) \\
&= \left(1+\mathcal{O}(1)\right)k_n\Delta_n e^{C_1 k_n (k_n\Delta_n)^{2\mathfrak{a}}}.
\end{aligned}
\tag{77}
$$

We conclude (75) using

$$
\begin{aligned}
k_n\Delta_n e^{C_1 k_n (k_n\Delta_n)^{2\mathfrak{a}}} &= k_n\Delta_n \exp\left(C_1 C_2 \log(n/k_n)\right) \\
&= (k_n\Delta_n)^{1-C_1C_2} = \mathcal{O}(1).
\end{aligned}
$$

Case $\mathfrak{a}\le 1/2$: The only time we make use of the condition $\mathfrak{a}>1/2$ above is in Step 3 to obtain $n/k_n^2=\mathcal{O}(1)$. The necessity of this relation is due to the large number of blocks n/k_n, when operating with the entropy bounds. To establish the lower bound, this constraint can be removed by granting the experimenter even more additional information what is briefly sketched in the following. Indeed, suppose we know in addition that $\theta n\in\{k_n,2k_n,\dots,l_nk_n\}$ where $l_n=n^{\mathfrak{l}}\ll n/k_n$, $\mathfrak{l}>0$ arbitrarily small but strictly positive and such that $l_n\in\mathbb{N}$. Using the sufficiency argument of Step 2, we can gather all the information contained in $(\eta_i)_{l_nk_n<i\le n}$ in one single average $(n-(l_n+1)k_n)^{-1}\sum_{i=l_nk_n+1}^n \eta_i^2$. Then one can repeat Steps 3 and 4, subject to the weaker condition $l_n/k_n=\mathcal{O}(1)$. Selecting $\mathfrak{l}>0$ sufficiently small for each $0<\mathfrak{a}\le 1$, this is always possible. Substituting nk_n^{-1} by l_n in the sum and (squared) probability in Step 4, we obtain instead of (77)

$$
\left|\mathbb{E}_{P_0}[L_{0,1}^2]-1\right| = \left(1+\mathcal{O}(1)\right)l_n^{-1}e^{C_1 k_n (k_n\Delta_n)^{2\mathfrak{a}}}.
$$

For an appropriate choice of $c_k>0$ in (67), $k_n(k_n\Delta_n)^{2\mathfrak{a}}=C_2\log(n/k_n)$ with $C_2<\mathfrak{l}C_1^{-1}$. Hence, we conclude that the term tends to zero and the lower bound in Step 4 gives the same minimax detection boundary.

Let us now touch on the general case with some $\underline{\sigma}^2>0$ and sequences L_n. We can divide formulas in (73) by $\underline{\sigma}^2$ to rescale. Exactly the same arguments lead to $\lim_{n\to\infty}\inf_\psi \gamma_\psi(\mathfrak{a},b_n)=1$ for k_n given in (67) with $b_n\le L_n(k_n\Delta_n)^{\mathfrak{a}}\underline{\sigma}^2$. Finally, for regularity alternative H_1^{R}, the proof is along the same lines as for jumps where instead of a jump of size $L_n(k_n\Delta_n)^{\mathfrak{a}}$ we observe a sudden, more regular increase in σ_t^2 of size $L_n(k_n\Delta_n)^{\mathfrak{a}+\mathfrak{a}'}$. Then the arguments are almost identical. □

Acknowledgment. We thank two anonymous referees for valuable comments on earlier versions as well as Charlie Vollmer and Ou Zhao for proofreading the manuscript. We also thank Marc Hoffmann for helpful remarks on testing hypotheses of smoothness classes.

SUPPLEMENTARY MATERIAL

Complete proofs (DOI: 10.1214/16-AOS1499SUPPA; .pdf). We provide all remaining proofs for the results from Sections 3, 4 and 5.

Application and simulations (DOI: 10.1214/16-AOS1499SUPPB; .pdf). We present complementary simulations for different sample sizes accompanied by a sensitivity analysis of the dependence on k_n and a discussion of data applications.

REFERENCES

AÏT-SAHALIA, Y. and JACOD, J. (2009). Testing for jumps in a discretely observed process. *Ann. Statist.* **37** 184–222. MR2488349

ALVAREZ, A., PANLOUP, F., PONTIER, M. and SAVY, N. (2012). Estimation of the instantaneous volatility. *Stat. Inference Stoch. Process.* **15** 27–59. MR2892587

ANDERSEN, T. G. and BOLLERSLEV, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *Internat. Econom. Rev.* **39** 885–905.

ANDREWS, D. W. K. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica* **61** 821–856. MR1231678

ARCONES, M. A. (1994). Limit theorems for nonlinear functionals of a stationary Gaussian sequence of vectors. *Ann. Probab.* **22** 2242–2274. MR1331224

AUE, A., HÖRMANN, S., HORVÁTH, L. and REIMHERR, M. (2009). Break detection in the covariance structure of multivariate time series models. *Ann. Statist.* **37** 4046–4087. MR2572452

BAI, J. and PERRON, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica* **66** 47–78. MR1616121

BARNDORFF-NIELSEN, O. E. and SHEPHARD, N. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **64** 253–280. MR1904704

BIBINGER, M., JIRAK, M. and VETTER, M. (2016a). Supplement to "Nonparametric change-point analysis of volatility." DOI:10.1214/16-AOS1499SUPPA.

BIBINGER, M., JIRAK, M. and VETTER, M. (2016b). Supplement to "Nonparametric change-point analysis of volatility." DOI:10.1214/16-AOS1499SUPPB.

BOBKOV, S. G., CHISTYAKOV, G. P. and GÖTZE, F. (2013). Rate of convergence and Edgeworth-type expansion in the entropic central limit theorem. *Ann. Probab.* **41** 2479–2512. MR3112923

BROWN, L. D. and LOW, M. G. (1996). Asymptotic equivalence of nonparametric regression and white noise. *Ann. Statist.* **24** 2384–2398. MR1425958

COEURJOLLY, J.-F. (2000). Simulation and identification of the fractional Brownian motion: A bibliographical and comparative study. *J. Stat. Softw.* **50** 1–53.

COMTE, F. and RENAULT, E. (1998). Long memory in continuous-time stochastic volatility models. *Math. Finance* **8** 291–323. MR1645101

CSÖRGŐ, M. and HORVÁTH, L. (1997). *Limit Theorems in Change-Point Analysis*. Wiley, Chichester. MR2743035

EMBRECHTS, P. and MAEJIMA, M. (2000). An introduction to the theory of self-similar stochastic processes. In *Proceedings of the Summer School on Mathematical Physics* 1999: *Methods of Renormalization Group* (*Tokyo*) **14** 1399–1420. Internat. J. Modern Phys. B, 12-13. MR1791158

GATHERAL, J., JAISSON, T. and ROSENBAUM, M. (2014). Volatility is rough. Preprint. Available at arXiv:1410.3394.

HALL, P. (1991). On convergence rates of suprema. *Probab. Theory Related Fields* **89** 447–455. MR1118558

HINKLEY, D. V. (1971). Inference about the change-point from cumulative sum tests. *Biometrika* **58** 509–523. MR0312623

HOFFMANN, M. and NICKL, R. (2011). On adaptive inference and confidence bands. *Ann. Statist.* **39** 2383–2409. MR2906872

IACUS, S. M. and YOSHIDA, N. (2012). Estimation for the change point of volatility in a stochastic differential equation. *Stochastic Process. Appl.* **122** 1068–1092. MR2891447

INGSTER, Y. I. (1993). Asymptotically minimax hypothesis testing for nonparametric alternatives. I. *Math. Methods Statist.* **2** 85–114. MR1257978

INGSTER, Y. I. and SUSLINA, I. A. (2003). *Nonparametric Goodness-of-Fit Testing Under Gaussian Models. Lecture Notes in Statistics* **169**. Springer, New York. MR1991446

JACOD, J. (1997). On continuous conditional Gaussian martingales and stable convergence in law. In *Séminaire de Probabilités, XXXI. Lecture Notes in Math.* **1655** 232–246. Springer, Berlin. MR1478732

JACOD, J. (2008). Asymptotic properties of realized power variations and related functionals of semimartingales. *Stochastic Process. Appl.* **118** 517–559. MR2394762

JACOD, J. and ROSENBAUM, M. (2013). Quarticity and other functionals of volatility: Efficient estimation. *Ann. Statist.* **41** 1462–1484. MR3113818

JACOD, J. and TODOROV, V. (2010). Do price and volatility jump together? *Ann. Appl. Probab.* **20** 1425–1469. MR2676944

MANCINI, C. (2009). Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scand. J. Stat.* **36** 270–296. MR2528985

MARSAGLIA, G., TSANG, WAN, W. and WANG, J. (2003). Evaluating Kolmogorov's distribution. *J. Stat. Softw.* **8** 1–4.

MÓRICZ, F. A., SERFLING, R. J. and STOUT, W. F. (1982). Moment and probability bounds with quasisuperadditive structure for the maximum partial sum. *Ann. Probab.* **10** 1032–1040. MR0672303

MÜLLER, H.-G. (1992). Change-points in nonparametric regression analysis. *Ann. Statist.* **20** 737–761. MR1165590

MÜLLER, H.-G. and STADTMÜLLER, U. (1999). Discontinuous versus smooth regression. *Ann. Statist.* **27** 299–337. MR1701113

MYKLAND, P. A. (2012). A Gaussian calculus for inference from high frequency data. *Ann. Finance* **8** 235–258. MR2922797

MYKLAND, P. A. and ZHANG, L. (2009). Inference for continuous semimartingales observed at high frequency. *Econometrica* **77** 1403–1445. MR2561071

PAGE, E. S. (1955). A test for a change in a parameter occurring at an unknown point. *Biometrika* **42** 523–527. MR0072412

PETTITT, A. N. (1980). A simple cumulative sum type statistic for the change-point problem with zero-one observations. *Biometrika* **67** 79–84. MR0570508

PHILLIPS, P. C. B. (1987). Time series regression with a unit root. *Econometrica* **55** 277–301. MR0882096

SPOKOINY, V. G. (1998). Estimation of a function with discontinuities via local polynomial fit with an adaptive window choice. *Ann. Statist.* **26** 1356–1378. MR1647669

SPOKOINY, V. (2009). Multiscale local change point detection with applications to value-at-risk. *Ann. Statist.* **37** 1405–1436. MR2509078

TODOROV, V. and TAUCHEN, G. (2011). Volatility jumps. *J. Bus. Econom. Statist.* **29** 356–371. MR2848503

WU, W. B. (2007). Strong invariance principles for dependent random variables. *Ann. Probab.* **35** 2294–2320. MR2353389

WU, W. B. and ZHAO, Z. (2007). Inference of trends in time series. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **69** 391–410. MR2323759

MARKUS BIBINGER
FACULTY OF MATHEMATICS AND COMPUTER SCIENCE
PHILIPPS-UNIVERSITÄT MARBURG
HANS-MEERWEIN-STRASSE 6
35032 MARBURG
GERMANY
E-MAIL: bibinger@mathematik.uni-marburg.de

MORITZ JIRAK
INSTITUT FÜR MATHEMATISCHE STOCHASTIK
TECHNISCHE UNIVERSITÄT BRAUNSCHWEIG
POCKELSSTRASSE 14
38106 BRAUNSCHWEIG
GERMANY
E-MAIL: m.jirak@tu-braunschweig.de

MATHIAS VETTER
MATHEMATISCHES SEMINAR
CHRISTIAN-ALBRECHTS-UNIVERSITÄT ZU KIEL
LUDEWIG-MEYN-STRASSE 4
24118 KIEL
GERMANY
E-MAIL: vetter@math.uni-kiel.de