Electronic Journal of Statistics
Vol. 13 (2019) 3654–3709
ISSN: 1935-7524
https://doi.org/10.1214/19-EJS1610

On detecting changes in the jumps of arbitrary size of a time-continuous stochastic process*

Michael Hoffmann

Ruhr-Universität Bochum, Fakultät für Mathematik, 44780 Bochum, Germany
e-mail: michael.hoffmann@rub.de

Holger Dette

Ruhr-Universität Bochum, Fakultät für Mathematik, 44780 Bochum, Germany
e-mail: holger.dette@rub.de

Abstract: This paper introduces test and estimation procedures for abrupt and gradual changes in the entire jump behaviour of a discretely observed Itō semimartingale. In contrast to existing work we analyse jumps of arbitrary size which are not restricted to a minimum height. Our methods are based on weak convergence of a truncated sequential empirical distribution function of the jump characteristic of the underlying Itō semimartingale. Critical values for the new tests are obtained by a multiplier bootstrap approach and we investigate the performance of the tests also under local alternatives. An extensive simulation study shows the finite-sample properties of the new procedures.

MSC 2010 subject classifications: 60F17, 60G51, 62G10, 62M99.
Keywords and phrases: Lévy measure, jump compensator, transition kernel, empirical processes, weak convergence, multiplier bootstrap, change points, gradual changes.

Received February 2018.

1. Introduction

Stochastic processes are widely used in science nowadays, as they allow for a flexible modelling of time-dependent phenomena. For example, in physics stochastic processes are used to explain the behaviour of quantum systems (see van Kampen, 2007), but stochastic processes are also suitable for financial modelling. The seminal paper by Delbaen and Schachermayer (1994) suggests to use the special class of Itō semimartingales in continuous time. Financial models based on Itō semimartingales satisfy a certain condition on the absence of arbitrage and moreover they are still rich enough to accommodate stylized facts such as volatility clustering, leverage effects and jumps. As a consequence, in recent

*This work has been supported by the Collaborative Research Center "Statistical modeling of nonlinear dynamic processes" (SFB 823, Project A1) and the Research Training Group "High-dimensional phenomena in probability – fluctuations and discontinuity" (RTG 2131) of the German Research Foundation (DFG).

years a lot of research was focused on the development of statistical procedures for characteristics of Itō semimartingales based on discrete observations. In particular, the importance of the jump component has been enforced by recent research (see Aït-Sahalia and Jacod, 2009a and Aït-Sahalia and Jacod, 2009b) and common methods in this field are gathered in the recent monographs by Jacod and Protter (2012) and Aït-Sahalia and Jacod (2014).

A fundamental topic in statistics for stochastic processes is the analysis of structural breaks. Corresponding test procedures, commonly referred to as change point tests, have their origin in quality control (see Page, 1954; Page, 1955) and nowadays, these techniques are widely used in many fields of science such as economics (Perron, 2006), finance (Andreou and Ghysels, 2009), climatology (Reeves et al., 2007) and engineering (Stoumbos et al., 2000). The contributions of the present paper to this field of research are new statistical procedures for the detection of changes in the jump behaviour of an Itō semimartingale. In contrast to the existing works Bücher et al. (2017) and Hoffmann et al. (2017) this paper introduces methods of inference on the jump behaviour of the underlying process in general, while in the previously mentioned references the authors restrict the analysis to jumps which exceed a minimum size $\varepsilon > 0$.

Throughout this work we assume that we have high-frequency data $X_{i\Delta_n}$ ($i = 0, 1, \ldots, n$) with $\Delta_n \to 0$, where the process $(X_t)_{t \in \mathbb{R}_+}$ is an Itō semimartingale with the following decomposition

$$X_t = X_0 + \int_0^t b_s\, ds + \int_0^t \sigma_s\, dW_s + \int_0^t \int_{\mathbb{R}} u 1_{\{|u| \le 1\}} (\mu - \bar{\mu})(ds, du) + \int_0^t \int_{\mathbb{R}} u 1_{\{|u|>1\}} \mu(du, dz).$$

Here W is a standard Brownian motion and μ is a Poisson random measure on $\mathbb{R}^+ \times \mathbb{R}$ with predictable compensator $\bar{\mu}$ satisfying $\bar{\mu}(ds, du) = ds\, \nu_s(du)$. Our approach is completely non-parametric, that is we only impose structural assumptions on the characteristic triplet (b_s, σ_s, ν_s) of $(X_t)_{t \in \mathbb{R}_+}$. The crucial quantity here is the transition kernel ν_s which controls the number and the size of the jumps around time $s \in \mathbb{R}_+$. Our aim is to test the null hypothesis

$$\mathbf{H}_0 : \nu_s(dz) = \nu_0(dz)$$

against various alternatives involving the non-constancy of ν_s. In particular, the detection of abrupt changes in a stochastic feature has been discussed extensively in the literature (see Aue and Horváth, 2013 and Jandhyala et al., 2013 for an overview in a time series context). The first part of this paper belongs to this area of research and introduces tests for $\mathbf{H}_0$ versus alternatives of an abrupt change of the form

$$\mathbf{H}_1^{(ab)} : \nu_s^{(n)}(dz) = \mathbb{1}_{\{s < \lfloor n\theta_0 \rfloor \Delta_n\}} \nu_1(dz) + \mathbb{1}_{\{s \ge \lfloor n\theta_0 \rfloor \Delta_n\}} \nu_2(dz),$$

for some unknown $\theta_0 \in (0, 1)$ and two distinct Lévy measures $\nu_1 \neq \nu_2$. Similar to the classical setup of detecting changes in the mean of a time series it is only possible to define the change point relative to the length of the data set which in our case is the time horizon $n\Delta_n$. However, for inference on the jump behaviour the time horizon has to tend to infinity ($n\Delta_n \to \infty$) since there are only finitely many jumps of a certain size on every compact interval. Furthermore, we also discuss how to estimate the unknown change point θ_0, if the alternative $\mathbf{H}_1^{(ab)}$ is true.

A more difficult problem is the detection of gradual (smooth, continuous) changes in a stochastic feature. As a consequence, the setup in most papers on this topic is restricted to non-parametric location or parametric models with independently distributed observations (see e.g. Bissell, 1984, Gan, 1991, Siegmund and Zhang, 1994, Hušková, 1999, Hušková and Steinebach, 2002 and Mallik et al., 2013). Gradual changes in a time series context are for instance discussed in Aue and Steinebach (2002) and Vogt and Dette (2015). In the second part of this paper we contribute to this development by introducing new procedures for gradual changes in the kernel ν_s, where we basically test $\mathbf{H}_0$ against the general alternative

$$\mathbf{H}_1^{(gra)} : \nu_s(dz) \text{ is not Lebesgue-almost everywhere constant in } s \in [0, n\Delta_n].$$

Moreover, we introduce an estimator for the first time point where the jump behaviour deviates from the null hypothesis.

The remaining paper is organized as follows: In Section 2 we give the basic assumptions on the characteristics of the underlying process and the observation scheme. Section 3 introduces test and estimation procedures for abrupt changes in the jump behaviour in general by using CUSUM processes. In Section 4 we discuss how to detect and estimate gradual changes in the entire jump behaviour. Section 5 contains an extensive simulation study investigating the finite-sample performance of the new procedures. Finally, all proofs are relegated to Section 6 and the technical appendices A–E.

2. The basic assumptions

In order to accommodate both abrupt and gradual changes in our approach we follow Hoffmann et al. (2017) and assume that there is a driving law behind the evolution of the jump behaviour in time which is common for all $n \in \mathbb{N}$. That is we assume that at step $n \in \mathbb{N}$ we observe an Itō semimartingale $X^{(n)}$ with characteristics $(b_s^{(n)}, \sigma_s^{(n)}, \nu_s^{(n)})$ at the equidistant time points $i\Delta_n$ with $i = 0, 1, \ldots, n$ which satisfies the following rescaling assumption

$$\nu_s^{(n)}(dz) = g\Big(\frac{s}{n\Delta_n}, dz\Big) \tag{2.1}$$

for a transition kernel $g(y, dz)$ from $([0,1], \mathbb{B}([0,1]))$ into $(\mathbb{R}, \mathbb{B})$, where here and below $\mathbb{B}(A)$ denotes the trace σ-algebra on $A \subset \mathbb{R}$ of the Borel σ-algebra $\mathbb{B}$ of $\mathbb{R}$. In order to detect changes in the jump behaviour of the underlying Itō semimartingale in general, we have to draw inference on the kernel $g(y, B)$ for sets $B \in \mathbb{B}$ containing the origin. However, g has locally the properties of a Lévy measure. Thus, if we deviate from the (simple) case of finite activity jumps the total mass of g on every neighbourhood of the origin is infinite and we cannot estimate $g(y,\cdot)$ on sets containing 0 directly. We address this problem by weighting the kernel g according to an auxiliary function, precisely for change point detection we consider

$$N_\rho(g;\theta,t) := \int_0^\theta \int_{-\infty}^t \rho(z) g(y,dz)dy, \tag{2.2}$$

for $(\theta,t) \in [0,1]\times\mathbb{R}$, where ρ is chosen appropriately such that the integral is always defined. Under weak conditions on ρ, this so-called Lévy distribution function N_ρ determines the entire kernel g and therefore the evolution of the jump behaviour in time. The natural approach to draw inference on N_ρ is the following sequential generalization of an estimator in Nickl et al. (2016)

$$\tilde{N}_\rho^{(n)}(\theta,t) = \frac{1}{n\Delta_n}\sum_{i=1}^{\lfloor n\theta\rfloor} \rho(\Delta_i^n X^{(n)}) \mathbb{1}_{(-\infty,t]}(\Delta_i^n X^{(n)}),$$

for $(\theta,t) \in [0,1]\times\mathbb{R}$, where $\Delta_i^n X^{(n)} = X_{i\Delta_n}^{(n)} - X_{(i-1)\Delta_n}^{(n)}$. Using a spectral approach similar to Nickl and Reiß (2012) these authors prove weak convergence of $\sqrt{n\Delta_n}\big(\tilde{N}_\rho^{(n)}(1,t) - N_\rho(g;1,t)\big)$ in $\ell^\infty(\mathbb{R})$ to a tight Gaussian process, but only for Lévy processes without a diffusion component, i.e. in particular for constant $g(y,\cdot) \equiv \nu(\cdot)$. The main difficulty in generalizing this result is the superposition of small jumps with the roughly fluctuating Brownian component of the process. We solve this problem by using a truncation approach which has originally been used by Mancini (2009) to cut off jumps in order to draw inference on integrated volatility. More precisely, we follow Hoffmann and Vetter (2017) and identify jumps by inverting the truncation technique of Mancini (2009), i.e. all test statistics and estimators investigated below are functionals of the sequential truncated empirical Lévy distribution function

$$N_\rho^{(n)}(\theta,t) = \frac{1}{n\Delta_n}\sum_{i=1}^{\lfloor n\theta\rfloor} \rho(\Delta_i^n X^{(n)}) \mathbb{1}_{(-\infty,t]}(\Delta_i^n X^{(n)}) \mathbb{1}_{\{|\Delta_i^n X^{(n)}|>v_n\}}, \tag{2.3}$$

$((\theta,t) \in [0,1]\times\mathbb{R})$ for some suitable null sequence $v_n \to 0$.

As a further improvement to previous studies we analyse the asymptotic behaviour of our tests under local alternatives. That is, in the rescaling assumption (2.1) we let $g = g^{(n)}$ depend on $n \in \mathbb{N}$, where there exist transition kernels g_0, g_1, g_2 satisfying some additional regularity assumptions such that for each $y \in [0,1]$

$$g^{(n)}(y,dz) = g_0(y,dz) + \frac{1}{\sqrt{n\Delta_n}} g_1(y,dz) + \mathcal{R}_n(y,dz) \tag{2.4}$$

and for each $y \in [0,1]$, $B \in \mathbb{B}$ and $n \in \mathbb{N}$ the remainder kernel $\mathcal{R}_n$ satisfies

$$\mathcal{R}_n(y,B) \le a_n g_2(y,B)$$

for a sequence $a_n = o((n\Delta_n)^{-1/2})$ of non-negative real numbers. For constant $g_0(y,\cdot) \equiv \nu_0(\cdot)$ assumption (2.4) is exactly the local alternative where the jump behaviour converges to the null hypothesis $g_0(y,\cdot) \equiv \nu_0(\cdot)$ from the direction defined by g_1 at rate $(n\Delta_n)^{-1/2}$. In this sense, Theorem 6.3, in which we prove weak convergence of the stochastic process

$$G_\rho^{(n)}(\theta,t) = \sqrt{n\Delta_n}\big(N_\rho^{(n)}(\theta,t) - N_\rho(g^{(n)};\theta,t)\big), \quad (\theta,t) \in [0,1]\times\mathbb{R}$$

to a tight Gaussian process in $\ell^\infty([0,1]\times\mathbb{R})$, is a generalization of the results in Hoffmann and Vetter (2017) to sequential processes for time dependent variable jump behaviour as in (2.4).

Critical values for the test procedures introduced below and the optimal choice of a regularization parameter of the new estimator for gradual change points are obtained by a multiplier bootstrap approach. Precisely, Theorem 6.4, in which we prove conditional weak convergence in a suitable sense of the bootstrapped version

$$\hat{G}_\rho^{(n)}(\theta,t) = \frac{1}{\sqrt{n\Delta_n}} \sum_{i=1}^{\lfloor n\theta \rfloor} \xi_i \rho\big(\Delta_i^n X^{(n)}\big) \mathbb{1}_{(-\infty,t]}\big(\Delta_i^n X^{(n)}\big) \mathbb{1}_{\{|\Delta_i^n X^{(n)}|>v_n\}},$$

($(\theta,t) \in [0,1]\times\mathbb{R}$) of $G_\rho^{(n)}$ to a Gaussian process, where $(\xi_i)_{i\in\mathbb{N}}$ is a sequence of i.i.d. multipliers with mean 0 and variance 1, complements the paper Hoffmann and Vetter (2017).

For the rescaling assumptions (2.1) and (2.4) we consider transition kernels $g_i(y,dz)$ of the set $\mathcal{G}(\beta,p)$ depending on parameters $\beta \in (0,2), p > 0$. In order to define this set we denote by λ the one-dimensional Lebesgue measure defined on the Lebesgue σ-algebra $\mathcal{L}_1$ of $\mathbb{R}$ and we denote by λ_1 the restriction of λ to the trace σ-algebra $[0,1]\cap\mathcal{L}_1$.

Definition 2.1. *For $\beta \in (0,2)$ and $p > 0$ the set $\mathcal{G}(\beta,p)$ consists of all transition kernels $g(y,dz)$ from $([0,1],\mathbb{B}([0,1]))$ into $(\mathbb{R},\mathbb{B})$, such that for each $y \in [0,1]$ the measure $g(y,dz)$ has a Lebesgue density $h_y(z)$ and there exist $\eta, M > 0$ as well as a Lebesgue null set $L \in [0,1]\cap\mathcal{L}_1$ such that the following items are satisfied:*

(1) $h_y(z) \le K|z|^{-(1+\beta)}$ holds for all $z \in (-\eta,\eta)$, $y \in [0,1]\setminus L$ and for some $K > 0$.
(2) For $n \in \mathbb{N}$ let $C_n := \{z \in \mathbb{R} \mid \frac{1}{n} \le |z| \le n\}$. Then for each $n \in \mathbb{N}$ there exists a $K_n > 0$ with $h_y(z) \le K_n$ for each $z \in C_n$ and all $y \in [0,1]\setminus L$.
(3) $h_y(z) \le K|z|^{-(2p\vee 2)-\epsilon}$ whenever $|z| \ge M$ and $y \in [0,1]\setminus L$, for some $K > 0$ and some $\epsilon > 0$.

The items above basically say that the densities h_y are bounded by a continuous Lévy density of a Lévy measure which behaves near zero like the one of a β-stable process, whereas this density has to decay sufficiently fast at infinity. Such conditions are well-known in the literature and often used in similar works on high-frequency statistics; see e.g. Aït-Sahalia and Jacod (2009a) or Aït-Sahalia and Jacod (2010). From Assumption 6.1 and Proposition 6.2 in

Section 6 it can be seen that it is even possible to work with a wider class of transition kernels $g(y, dz)$ which does not require Lebesgue densities. Nevertheless, we stick to the set $\mathcal{G}(\beta, p)$ defined above which is much simpler to interpret. The following example shows that alternatives of abrupt changes in the jump behaviour can be described by transition kernels in the set $\mathcal{G}(\beta, p)$.

Example 2.2 (Abrupt changes)**.** *In Section 3 we introduce statistical procedures for inference of abrupt changes in the jump behaviour. In this case the kernel* g_0 *is typically of the form as discussed below. For* $\beta \in (0, 2)$ *and* $p > 0$ *let* $\mathcal{M}(\beta, p)$ *be the set of all Lévy measures* ν *such that the constant transition kernel* $g(y, dz) = \nu(dz)$ *belongs to* $\mathcal{G}(\beta, p)$*.*

Let $\theta_0 \in (0, 1]$ *and let* $\nu_1, \nu_2 \in \mathcal{M}(\beta, p)$ *be two Lévy measures. Then the transition kernel* g_0 *given by*

$$g_0(y, dz) = \begin{cases} \nu_1(dz), & \text{for } y \in [0, \theta_0] \\ \nu_2(dz), & \text{for } y \in (\theta_0, 1] \end{cases} \tag{2.5}$$

is an element of $\mathcal{G}(\beta, p)$*. In the context of change-point tests* $\theta_0 = 1$ *corresponds to the null hypothesis of no change in the jump behaviour, whereas* (2.5) *describes an abrupt change for* $\theta_0 \in (0, 1)$ *and* $\nu_1 \neq \nu_2$*.*

The variance gamma process is a common model for the log stock price in finance (see for instance Madan et al. (1998)). Moreover, the Lévy measure of a variance gamma process has the form $\nu(dz) = (a_1 z^{-1} e^{-b_1 z} - a_2 z^{-1} e^{-b_2 z})\, dz$ for $a_1, a_2, b_1, b_2 > 0$. Thus, the transition kernel $g_0(y, dz)$ belongs to $\mathcal{G}(\beta, p)$ for all $\beta \in (0, 2)$ and $p > 0$, if similar as in (2.5) g_0 is piecewise constant in $y \in [0, 1]$ and on the domains of constancy it is equal to the Lévy measure of a variance gamma process.

For the asymptotic statements in this paper we require the following assumptions. Our results are also correct under less restrictive but more technical conditions. For the sake of a transparent presentation these are not presented here but deferred to Section 6.1.

Assumption 2.3. *Let* $0 < \beta < 2$ *and* $0 < \tau < (1/5 \wedge \frac{2-\beta}{2+5\beta})$*. Furthermore, let* $p > \beta + ((\frac{1}{2} + \frac{3}{2}\beta) \vee \frac{2}{1+5\tau})$*. At step* $n \in \mathbb{N}$ *we observe an Itō semimartingale* $X^{(n)}$ *adapted to the filtration of some filtered probability space* $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}_+}, \mathbb{P})$ *with characteristics* $(b_s^{(n)}, \sigma_s^{(n)}, \nu_s^{(n)})$ *at the equidistant time points* $\{i\Delta_n \mid i = 0, 1, \ldots, n\}$ *such that the following items are satisfied:*

(a) Assumptions on the jump characteristic and the function ρ*:*

(1) For each $n \in \mathbb{N}$ *and* $s \in [0, n\Delta_n]$ *we have*

$$\nu_s^{(n)}(dz) = g^{(n)}\Big(\frac{s}{n\Delta_n}, dz\Big), \tag{2.6}$$

where there exist transition kernels $g_0, g_1, g_2 \in \mathcal{G}(\beta, p)$ *such that for each* $y \in [0, 1]$

$$g^{(n)}(y, dz) = g_0(y, dz) + \frac{1}{\sqrt{n\Delta_n}} g_1(y, dz) + \mathcal{R}_n(y, dz) \tag{2.7}$$

and for each $y \in [0,1]$, $B \in \mathbb{B}$ and $n \in \mathbb{N}$ the kernel $\mathcal{R}_n$ satisfies $\mathcal{R}_n(y,B) \leq a_n g_2(y,B)$ for a sequence $a_n = o((n\Delta_n)^{-1/2})$ of non-negative real numbers.

(2) *$\rho\colon \mathbb{R} \to \mathbb{R}$ is a bounded $\mathcal{C}^1$-function with $\rho(0) = 0$ and its derivative satisfies $|\rho'(z)| \leq K|z|^{p-1}$ for all $z \in \mathbb{R}$ and some constant $K > 0$.*

(3) *$\rho(z) \neq 0$ for each $z \neq 0$.*

(4) *For every $t \in \mathbb{R}$ there exists a finite set $M_{(t)} \subset [0,1]$, such that the function*

$$y \mapsto \int_{-\infty}^{t} \rho(z) g_0(y,dz)$$

is continuous on $[0,1] \setminus M_{(t)}$.

(b) *Assumptions on the truncation sequence v_n and the observation scheme: The truncation sequence v_n satisfies $v_n := \gamma \Delta_n^{\overline{w}}$, with $\overline{w} = (1+5\tau)/4$ and some $\gamma > 0$. Define further $t_1 := (1+\tau)^{-1}$ and $t_2 := ((7\tau+1)/2)^{-1} \wedge 1$ (note that $0 < t_1 < t_2 \leq 1$) and we suppose that the observation scheme satisfies for some $\delta > 0$*

$$\Delta_n = o(n^{-t_1}) \quad \text{and} \quad n^{-t_2+\delta} = o(\Delta_n).$$

(c) *Assumptions on the drift and the diffusion coefficient: For $m_b = \frac{6+10\tau}{3-5\tau} \leq 4$ and $m_\sigma = \frac{6+10\tau}{1-5\tau}$ we have*

$$\sup_{n \in \mathbb{N}} \sup_{s \in \mathbb{R}_+} \left\{ \mathbb{E}\big|b_s^{(n)}\big|^{m_b} \vee \mathbb{E}\big|\sigma_s^{(n)}\big|^{m_\sigma} \right\} < \infty.$$

Remark 2.4. *Suppose we have complete knowledge of the distribution function $N_\rho(g_0;\theta,t)$. Obviously, the measure with density $M(dy,dz) := \rho(z)g_0(y,dz)dy$ is completely determined from knowledge of the entire function $N_\rho(g_0;\cdot,\cdot)$ and does not charge $[0,1]\times\{0\}$. Therefore, due to Assumption 2.3(a3) $1/\rho(z)M(dy,dz) = g_0(y,dz)dy$ and consequently the jump behaviour corresponding to g_0 is known as well. Furthermore, Assumption 2.3(a4) ensures that a characteristic quantity for a gradual change, which we introduce in Section 4 is zero if and only if the jump behaviour corresponding to g_0 is constant in time. All convergence results in this paper also hold without Assumption 2.3(a3) and (a4). Moreover, the function*

$$\tilde{\rho}(x) = \begin{cases} 0, & \text{if } x = 0, \\ e^{-1/|x|}, & \text{if } |x| > 0, \end{cases}$$

is suitable for any choice of the constants β and τ. In practice, however, one would like to work with a polynomial decay at zero, in which case the condition on p comes into play. Here, the smaller the parameter β, the smaller p can be chosen. For example, for $\beta < 3/5$ and $\tau > 3/35$ even a choice $p < 2$ is possible.

Furthermore, it is also important to choose the observation scheme suitably. Obviously, we have $\Delta_n \to 0$ and $n\Delta_n \to \infty$ because of $0 < t_1 < t_2 \leq 1$, and a typical choice is $\Delta_n = O(n^{-y})$ and $n^{-y} = O(\Delta_n)$ for some $0 < t_1 < y < t_2 \leq 1$. Finally, Assumption 2.3(c) requires only a bound on the moments of the remaining characteristics and is therefore extremely mild.

In the remaining part of this section we illustrate an example of a kernel $g_0 \in \mathcal{G}(\beta, p)$ for some suitable β, p and a function ρ satisfying Assumption 2.3(a2) and (a3).

Example 2.5 (Gradual changes)**.** *In Section 4, which is dedicated to inference of gradual changes, we basically test against the general alternative that the jump behaviour is non-constant. In the following we introduce an example of a kernel g_0 which can be used to describe a gradual change in the jump behaviour and a corresponding function ρ satisfying Assumption 2.3(a2) and (a3). To this end, for $L > 0$, $p > 1$ let*

$$\rho_{L,p}(z) := L \times \begin{cases} 2|z|^p, & \text{for } |z| \leq 1 \\ 4p|z| - pz^2 + 2 - 3p, & \text{for } 1 \leq |z| \leq 2 \\ 2 + p, & \text{for } |z| \geq 2 \end{cases} \tag{2.8}$$

and for $0 < \beta < 2$, $p > 1$ consider the Lévy density

$$h_{\beta,p}(z) := |z|^{-(1+\beta)} \mathbb{1}_{\{0<|z|<1\}} + \mathbb{1}_{\{1 \leq |z| \leq 2\}} + |z|^{-p} \mathbb{1}_{\{|z|>2\}}.$$

Furthermore, for $0 < \hat{\beta} < 2$ and $\hat{p} > 1 \vee \hat{\beta}$ let $A : [0,1] \to (0, \infty)$, $\beta : [0,1] \to (0, \hat{\beta}]$ and $p : [0,1] \to [2\hat{p} + \varepsilon, \infty)$ for some $\varepsilon > 0$ be Borel measurable functions such that A is bounded. Then, the kernel

$$g_0(y, dz) = A(y) h_{\beta(y),p(y)}(z) dz, \quad y \in [0,1] \tag{2.9}$$

belongs to $\mathcal{G}(\hat{\beta}, \hat{p})$ and for arbitrary $L > 0$ the function $\rho_{L,\hat{p}}$ satisfies Assumption 2.3(a2) and (a3).

3. Statistical inference for abrupt changes

In this section we deduce test and estimation procedures for abrupt changes in the jump behaviour of the underlying process, that is we investigate the situation of Example 2.2. To this end, we test the null hypothesis of no change in the jump behaviour

$\mathbf{H}_0$: Assumption 2.3 is satisfied for $g_1 = g_2 = 0$ and there exists a Lévy measure ν_0 such that $g_0(y, dz) = \nu_0(dz)$ for Lebesgue almost every $y \in [0,1]$.

against the alternative that the jump behaviour is constant on two intervals

$\mathbf{H}_1$: Assumption 2.3 is satisfied for $g_1 = g_2 = 0$ and there exists some $\theta_0 \in (0,1)$ and two Lévy measures $\nu_1 \neq \nu_2$ such that g_0 has the form (2.5).

The corresponding alternative for fixed $t_0 \in \mathbb{R}$ is given by:

$\mathbf{H}_1^{(\rho,t_0)}$: We have the situation from $\mathbf{H}_1$, but with $N_\rho(\nu_1; t_0) \neq N_\rho(\nu_2; t_0)$, where

$$N_\rho(\nu; t) = \int_{-\infty}^{t} \rho(z) \nu(dz) \tag{3.1}$$

for a Lévy measure ν.

Moreover, we investigate the behaviour of the tests introduced in this section under local alternatives which tend to the null hypothesis as $n \to \infty$:

$\mathbf{H}_1^{(loc)}$: Assumption 2.3 is satisfied with $g_0(y, dz) = \nu_0(dz)$ for Lebesgue-a.e. $y \in [0,1]$ for some Lévy measure ν_0 and with some transition kernels $g_1, g_2 \in \mathcal{G}(\beta, p)$.

3.1. Weak convergence of test statistics

Following Inoue (2001) a suitable approach to introduce tests for the hypotheses above is to investigate the convergence behaviour of the CUSUM process

$$\mathbb{T}_\rho^{(n)}(\theta, t) = \sqrt{n\Delta_n}\Big(N_\rho^{(n)}(\theta, t) - \frac{\lfloor n\theta \rfloor}{n} N_\rho^{(n)}(1, t)\Big), \tag{3.2}$$

with $N_\rho^{(n)}(\theta, t)$ defined in (2.3). The corresponding test rejects the null hypothesis $\mathbf{H}_0$ for large values of the Kolmogorov-Smirnov-type statistic

$$T_\rho^{(n)} = \sup_{(\theta,t)\in[0,1]\times\mathbb{R}} \big|\mathbb{T}_\rho^{(n)}(\theta, t)\big|.$$

The theorem below establishes functional weak convergence of $\mathbb{T}_\rho^{(n)}$ in the general case of local alternatives.

Theorem 3.1. *Under $\mathbf{H}_1^{(loc)}$ the process $\mathbb{T}_\rho^{(n)}$ converges weakly in $\ell^\infty([0,1]\times\mathbb{R})$ to the process $\mathbb{T}_\rho + \mathbb{T}_{\rho,g_1}$, where the tight mean zero Gaussian process $\mathbb{T}_\rho$ has the covariance structure*

$$\mathbb{E}\{\mathbb{T}_\rho(\theta_1, t_1)\mathbb{T}_\rho(\theta_2, t_2)\} = \{(\theta_1 \wedge \theta_2) - \theta_1\theta_2\} \int_{-\infty}^{t_1 \wedge t_2} \rho^2(z)\nu_0(dz) \tag{3.3}$$

and the deterministic function $\mathbb{T}_{\rho,g_1} \in \ell^\infty([0,1]\times\mathbb{R})$ is given by

$$\mathbb{T}_{\rho,g_1}(\theta, t) = N_\rho(g_1; \theta, t) - \theta N_\rho(g_1; 1, t), \tag{3.4}$$

where $N_\rho(g_1; \cdot, \cdot)$ is defined in (2.2).

As an immediate consequence of the previous result and the continuous mapping theorem we obtain weak convergence of the statistic $T_\rho^{(n)}$.

Corollary 3.2. *Suppose $\mathbf{H}_1^{(loc)}$ is true, then we have*

$$T_\rho^{(n)} \rightsquigarrow T_{\rho,g_1} := \sup_{(\theta,t)\in[0,1]\times\mathbb{R}} \big|\mathbb{T}_\rho(\theta, t) + \mathbb{T}_{\rho,g_1}(\theta, t)\big|, \tag{3.5}$$

in $(\mathbb{R}, \mathbb{B})$ with $\mathbb{T}_\rho + \mathbb{T}_{\rho,g_1}$ the limit process in Theorem 3.1.

In applications the Lévy measure ν_0 which describes the limiting jump behaviour of the underlying process is usually unknown. If one is only interested in the detection of changes in the distribution function $N_\rho(\nu_0; t_0)$ for a fixed $t_0 \in \mathbb{R}$, the processes

$$\mathbb{V}^{(n)}_{\rho,t_0}(\theta) := \frac{\mathbb{T}^{(n)}_\rho(\theta, t_0)}{\sqrt{N^{(n)}_{\rho^2}(1, t_0)}} \mathbb{1}_{\{N^{(n)}_{\rho^2}(1,t_0)>0\}}, \quad \theta \in [0,1]$$

converge weakly to a shifted version of a pivotal limit process.

Proposition 3.3. *Under $\mathbf{H}_1^{(loc)}$ for each fixed $t_0 \in \mathbb{R}$ with $N_{\rho^2}(\nu_0; t_0) > 0$ we have $\mathbb{V}^{(n)}_{\rho,t_0} \rightsquigarrow \mathbb{K} + \bar{\mathbb{V}}^{(g_1)}_{\rho,t_0}$ in $\ell^\infty([0,1])$, where $\mathbb{K}$ denotes a standard Brownian bridge and with the deterministic function*

$$\bar{\mathbb{V}}^{(g_1)}_{\rho,t_0}(\theta) := \frac{\mathbb{T}_{\rho,g_1}(\theta, t_0)}{\sqrt{N_{\rho^2}(\nu_0; t_0)}} \in \ell^\infty([0,1]),$$

where $N_{\rho^2}(\nu_0; \cdot)$ is defined in (3.1). In particular,

$$V^{(n)}_{\rho,t_0} := \sup_{\theta\in[0,1]} \big|\mathbb{V}^{(n)}_{\rho,t_0}(\theta)\big| \rightsquigarrow \bar{V}^{(g_1)}_{\rho,t_0} := \sup_{\theta\in[0,1]} \big|\mathbb{K}(\theta) + \bar{\mathbb{V}}^{(g_1)}_{\rho,t_0}(\theta)\big|. \tag{3.6}$$

Quantiles of functionals of the limit process $\mathbb{T}_\rho + \mathbb{T}_{\rho,g_1}$ in Theorem 3.1 are not easily accessible since the distribution of such functionals usually depends in a complicated way on the unknown quantities ν_0 and g_1 in the jump characteristic of the underlying process. In order to obtain reasonable approximations for these quantiles we use a multiplier bootstrap approach. That is, in the following we consider bootstrapped processes, $\hat{Y}_n = \hat{Y}_n(X_1, \dots, X_n, \xi_1, \dots, \xi_n)$, which depend on random variables $X_1, \dots, X_n$ defined on a probability space $(\Omega_X, \mathcal{F}_X, \mathbb{P}_X)$ and on random weights $\xi_1, \dots, \xi_n$ which are defined on a distinct probability space $(\Omega_\xi, \mathcal{F}_\xi, \mathbb{P}_\xi)$. Thus, the processes $\hat{Y}_n$ live on the product space $(\Omega, \mathcal{A}, \mathbb{P}) := (\Omega_X, \mathcal{A}_X, \mathbb{P}_X) \otimes (\Omega_\xi, \mathcal{A}_\xi, \mathbb{P}_\xi)$. Below we use the notion of weak convergence conditional on the sequence $(X_i)_{i\in\mathbb{N}}$ in probability. It can be found in Kosorok (2008) on pp. 19–20.

Definition 3.4. *Let $\hat{Y}_n = \hat{Y}_n(X_1, \dots, X_n; \xi_1, \dots, \xi_n)\colon (\Omega, \mathcal{A}, \mathbb{P}) \to \mathbb{D}$ be a random element taking values in some metric space $\mathbb{D}$ depending on some random variables $X_1, \dots, X_n$ and some random weights $\xi_1, \dots, \xi_n$. Moreover, let Y be a tight, Borel measurable random variable into $\mathbb{D}$. Then $\hat{Y}_n$ converges weakly to Y conditional on the data $X_1, X_2, \dots$ in probability, if and only if*

(a) $\sup_{f\in BL_1(\mathbb{D})} |\mathbb{E}_\xi f(\hat{Y}_n) - \mathbb{E} f(Y)| \overset{\mathbb{P}^*}{\to} 0,$

(b) $\mathbb{E}_\xi f(\hat{Y}_n)^* - \mathbb{E}_\xi f(\hat{Y}_n)_* \overset{\mathbb{P}}{\to} 0$ *for all* $f \in BL_1(\mathbb{D})$.

Here, $\mathbb{E}_\xi$ denotes the conditional expectation over the weights ξ given the data $X_1, \dots, X_n$, whereas $BL_1(\mathbb{D})$ is the space of all real-valued Lipschitz continuous functions f on $\mathbb{D}$ with sup-norm $\|f\|_{\mathbb{D}} \le 1$ and Lipschitz constant 1. Here and below we denote the sup-norm of a real valued function f on a set M by $\|f\|_M$. Furthermore, in item (b) $f(\hat{Y}_n)^*$ and $f(\hat{Y}_n)_*$ denote a minimal measurable majorant and a maximal measurable minorant with respect to the joint probability space $(\Omega, \mathcal{A}, \mathbb{P})$. The type of convergence defined above is denoted by $\hat{Y}_n \rightsquigarrow_\xi Y$.

Remark 3.5.

(i) Throughout this work all expressions $f(\hat{Y}_n)$, with a bootstrapped statistic $\hat{Y}_n$ and a Lipschitz continuous function f, are measurable functions of the random weights. To this end we do not use a measurable majorant or minorant in item (a) in the definition above.

(ii) The implication "(ii) ⇒ (i)" in the proof of Theorem 2.9.6 in Van der Vaart and Wellner (1996) shows that conditional weak convergence $\rightsquigarrow_\xi$ implies unconditional weak convergence $\rightsquigarrow$ with respect to the product measure $\mathbb{P}$.

For the results on conditional weak convergence of the bootstrapped processes below we require a rather mild additional assumption on the sequence of multipliers, which is satisfied for many common distributions such as for instance the Gaussian, the Poisson or the Binomial distribution.

Assumption 3.6. *The sequence $(\xi_i)_{i\in\mathbb{N}}$ is defined on a distinct probability space than the one generating the data $\{X^{(n)}_{i\Delta_n} \mid i = 0, 1, \ldots, n\}$ as described above, is i.i.d. with mean zero, variance one and there exists an $M > 0$ such that for each integer $m \geq 2$ we have*

$$\mathbb{E}|\xi_1|^m \leq m! M^m.$$

Reasonable bootstrap counterparts $\hat{\mathbb{T}}^{(n)}_\rho$ of the processes $\mathbb{T}^{(n)}_\rho$ are given by

$$\begin{aligned}
\hat{\mathbb{T}}^{(n)}_\rho(\theta, t) &:= \hat{\mathbb{T}}^{(n)}_\rho(X^{(n)}_{\Delta_n}, \ldots, X^{(n)}_{n\Delta_n}; \xi_1, \ldots, \xi_n; \theta, t) := \\
&\sqrt{n\Delta_n}\frac{\lfloor n\theta\rfloor}{n}\frac{n-\lfloor n\theta\rfloor}{n}\Big[\frac{1}{\lfloor n\theta\rfloor\Delta_n}\sum_{j=1}^{\lfloor n\theta\rfloor}\xi_j\rho(\Delta_j^n X^{(n)})\times \\
&\qquad\qquad \times \mathbb{1}_{(-\infty,t]}(\Delta_j^n X^{(n)})\mathbb{1}_{\{|\Delta_j^n X^{(n)}|>v_n\}} \\
&\quad - \frac{1}{(n-\lfloor n\theta\rfloor)\Delta_n}\sum_{j=\lfloor n\theta\rfloor+1}^{n}\xi_j\rho(\Delta_j^n X^{(n)})\mathbb{1}_{(-\infty,t]}(\Delta_j^n X^{(n)})\mathbb{1}_{\{|\Delta_j^n X^{(n)}|>v_n\}}\Big].
\end{aligned} \tag{3.7}$$

In the following theorem we establish conditional weak convergence of $\hat{\mathbb{T}}^{(n)}_\rho$ under the general assumptions of Section 2.

Theorem 3.7. *Let Assumption 2.3 be valid and let the multipliers $(\xi_j)_{j\in\mathbb{N}}$ satisfy Assumption 3.6. Then we have*

$$\hat{\mathbb{T}}^{(n)}_\rho \rightsquigarrow_\xi \mathbb{T}_\rho$$

in $\ell^\infty([0,1]\times\mathbb{R})$, where $\mathbb{T}_\rho$ is a tight mean zero Gaussian process in $\ell^\infty([0,1]\times\mathbb{R})$ with covariance function

$$\mathbb{E}\{\mathbb{T}_\rho(\theta_1, t_1)\mathbb{T}_\rho(\theta_2, t_2)\} \tag{3.8}$$

$$= \int_0^{\theta_1 \wedge \theta_2} \int_{-\infty}^{t_1 \wedge t_2} \rho^2(z) g_0(y, dz) dy - \theta_1 \int_0^{\theta_2} \int_{-\infty}^{t_1 \wedge t_2} \rho^2(z) g_0(y, dz) dy \\ - \theta_2 \int_0^{\theta_1} \int_{-\infty}^{t_1 \wedge t_2} \rho^2(z) g_0(y, dz) dy + \theta_1 \theta_2 \int_0^1 \int_{-\infty}^{t_1 \wedge t_2} \rho^2(z) g_0(y, dz) dy.$$

Remark 3.8. *The aim of our bootstrap procedure is to mimic the convergence behaviour of $\mathbb{T}_\rho^{(n)}$. The covariance function of the limiting process in Theorem 3.7 differs from* (3.3)*, because Theorem 3.7 holds under the general conditions introduced in Assumption 2.3, i.e. for an arbitrary kernel $g_0 \in \mathcal{G}(\beta, p)$. Under the null hypothesis $\mathbf{H}_0$, where we have $g_0(\cdot, dz) = \nu_0(dz)$, the covariance function* (3.8) *coincides with* (3.3).

The limit distribution of the Kolmogorov-Smirnov-type test statistic $T_\rho^{(n)}$ in Corollary 3.2 can be approximated under $\mathbf{H}_0$ by the bootstrap statistics in the following corollary, which is an immediate consequence of Proposition 10.7 in Kosorok (2008).

Corollary 3.9. *If Assumption 2.3 and Assumption 3.6 are satisfied, we have*

$$\hat{T}_\rho^{(n)} := \sup_{(\theta,t) \in [0,1] \times \mathbb{R}} |\hat{\mathbb{T}}_\rho^{(n)}(\theta, t)| \rightsquigarrow_\xi T_\rho := \sup_{(\theta,t) \in [0,1] \times \mathbb{R}} \big|\mathbb{T}_\rho(\theta, t)\big|,$$

with $\mathbb{T}_\rho$ the limit process in Theorem 3.7.

3.2. Test procedures for abrupt changes

The weak convergence results of the previous section make it possible to define test procedures for abrupt changes in the jump behaviour of the underlying process based on Lévy distribution functions of type (2.2). In the following let $B \in \mathbb{N}$ be some large number and let $(\xi^{(b)})_{b=1,\dots,B}$ be independent vectors of i.i.d. random variables $\xi^{(b)} = (\xi_j^{(b)})_{j=1,\dots,n}$ with mean zero and variance one, which satisfy Assumption 3.6. With $\hat{\mathbb{T}}_{\rho,\xi^{(b)}}^{(n)}$ and $\hat{T}_{\rho,\xi^{(b)}}^{(n)}$ we denote the corresponding bootstrapped quantity calculated with respect to the data and the b-th multiplier sequence $\xi^{(b)}$. For a given level $\alpha \in (0, 1)$, we propose to reject $\mathbf{H}_0$ in favor of $\mathbf{H}_1$, if

$$T_\rho^{(n)} \geq \hat{q}_{1-\alpha}^{(B)}\Big(T_\rho^{(n)}\Big), \tag{3.9}$$

where $\hat{q}_{1-\alpha}^{(B)}(T_\rho^{(n)})$ denotes the $(1-\alpha)$-sample quantile of $\hat{T}_{\rho,\xi^{(1)}}^{(n)}, \dots, \hat{T}_{\rho,\xi^{(B)}}^{(n)}$. Similarly, for $t_0 \in \mathbb{R}$, $\mathbf{H}_0$ is rejected in favor of $\mathbf{H}_1^{(\rho,t_0)}$, if

$$W_\rho^{(n,t_0)} := \sup_{\theta \in [0,1]} |\mathbb{T}_\rho^{(n)}(\theta, t_0)| \geq \hat{q}_{1-\alpha}^{(B)}\Big(W_\rho^{(n,t_0)}\Big), \tag{3.10}$$

where $\hat{q}_{1-\alpha}^{(B)}(W_\rho^{(n,t_0)})$ denotes the $(1-\alpha)$-sample quantile of $\hat{W}_{\rho,\xi^{(1)}}^{(n,t_0)}, \dots, \hat{W}_{\rho,\xi^{(B)}}^{(n,t_0)}$, and where $\hat{W}_{\rho,\xi^{(b)}}^{(n,t_0)} := \sup_{\theta \in [0,1]} |\hat{\mathbb{T}}_{\rho,\xi^{(b)}}^{(n)}(\theta, t_0)|$ for $b = 1, \dots, B$. Furthermore, according to Proposition 3.3 we define an exact test procedure, that is $\mathbf{H}_0$ is rejected in favor of the point-wise alternative $\mathbf{H}_1^{(\rho,t_0)}$, if

$$V_{\rho,t_0}^{(n)} \geq q_{1-\alpha}^K, \tag{3.11}$$

where $q_{1-\alpha}^K$ is the $(1-\alpha)$-quantile of the Kolmogorov-Smirnov-distribution, that is the distribution of the supremum of a standard Brownian bridge $K = \sup_{\theta\in[0,1]} |\mathbb{K}(\theta)|$.

The following results show the behaviour of the previously introduced tests under the null hypothesis, local alternatives and the alternatives of an abrupt change. In particular, these tests are consistent asymptotic level α tests. First, recall the tight centered Gaussian process $\mathbb{T}_\rho$ in $\ell^\infty([0,1]\times\mathbb{R})$ with covariance function (3.3), let $L_\rho : (\mathbb{R},\mathbb{B}) \to (\mathbb{R},\mathbb{B})$ be the distribution function of the supremum variable $\sup_{(\theta,t)\in[0,1]\times\mathbb{R}} |\mathbb{T}_\rho(\theta,t)|$ and let $L_\rho^{(t_0)}$ be the distribution function of $\sup_{\theta\in[0,1]} |\mathbb{T}_\rho(\theta,t_0)|$. Furthermore, recall the random variable

$$T_{\rho,g_1} = \sup_{(\theta,t)\in[0,1]\times\mathbb{R}} \big|\mathbb{T}_\rho(\theta,t) + \mathbb{T}_{\rho,g_1}(\theta,t)\big|,$$

defined in (3.5) with the deterministic function

$$\mathbb{T}_{\rho,g_1}(\theta,t) = N_\rho(g_1;\theta,t) - \theta N_\rho(g_1;1,t),$$

defined in (3.4) and let

$$T_{\rho,g_1}^{(t_0)} := \sup_{\theta\in[0,1]} \big|\mathbb{T}_\rho(\theta,t_0) + \mathbb{T}_{\rho,g_1}(\theta,t_0)\big|.$$

Then the results on consistency of the tests are as follows.

Proposition 3.10. *Under $\mathbf{H}_1^{(loc)}$ with $\nu_0 \neq 0$*

$$\begin{aligned}\mathbb{P}\big(L_\rho(T_{\rho,g_1}) > 1-\alpha\big) &\leq \liminf_{B\to\infty} \lim_{n\to\infty} \mathbb{P}\big(T_\rho^{(n)} \geq \hat{q}_{1-\alpha}^{(B)}\big(T_\rho^{(n)}\big)\big) \\ &\leq \limsup_{B\to\infty} \lim_{n\to\infty} \mathbb{P}\big(T_\rho^{(n)} \geq \hat{q}_{1-\alpha}^{(B)}\big(T_\rho^{(n)}\big)\big) \leq \mathbb{P}\big(L_\rho(T_{\rho,g_1}) \geq 1-\alpha\big)\end{aligned} \tag{3.12}$$

holds for each $\alpha \in (0,1)$ and additionally if $N_{\rho^2}(\nu_0,t_0) > 0$ then for all $\alpha\in(0,1)$ we have

$$\begin{aligned}\mathbb{P}\big(\bar{V}_{\rho,t_0}^{(g_1)} > q_{1-\alpha}^K\big) &\leq \liminf_{n\to\infty} \mathbb{P}\big(V_{\rho,t_0}^{(n)} \geq q_{1-\alpha}^K\big) \\ &\leq \limsup_{n\to\infty} \mathbb{P}\big(V_{\rho,t_0}^{(n)} \geq q_{1-\alpha}^K\big) \leq \mathbb{P}\big(\bar{V}_{\rho,t_0}^{(g_1)} \geq q_{1-\alpha}^K\big),\end{aligned} \tag{3.13}$$

with $V_{\rho,t_0}^{(n)}$ and $\bar{V}_{\rho,t_0}^{(g_1)}$ defined in (3.6), as well as

$$\begin{aligned}\mathbb{P}\big(L_\rho^{(t_0)}\big(T_{\rho,g_1}^{(t_0)}\big) > 1-\alpha\big) &\leq \liminf_{B\to\infty} \lim_{n\to\infty} \mathbb{P}\big(W_\rho^{(n,t_0)} \geq \hat{q}_{1-\alpha}^{(B)}\big(W_\rho^{(n,t_0)}\big)\big) \\ \leq \limsup_{B\to\infty} \lim_{n\to\infty} \mathbb{P}\big(W_\rho^{(n,t_0)} \geq \hat{q}_{1-\alpha}^{(B)}\big(W_\rho^{(n,t_0)}\big)\big) &\leq \mathbb{P}\big(L_\rho^{(t_0)}\big(T_{\rho,g_1}^{(t_0)}\big) \geq 1-\alpha\big).\end{aligned} \tag{3.14}$$

Remark 3.11. *According to Corollary 1.3 and Remark 4.1 in Gaenssler et al. (2007) the distribution function* L_ρ *is continuous on* $\mathbb{R}$ *and strictly increasing on* $\mathbb{R}_+$*. Thus, (3.12) basically states that under the local alternative for large* $B, n \in \mathbb{N}$ *the probability that the test (3.9) rejects the null hypothesis is approximately equal to the probability that the supremum of the shifted version* T_{ρ,g_1} *exceeds the* $(1-\alpha)$*-quantile of the non-shifted version* $T_{\rho,0}$*. An analysis of the latter probability, which is beyond the scope of this paper, then shows in which direction, i.e. for which* g_1*, it is harder to distinguish the null hypothesis from the alternative. The assertions (3.13) and (3.14) can be interpreted in the same way.*

Corollary 3.12. *Under* $\mathbf{H}_0$ *the tests (3.9), (3.10) and (3.11) have asymptotic level* α*, that is if* $\nu_0 \neq 0$ *we have for each* $\alpha \in (0,1)$

$$\lim_{B\to\infty}\lim_{n\to\infty} \mathbb{P}\big(T_\rho^{(n)} \geq \hat{q}_{1-\alpha}^{(B)}(T_\rho^{(n)})\big) = \alpha \tag{3.15}$$

and furthermore

$$\lim_{n\to\infty} \mathbb{P}\big(V_{\rho,t_0}^{(n)} \geq q_{1-\alpha}^K\big) = \alpha, \quad \lim_{B\to\infty}\lim_{n\to\infty} \mathbb{P}\big(W_\rho^{(n,t_0)} \geq \hat{q}_{1-\alpha}^{(B)}(W_\rho^{(n,t_0)})\big) = \alpha, \tag{3.16}$$

holds for all $\alpha \in (0,1)$*, if* $N_{\rho^2}(\nu_0; t_0) > 0$.

Proposition 3.13. *The tests (3.9), (3.10) and (3.11) are consistent in the following sense: Under* $\mathbf{H}_1$*, for all* $\alpha \in (0,1)$ *and all* $B \in \mathbb{N}$*, we have*

$$\lim_{n\to\infty} \mathbb{P}\big(T_\rho^{(n)} \geq \hat{q}_{1-\alpha}^{(B)}(T_\rho^{(n)})\big) = 1.$$

Under $\mathbf{H}_1^{(\rho,t_0)}$*, we have for all* $\alpha \in (0,1)$ *and all* $B \in \mathbb{N}$*,*

$$\lim_{n\to\infty} \mathbb{P}\big(V_{\rho,t_0}^{(n)} \geq q_{1-\alpha}^K\big) = 1 \quad \text{and} \quad \lim_{n\to\infty} \mathbb{P}\big(W_\rho^{(n,t_0)} \geq \hat{q}_{1-\alpha}^{(B)}(W_\rho^{(n,t_0)})\big) = 1.$$

3.3. Argmax-estimators

If one of the aforementioned tests rejects the null hypothesis in favor of an abrupt alternative the natural question arises of how to estimate the unknown break point θ_0. A typical approach in change-point analysis to this estimation problem is the so-called argmax-estimator, that is we basically take the argmax of the function $\theta \mapsto \sup_{t\in\mathbb{R}} |\mathbb{T}_\rho^{(n)}(\theta,t)|$ as an estimate for θ_0. Consistency of our estimators follows with the argmax continuous mapping theorem of Kim and Pollard (1990) using the following auxiliary result.

Proposition 3.14. *Under* $\mathbf{H}_1$*, the random function* $(\theta,t) \mapsto (n\Delta_n)^{-1/2}\mathbb{T}_\rho^{(n)}(\theta,t)$ *converges in* $\ell^\infty([0,1]\times\mathbb{R})$ *to the function*

$$T_{(1)}^\rho(\theta,t) := \begin{cases} \theta(1-\theta_0)\{N_\rho(\nu_1;t) - N_\rho(\nu_2;t)\}, & \text{if } \theta \leq \theta_0 \\ \theta_0(1-\theta)\{N_\rho(\nu_1;t) - N_\rho(\nu_2;t)\}, & \text{if } \theta \geq \theta_0 \end{cases}$$

in outer probability, where $N_\rho(\nu;\cdot)$ *is defined in (3.1).*

For the test problem $\mathbf{H}_0$ versus $\mathbf{H}_1$ we consider the estimator

$$\tilde{\theta}_\rho^{(n)} := \arg\max_{\theta\in[0,1]} \sup_{t\in\mathbb{R}} \big|\mathbb{T}_\rho^{(n)}(\theta,t)\big| \tag{3.17}$$

and in the setup $\mathbf{H}_0$ versus $\mathbf{H}_1^{(\rho,t_0)}$ a suitable estimator for the change point is given by

$$\tilde{\theta}_{\rho,t_0}^{(n)} := \arg\max_{\theta\in[0,1]} \big|\mathbb{T}_\rho^{(n)}(\theta,t_0)\big|.$$

The following proposition establishes consistency of these estimators.

Proposition 3.15. *Under $\mathbf{H}_1$ we have $\tilde{\theta}_\rho^{(n)} = \theta_0 + o_{\mathbb{P}}(1)$ for $n\to\infty$ and if the special case $\mathbf{H}_1^{(\rho,t_0)}$ is true we obtain $\tilde{\theta}_{\rho,t_0}^{(n)} = \theta_0 + o_{\mathbb{P}}(1)$.*

Remark 3.16. *For the sake of convenience we have focused on the case of one single break. The results on the tests in Section 3.2 also hold for alternatives with finitely many abrupt changes. Moreover, the estimation methods depicted above can easily be extended to detect multiple change points by a standard binary segmentation algorithm dating back to Vostrikova (1981). We illustrate this in Section 5.2, where a real data example is discussed.*

4. Statistical inference for gradual changes

As a generalization of Proposition 3.14 one can show that $(n\Delta_n)^{-1/2}\mathbb{T}_\rho^{(n)}(\theta,t)$ converges in $\ell^\infty([0,1]\times\mathbb{R})$ in outer probability to the function $\mathbb{T}_{\rho,g_0}$ defined in (3.4) whenever Assumption 2.3 is satisfied. Thus, under some regularity conditions, $\mathrm{argmax}_{\theta\in[0,1]}|\mathbb{T}_\rho^{(n)}(\theta,t)|$ is a consistent estimator of $\mathrm{argmax}_{\theta\in[0,1]}$ $|\mathbb{T}_{\rho,g_0}(\theta,t)|$. However, if the jump behaviour changes gradually at θ_0, the function $\theta\mapsto|\mathbb{T}_{\rho,g_0}(\theta,t)|$ is usually maximal at a point $\theta_1>\theta_0$. As a consequence the argmax-estimators investigated in Section 3.3 usually overestimate a change point, if the change is not abrupt. Therefore, in this section we introduce test and estimation procedures which are tailored for gradual changes in the entire jump behaviour.

4.1. A measure of time variation for the entire jump behaviour

If the jump behaviour is given by (2.1) for some suitable transition kernel $g=g_0$ from $([0,1],\mathbb{B}([0,1]))$ into $(\mathbb{R},\mathbb{B})$, we follow Vogt and Dette (2015) and base our analysis of gradual changes on the quantity

$$D_\rho^{(g_0)}(\zeta,\theta,t) := N_\rho(g_0;\zeta,t) - \frac{\zeta}{\theta}N_\rho(g_0;\theta,t), \quad (\zeta,\theta,t)\in C\times\mathbb{R} \tag{4.1}$$

with

$$C := \{(\zeta,\theta)\in[0,1]^2 \mid \zeta\le\theta\} \tag{4.2}$$

and where $N_\rho(g_0;\cdot,\cdot)$ is defined in (2.2). Here and throughout this paper we use the convention $\frac{0}{0}:=1$. We will address $D_\rho^{(g_0)}$ as the measure of time variation (with respect to ρ) of the entire jump behaviour of the underlying process, because the following lemma shows that $D_\rho^{(g_0)}$ indicates whether there is a change in the jump behaviour.

Lemma 4.1. *Let $\theta \in [0,1]$. Then $D_\rho^{(g_0)}(\zeta,\theta,t) = 0$ for all $0 \le \zeta \le \theta$ and $t \in \mathbb{R}$ if and only if the kernel $g_0(\cdot, dz)$ is Lebesgue almost everywhere constant on $[0,\theta]$.*

According to the preceding lemma there exists a (gradual) change in the jump behaviour given by g_0 if and only if

$$\sup_{\theta\in[0,1]} \tilde{\mathcal{D}}_\rho^{(g_0)}(\theta) > 0,$$

where $\tilde{\mathcal{D}}_\rho^{(g_0)}(\theta) := \sup\limits_{t\in\mathbb{R}} \sup_{0\le\zeta\le\theta} \left|D_\rho^{(g_0)}(\zeta,\theta,t)\right|$. As a consequence, the first point of a change in the jump behaviour is given by

$$\theta_0 := \inf\left\{\theta \in [0,1] \mid \tilde{\mathcal{D}}_\rho^{(g_0)}(\theta) > 0\right\}, \tag{4.3}$$

where we set $\inf\varnothing := 1$. We call θ_0 the change point of the jump behaviour of the underlying process. Notice that by the discussion after (4.2) the definition in (4.3) is independent of ρ. In Section 4.3 we construct an estimator for θ_0, where we only consider the quantity

$$\mathcal{D}_\rho^{(g_0)}(\theta) := \sup_{t\in\mathbb{R}} \sup_{0\le\zeta\le\theta'\le\theta} \left|D_\rho^{(g_0)}(\zeta,\theta',t)\right|, \tag{4.4}$$

instead of $\tilde{\mathcal{D}}_\rho^{(g_0)}$. On the one hand the monotonicity of $\mathcal{D}_\rho^{(g_0)}$ simplifies our entire presentation and on the other hand the first time point where $\mathcal{D}_\rho^{(g_0)}$ deviates from 0 is also given by θ_0, so it is equivalent to consider $\mathcal{D}_\rho^{(g_0)}$ instead. Our analysis of gradual changes is based on a consistent estimator $\mathbb{D}_\rho^{(n)}$ of $D_\rho^{(g_0)}$ which we construct in Section 4.2. Before that we illustrate the quantities introduced in (4.3) and (4.4) in the situations of Example 2.2 and Example 2.5.

Example 4.2. *Recall the situation of an abrupt change as in Example 2.2. Precisely, let $\beta \in (0,2)$, $p > 0$ and $\nu_1, \nu_2 \in \mathcal{M}(\beta,p)$ with $\nu_1 \ne \nu_2$ such that for some $\theta_0 \in (0,1)$ the transition kernel g_0 has the form*

$$g_0(y,dz) = \begin{cases} \nu_1(dz), & \text{for } y \in [0,\theta_0], \\ \nu_2(dz), & \text{for } y \in (\theta_0,1]. \end{cases} \tag{4.5}$$

Obviously, for some function $\rho : \mathbb{R} \to \mathbb{R}$ such that Assumption 2.3(a2) and (a3) are satisfied we have $D_\rho^{(g_0)}(\zeta,\theta',t) = 0$ for each $(\zeta,\theta',t) \in C \times \mathbb{R}$ with $\theta' \le \theta_0$ and consequently $\mathcal{D}_\rho^{(g_0)}(\theta) = 0$ for each $\theta \le \theta_0$. On the other hand, if $\theta_0 < \theta' \le 1$ and $\zeta \le \theta_0$ we have

$$\begin{aligned} D_\rho^{(g_0)}(\zeta,\theta',t) &= \zeta N_\rho(\nu_1;t) - \frac{\zeta}{\theta'}(\theta_0 N_\rho(\nu_1;t) + (\theta'-\theta_0)N_\rho(\nu_2;t)) \\ &= \zeta(N_\rho(\nu_2;t) - N_\rho(\nu_1;t))\big(\frac{\theta_0}{\theta'} - 1\big) \end{aligned}$$

with $N_\rho(\nu;t)$ defined in (3.1) and we obtain

$$\sup_{t\in\mathbb{R}}\sup_{\zeta\le\theta_0}|D_\rho^{(g_0)}(\zeta,\theta',t)| = V_0^\rho\theta_0\big(1-\frac{\theta_0}{\theta'}\big),$$

where $V_0^\rho = \sup_{t\in\mathbb{R}}|N_\rho(\nu_1;t)-N_\rho(\nu_2;t)| > 0$, because of $\nu_1\neq\nu_2$ and the assumptions on ρ. For $\theta_0<\zeta\le\theta'$ a similar calculation yields

$$D_\rho^{(g_0)}(\zeta,\theta',t) = \theta_0(N_\rho(\nu_2;t)-N_\rho(\nu_1;t))\big(\frac{\zeta}{\theta'}-1\big)$$

which gives

$$\sup_{t\in\mathbb{R}}\sup_{\theta_0<\zeta\le\theta'}\big|D_\rho^{(g_0)}(\zeta,\theta',t)\big| = V_0^\rho\theta_0\big(1-\frac{\theta_0}{\theta'}\big).$$

Therefore, it follows that the quantity defined in (4.3) is given by θ_0, because for $\theta>\theta_0$ we have

$$\begin{aligned}\mathcal{D}_\rho^{(g_0)}(\theta) &= \sup_{\theta_0<\theta'\le\theta}\max\Big\{\sup_{t\in\mathbb{R}}\sup_{\zeta\le\theta_0}\big|D_\rho^{(g_0)}(\zeta,\theta',t)\big|,\ \sup_{t\in\mathbb{R}}\sup_{\theta_0<\zeta\le\theta'}\big|D_\rho^{(g_0)}(\zeta,\theta',t)\big|\Big\}\\ &= V_0^\rho\theta_0\big(1-\frac{\theta_0}{\theta}\big). \end{aligned}\tag{4.6}$$

Example 4.3. *Recall the situation of Example 2.5. Let the transition kernel g_0 be of the form (2.9) such that there exist $\theta_0\in(0,1)$, $A_0\in(0,\infty)$, $\beta_0\in(0,\hat\beta]$ and $p_0\in[2\hat p+\varepsilon,\infty)$ for some $\varepsilon>0$ with*

$$A(y)=A_0,\quad \beta(y)=\beta_0\quad\text{ and }\quad p(y)=p_0 \tag{4.7}$$

for each $y\in[0,\theta_0]$. Additionally, let θ_0 be contained in an open interval U with a real analytic function $\bar A: U\to(0,\infty)$ and affine linear functions $\bar\beta: U\to(0,\hat\beta]$, $\bar p: U\to[2\hat p+\varepsilon,\infty)$ such that at least one of the functions $\bar A$, $\bar\beta$ and $\bar p$ is non-constant and

$$A(y)=\bar A(y),\quad \beta(y)=\bar\beta(y),\quad\text{ as well as }\quad p(y)=\bar p(y) \tag{4.8}$$

for all $y\in[\theta_0,1)\cap U$. Then the quantity defined in (4.3) is given by θ_0.

4.2. The empirical measure of time variation and its convergence behaviour

Suppose we have established that $N_\rho^{(n)}(\cdot,\cdot)$ is a consistent estimator for $N_\rho(g_0;\cdot,\cdot)$. Then with the set C defined in (4.2) it is reasonable to consider

$$\mathbb{D}_\rho^{(n)}(\zeta,\theta,t) := N_\rho^{(n)}(\zeta,t)-\frac{\zeta}{\theta}N_\rho^{(n)}(\theta,t),\quad (\zeta,\theta,t)\in C\times\mathbb{R}, \tag{4.9}$$

as an estimate for the measure of time variation of the entire jump behaviour $D_\rho^{(g_0)}$ defined in (4.1). In the following we want to establish consistency of the empirical measure of time variation $\mathbb{D}_\rho^{(n)}$. To be precise, the following two theorems show that the process

$$\mathbb{H}_\rho^{(n)}(\zeta,\theta,t) := \sqrt{n\Delta_n}\big(\mathbb{D}_\rho^{(n)}(\zeta,\theta,t) - D_\rho^{(g_0)}(\zeta,\theta,t)\big), \tag{4.10}$$

and its bootstrapped counterpart converge weakly or weakly conditional on the data in probability, respectively, to a suitable tight mean zero Gaussian process.

Theorem 4.4. *If Assumption 2.3 is satisfied, then the process $\mathbb{H}_\rho^{(n)}$ defined in (4.10) converges weakly, that is $\mathbb{H}_\rho^{(n)} \rightsquigarrow \mathbb{H}_\rho + D_\rho^{(g_1)}$ in $\ell^\infty(C\times\mathbb{R})$, where $\mathbb{H}_\rho$ is a tight mean zero Gaussian process with covariance function*

$$\begin{aligned}
&\operatorname{Cov}\big(\mathbb{H}_\rho(\zeta_1,\theta_1,t_1),\mathbb{H}_\rho(\zeta_2,\theta_2,t_2)\big)\\
&= \int_0^{\zeta_1\wedge\zeta_2}\int_{-\infty}^{t_1\wedge t_2}\rho^2(z)g_0(y,dz)dy - \frac{\zeta_1}{\theta_1}\int_0^{\zeta_2\wedge\theta_1}\int_{-\infty}^{t_1\wedge t_2}\rho^2(z)g_0(y,dz)dy\\
&\quad - \frac{\zeta_2}{\theta_2}\int_0^{\zeta_1\wedge\theta_2}\int_{-\infty}^{t_1\wedge t_2}\rho^2(z)g_0(y,dz)dy + \frac{\zeta_1\zeta_2}{\theta_1\theta_2}\int_0^{\theta_1\wedge\theta_2}\int_{-\infty}^{t_1\wedge t_2}\rho^2(z)g_0(y,dz)dy.
\end{aligned} \tag{4.11}$$

For the statistical change-point inference proposed in the following sections we require quantiles of functionals of the limiting distribution in Theorem 4.4. (4.11) shows that this distribution depends in a complicated way on the unknown underlying kernel g_0 and therefore corresponding quantiles are difficult to estimate. In order to solve this problem we want to use a multiplier bootstrap approach similar to Section 3. To this end, we define the following bootstrap counterpart of the process $\mathbb{H}_\rho^{(n)}$

$$\begin{aligned}
\hat{\mathbb{H}}_\rho^{(n)}(\zeta,\theta,t) &:= \hat{\mathbb{H}}_\rho^{(n)}(X_{\Delta_n}^{(n)},\dots,X_{n\Delta_n}^{(n)};\xi_1,\dots,\xi_n;\zeta,\theta,t)\\
&:= \frac{1}{\sqrt{n\Delta_n}}\Bigg[\sum_{j=1}^{\lfloor n\zeta\rfloor}\xi_j\rho(\Delta_j^nX^{(n)})\mathbb{1}_{(-\infty,t]}(\Delta_j^nX^{(n)})\mathbb{1}_{\{|\Delta_j^nX^{(n)}|>v_n\}}\\
&\qquad - \frac{\zeta}{\theta}\sum_{j=1}^{\lfloor n\theta\rfloor}\xi_j\rho(\Delta_j^nX^{(n)})\mathbb{1}_{(-\infty,t]}(\Delta_j^nX^{(n)})\mathbb{1}_{\{|\Delta_j^nX^{(n)}|>v_n\}}\Bigg].
\end{aligned} \tag{4.12}$$

The result below establishes consistency of $\hat{\mathbb{H}}_\rho^{(n)}$.

Theorem 4.5. *Let Assumption 2.3 be valid and let the multiplier sequence $(\xi_i)_{i\in\mathbb{N}}$ satisfy Assumption 3.6. Then we have $\hat{\mathbb{H}}_\rho^{(n)} \rightsquigarrow_\xi \mathbb{H}_\rho$ in $\ell^\infty(C\times\mathbb{R})$, where the tight mean zero Gaussian process $\mathbb{H}_\rho$ has the covariance structure (4.11).*

4.3. Estimating the gradual change point

For the sake of a unique definition of the (gradual) change point θ_0 in (4.3) we suppose throughout this section that Assumption 2.3 holds with $g_1 = g_2 = 0$. Recall the definition

$$\mathcal{D}_\rho^{(g_0)}(\theta) = \sup_{t\in\mathbb{R}}\sup_{0\le\zeta\le\theta'\le\theta}\big|D_\rho^{(g_0)}(\zeta,\theta',t)\big|$$

in (4.4), then by Theorem 4.4 the process $\mathbb{D}_\rho^{(n)}(\zeta,\theta,t)$ from (4.9) is a consistent estimator of $D_\rho^{(g_0)}(\zeta,\theta,t)$. Therefore, we set

$$\mathbb{D}_{\rho,*}^{(n)}(\theta) := \sup_{t\in\mathbb{R}} \sup_{0\le\zeta\le\theta'\le\theta} \big|\mathbb{D}_\rho^{(n)}(\zeta,\theta',t)\big|,$$

and an application of the continuous mapping theorem and Theorem 4.4 yields the following result.

Corollary 4.6. *If Assumption 2.3 is satisfied with $g_1 = g_2 = 0$, then $(n\Delta_n)^{1/2}$ $\mathbb{D}_{\rho,*}^{(n)} \rightsquigarrow \mathbb{H}_{\rho,*}$ in $\ell^\infty\big([0,\theta_0]\big)$, where $\mathbb{H}_{\rho,*}$ is the tight process in $\ell^\infty([0,1])$ defined by*

$$\mathbb{H}_{\rho,*}(\theta) := \sup_{t\in\mathbb{R}} \sup_{0\le\zeta\le\theta'\le\theta} |\mathbb{H}_\rho(\zeta,\theta',t)|,$$

with the centered Gaussian process $\mathbb{H}_\rho$ defined in Theorem 4.4.

Below we obtain that the rate of convergence of an estimator for θ_0 depends on the smoothness of the curve $\theta \mapsto \mathcal{D}_\rho^{(g_0)}(\theta)$ at θ_0. Thus, we impose a kind of Taylor expansion of the function $\mathcal{D}_\rho^{(g_0)}$. More precisely, we assume throughout this section that $\theta_0 < 1$ and that there exist constants $\iota, \eta, \varpi, c > 0$ such that $\mathcal{D}_\rho^{(g_0)}$ admits an expansion of the form

$$\mathcal{D}_\rho^{(g_0)}(\theta) = c\big(\theta-\theta_0\big)^{\varpi} + \aleph(\theta) \tag{4.13}$$

for all $\theta \in [\theta_0, \theta_0+\iota]$, where the remainder term satisfies $|\aleph(\theta)| \le K\big(\theta-\theta_0\big)^{\varpi+\eta}$ for some $K > 0$. According to Theorem 4.4 we have $(n\Delta_n)^{1/2}\mathbb{D}_{\rho,*}^{(n)}(\theta) \to \infty$ in probability for any $\theta \in (\theta_0, 1]$. Consequently, if the deterministic sequence $\varkappa_n \to \infty$ is chosen appropriately, the statistic

$$r_\rho^{(n)}(\theta) := \mathbb{1}_{\{(n\Delta_n)^{1/2}\mathbb{D}_{\rho,*}^{(n)}(\theta)\le\varkappa_n\}},$$

should satisfy

$$r_\rho^{(n)}(\theta) \xrightarrow{\mathbb{P}} \begin{cases} 1, & \text{if } \theta \le \theta_0, \\ 0, & \text{if } \theta > \theta_0. \end{cases}$$

Thus, we define the estimator for the change point by

$$\hat\theta_\rho^{(n)} = \hat\theta_\rho^{(n)}(\varkappa_n) := \int_0^1 r_\rho^{(n)}(\theta)d\theta. \tag{4.14}$$

The theorem below establishes consistency of the estimator $\hat\theta_\rho^{(n)}$ under mild additional assumptions on the sequence $(\varkappa_n)_{n\in\mathbb{N}}$.

Theorem 4.7. *If Assumption 2.3 is satisfied with $g_1 = g_2 = 0$, $\theta_0 < 1$, and (4.13) holds for some $\varpi > 0$, then*

$$\hat\theta_\rho^{(n)} - \theta_0 = O_{\mathbb{P}}\Big(\Big(\frac{\varkappa_n}{\sqrt{n\Delta_n}}\Big)^{1/\varpi}\Big),$$

for any sequence $\varkappa_n \to \infty$ with $\varkappa_n/\sqrt{n\Delta_n} \to 0$.

Theorem 4.7 describes how the curvature of $\mathcal{D}_\rho^{(g_0)}$ at θ_0 determines the convergence behaviour of the estimator: A lower degree of smoothness of $\mathcal{D}_\rho^{(g_0)}$ in θ_0 yields a better rate of convergence. However, the estimator depends on the choice of the threshold level $\varkappa_n$ and we explain below how to choose this sequence with bootstrap methods in order to control the probability of over- and underestimation. But before that the following theorem investigates the mean squared error

$$\mathrm{MSE}(\varkappa_n) = \mathbb{E}\Big[\big(\hat\theta_\rho^{(n)}(\varkappa_n) - \theta_0\big)^2\Big]$$

of the estimator $\hat\theta_\rho^{(n)}$. Recall the definition of $\mathbb{H}_\rho^{(n)}$ in (4.10) and define

$$\mathbb{H}_{\rho,*}^{(n)}(\theta) := \sup_{t\in\mathbb{R}} \sup_{0\le\zeta\le\theta'\le\theta} |\mathbb{H}_\rho^{(n)}(\zeta,\theta',t)|, \quad \theta \in [0,1],$$

which is an upper bound for the distance between the estimator $\mathbb{D}_{\rho,*}^{(n)}(\theta)$ and the true value $\mathcal{D}_\rho^{(g_0)}(\theta)$. For a sequence $\alpha_n \to \infty$ with $\alpha_n = o(\varkappa_n)$ we decompose the MSE into

$$\mathrm{MSE}_1^{(\rho)}(\varkappa_n,\alpha_n) := \mathbb{E}\Big[\big(\hat\theta_\rho^{(n)}(\varkappa_n) - \theta_0\big)^2 \mathbb{1}_{\left\{\mathbb{H}_{\rho,*}^{(n)}(1)\le\alpha_n\right\}}\Big],$$
$$\mathrm{MSE}_2^{(\rho)}(\varkappa_n,\alpha_n) := \mathbb{E}\Big[\big(\hat\theta_\rho^{(n)}(\varkappa_n) - \theta_0\big)^2 \mathbb{1}_{\left\{\mathbb{H}_{\rho,*}^{(n)}(1)>\alpha_n\right\}}\Big] \le \mathbb{P}\big(\mathbb{H}_{\rho,*}^{(n)}(1) > \alpha_n\big),$$

which can be considered as the MSE due to small and large estimation error.

Theorem 4.8. *Suppose that $\theta_0 < 1$, (4.13) and Assumption 2.3 with $g_1 = g_2 = 0$ are satisfied. Then for any sequence $\alpha_n \to \infty$ with $\alpha_n = o(\varkappa_n)$ we have*

$$K_1\Big(\frac{\varkappa_n}{\sqrt{n\Delta_n}}\Big)^{2/\varpi} \le \mathrm{MSE}_1^{(\rho)}(\varkappa_n,\alpha_n) \le K_2\Big(\frac{\varkappa_n}{\sqrt{n\Delta_n}}\Big)^{2/\varpi}$$
$$\mathrm{MSE}_2^{(\rho)}(\varkappa_n,\alpha_n) \le \mathbb{P}\big(\mathbb{H}_{\rho,*}^{(n)}(1) > \alpha_n\big),$$

for $n \in \mathbb{N}$ sufficiently large, where $K_1 = \big(\frac{1-\varphi}{c}\big)^{2/\varpi}$ and $K_2 = \big(\frac{1+\varphi}{c}\big)^{2/\varpi}$ for some $\varphi \in (0,1)$.

In the following we discuss the choice of the regularizing sequence $\varkappa_n$ for the estimator $\hat\theta_\rho^{(n)}$ in order to control the probability of over- and underestimation of the change point $\theta_0 \in (0,1)$. Let $\hat\theta_n^*$ be a preliminary consistent estimate of θ_0. For example, if (4.13) holds for some $\varpi > 0$, one can take $\hat\theta_n^* = \hat\theta_\rho^{(n)}(\varkappa_n)$ for a sequence $\varkappa_n \to \infty$ satisfying the assumptions of Theorem 4.7. In the sequel, let $B \in \mathbb{N}$ be some large number and let $(\xi^{(b)})_{b=1,\dots,B}$ denote independent sequences of random variables, $\xi^{(b)} := (\xi_j^{(b)})_{j\in\mathbb{N}}$, satisfying Assumption 3.6. We denote by $\hat{\mathbb{H}}_{\rho,*}^{(n,b)}$ the particular bootstrap statistics calculated with respect to the data and the bootstrap multipliers $\xi_1^{(b)},\dots,\xi_n^{(b)}$ from the b-th iteration, where

$$\hat{\mathbb{H}}_{\rho,*}^{(n)}(\theta) := \sup_{t\in\mathbb{R}} \sup_{0\le\zeta\le\theta'\le\theta} \big|\hat{\mathbb{H}}_\rho^{(n)}(\zeta,\theta',t)\big| \tag{4.15}$$

for $\theta \in [0,1]$. With these notations for $B, n \in \mathbb{N}$ and $0 < r \leq 1$ we define the following empirical distribution function

$$F_{n,B}^{(\rho,r)}(x) = \frac{1}{B}\sum_{i=1}^{B} \mathbb{1}_{\left\{\left(\hat{\mathbb{H}}_{\rho,*}^{(n,i)}(\hat{\theta}_n^*)\right)^r \leq x\right\}},$$

and we denote by $F_{n,B}^{(\rho,r)-}(y) := \inf\left\{x \in \mathbb{R} \;\middle|\; F_{n,B}^{(\rho,r)}(x) \geq y\right\}$ its pseudo-inverse. Then in the sense of the theorems below the optimal choice of the threshold is given by

$$\hat{\varkappa}_{n,B}^{(\alpha,\rho)}(r) := F_{n,B}^{(\rho,r)-}(1-\alpha), \tag{4.16}$$

for a confidence level $\alpha \in (0,1)$.

Theorem 4.9. *Let $0 < \alpha < 1$ and assume that Assumption 2.3 is satisfied with $g_1 = g_2 = 0$ and with $0 < \theta_0 < 1$ for θ_0 defined in (4.3). Suppose further that there exists some $t_0 \in \mathbb{R}$ with $N_{\rho^2}(g_0; \theta_0, t_0) > 0$. Then the limiting probability for underestimation of the change point θ_0 is bounded by α. Precisely,*

$$\limsup_{B\to\infty}\limsup_{n\to\infty}\mathbb{P}\Big(\hat{\theta}_\rho^{(n)}\big(\hat{\varkappa}_{n,B}^{(\alpha,\rho)}(1)\big) < \theta_0\Big) \leq \alpha.$$

Theorem 4.10. *Let Assumption 2.3 be satisfied with $g_1 = g_2 = 0$, let $0 < r < 1$ and for θ_0 defined in (4.3) let $0 < \theta_0 < 1$. Furthermore, suppose that (4.13) holds for some $\varpi, c > 0$ and that there exists a $t_0 \in \mathbb{R}$ satisfying $N_{\rho^2}(g_0; \theta_0, t_0) > 0$. Additionally, let the bootstrap multipliers be either bounded in absolute value or standard normal distributed. Then for each $K > \big(1/c\big)^{1/\varpi}$ and all sequences $(\alpha_n)_{n\in\mathbb{N}} \subset (0,1)$ with $\alpha_n \to 0$ and $(B_n)_{n\in\mathbb{N}} \subset \mathbb{N}$ with $B_n \to \infty$ such that $\alpha_n^2 B_n \to \infty$, $(n\Delta_n)^{\frac{1-r}{2r}}\alpha_n \to \infty$, $\alpha_n^{-1} n\Delta_n^{1+\tau} \to 0$ (with $\tau > 0$ from Assumption 2.3), we have*

$$\lim_{n\to\infty}\mathbb{P}\Big(\hat{\theta}_\rho^{(n)}\big(\hat{\varkappa}_{n,B_n}^{(\alpha_n,\rho)}(r)\big) > \theta_0 + K\varphi_n^*\Big) = 0, \tag{4.17}$$

where $\varphi_n^* = \big(\hat{\varkappa}_{n,B_n}^{(\alpha_n,\rho)}(r)/\sqrt{n\Delta_n}\big)^{1/\varpi} \xrightarrow{\mathbb{P}} 0$, while $\hat{\varkappa}_{n,B_n}^{(\alpha_n,\rho)}(r) \xrightarrow{\mathbb{P}} \infty$.

Theorem 4.10 is meaningless without the statement $\varphi_n^* \xrightarrow{\mathbb{P}} 0$. With the additional parameter $r \in (0,1)$ this assertion can be proved by using the assumptions $(n\Delta_n)^{\frac{1-r}{2r}}\alpha_n \to \infty$ and $\alpha_n^{-1} n\Delta_n^{1+\tau} \to 0$ only. However, it seems that for $r = 1$ the statement $\varphi_n^* \xrightarrow{\mathbb{P}} 0$ can only be verified under very restrictive conditions on the underlying process.

We conclude this section with an example which shows that the expansion (4.13) and the additional assumption $N_{\rho^2}(g_0; \theta_0, t_0) > 0$ of the preceding theorems are satisfied in the situations of Example 2.2 and Example 2.5. A proof for this example can be found in Section 6.4.

Example 4.11.

(1) *Recall the situation of an abrupt change considered in Example 4.2. In this case it follows from* (4.6) *that*

$$\mathcal{D}_\rho^{(g_0)}(\theta) = V_0^\rho \theta_0 \big(1 - \frac{\theta_0}{\theta}\big) = V_0^\rho(\theta - \theta_0) - \frac{V_0^\rho}{\theta}(\theta-\theta_0)^2 > 0,$$

whenever $\theta_0 < \theta \le 1$. *Consequently,* (4.13) *is satisfied with* $\varpi = 1$ *and* $\aleph(\theta) = -\frac{V_0^\rho}{\theta}(\theta-\theta_0)^2 = O((\theta-\theta_0)^2)$ *for* $\theta \to \theta_0$. *Moreover, if* $\nu_1 \neq 0$ *and the function* ρ *meets Assumption 2.3*(a3)*, the transition kernel given by* (4.5) *satisfies the additional assumption* $N_{\rho^2}(g_0;\theta_0,t_0) > 0$ *in Theorem 4.9 and Theorem 4.10 for some* $t_0 \in \mathbb{R}$.

(2) *In the situation discussed in Example 4.3 let*

$$\bar{N}(y,t) = \bar{A}(y)\int_{-\infty}^t \rho_{L,\hat{p}}(z) h_{\bar{\beta}(y),\bar{p}(y)}(z)dz$$

for $y \in U$ *and* $t \in \mathbb{R}$. *Then we have* $k_0 := \min\{k \in \mathbb{N} \mid \exists t \in \mathbb{R}\colon N_k(t) \neq 0\} < \infty$, *where for* $k \in \mathbb{N}_0$ *and* $t \in \mathbb{R}$

$$N_k(t) := \Big(\frac{\partial^k \bar{N}}{\partial y^k}\Big)\Big|_{(\theta_0,t)}$$

denotes the k*-th partial derivative of* $\bar{N}$ *with respect to* y *at* (θ_0,t)*, which is a bounded function on* $\mathbb{R}$. *Furthermore, there exists a* $\iota > 0$ *such that*

$$\mathcal{D}_{\rho_{L,\hat{p}}}^{(g_0)}(\theta) = \Big(\frac{1}{(k_0+1)!}\sup_{t\in\mathbb{R}}|N_{k_0}(t)|\Big)(\theta-\theta_0)^{k_0+1} + \aleph(\theta) \tag{4.18}$$

on $[\theta_0, \theta_0+\iota]$ *with* $|\aleph(\theta)| \le K(\theta-\theta_0)^{k_0+2}$ *for some* $K > 0$. *Obviously,* $N_{\rho^2_{L,\hat{p}}}(g_0;\theta_0,t_0) > 0$ *holds for some* $t_0 \in \mathbb{R}$.

4.4. Testing for a gradual change

In Section 3 we introduced change point tests for the situation of an abrupt change as in Example 2.2, where the jump behaviour is assumed to be constant before and after the change point. In this section we illustrate a reasonable way to derive test procedures for the existence of a gradual change in the data. In order to formulate suitable hypotheses for a gradual change point recall the definition of the measure of time variation for the entire jump behaviour $D_\rho^{(g_0)}$ in (4.1) and define for $t_0 \in \mathbb{R}$ and $\theta \in [0,1]$ the quantities

$$\mathcal{D}_\rho^{(g_0)}(\theta) := \sup_{t\in\mathbb{R}}\sup_{0\le\zeta\le\theta'\le\theta}\big|D_\rho^{(g_0)}(\zeta,\theta',t)\big|$$
$$\mathcal{D}_{\rho,t_0}^{(g_0)}(\theta) := \sup_{0\le\zeta\le\theta'\le\theta}\big|D_\rho^{(g_0)}(\zeta,\theta',t_0)\big|.$$

We test the null hypothesis

$\mathbf{H}_0$: Assumption 2.3 is satisfied with $g_1 = g_2 = 0$ and there exists a Lévy measure ν_0 such that $g_0(y, dz) = \nu_0(dz)$ holds for Lebesgue almost every $y \in [0, 1]$.

versus the general alternative of non-constant jump behaviour

$\mathbf{H}_1^*$: Assumption 2.3 holds with $g_1 = g_2 = 0$ and we have $\mathcal{D}_\rho^{(g_0)}(1) > 0$.

If one is interested in gradual changes in $N_\rho(\nu_s^{(n)}; t_0)$ for a fixed $t_0 \in \mathbb{R}$, one can consider the corresponding alternative

$\mathbf{H}_1^*(t_0)$: Assumption 2.3 is satisfied with $g_1 = g_2 = 0$ and we have $\mathcal{D}_{\rho,t_0}^{(g_0)}(1) > 0$.

Furthermore, we investigate the behaviour of the tests introduced below under local alternatives of the form

$\mathbf{H}_1^{(loc)}$: Assumption 2.3 holds with $g_0(y, dz) = \nu_0(dz)$ for Lebesgue-a.e. $y \in [0, 1]$ for some Lévy measure ν_0 and some transition kernels $g_1, g_2 \in \mathcal{G}(\beta, p)$.

Remark 4.12. *Note that the function $D_\rho^{(g_0)}$ in (4.1) is uniformly continuous in $(\zeta, \theta) \in C$ uniformly in $t \in \mathbb{R}$, that is for any $\eta > 0$ there exists a $\delta > 0$ such that*

$$\left|D_\rho^{(g_0)}(\zeta_1, \theta_1, t) - D_\rho^{(g_0)}(\zeta_2, \theta_2, t)\right| < \eta$$

holds for each $t \in \mathbb{R}$ and all pairs $(\zeta_1, \theta_1), (\zeta_2, \theta_2) \in C = \{(\zeta, \theta) \in [0, 1]^2 \mid \zeta \leq \theta\}$ with maximum distance $\|(\zeta_1, \theta_1) - (\zeta_2, \theta_2)\|_\infty < \delta$. Therefore, the function $D_\rho^*(g_0; \zeta, \theta) = \sup_{t \in \mathbb{R}} |D_\rho^{(g_0)}(\zeta, \theta, t)|$ is uniformly continuous on C and as a consequence $\mathcal{D}_\rho^{(g_0)}$ is continuous on $[0, 1]$. Thus, $\mathcal{D}_\rho^{(g_0)}(1) > 0$ holds if and only if the point θ_0 defined in (4.3) satisfies $\theta_0 < 1$.

The idea of the following tests is to reject the null hypothesis $\mathbf{H}_0$ for large values of the corresponding estimators $\mathbb{D}_{\rho,*}^{(n)}(1)$ and $\sup_{(\zeta,\theta) \in C} |\mathbb{D}_\rho^{(n)}(\zeta, \theta, t_0)|$ for $\mathcal{D}_\rho^{(g_0)}(1)$ and $\mathcal{D}_{\rho,t_0}^{(g_0)}(1)$, respectively. In order to obtain critical values we use the multiplier bootstrap approach introduced in Section 4.2. For this purpose we denote by $(\xi^{(b)})_{b=1,\ldots,B}$ for some large $B \in \mathbb{N}$ independent sequences $\xi^{(b)} = (\xi_j^{(b)})_{j \in \mathbb{N}}$ of multipliers satisfying Assumption 3.6. We denote by $\hat{\mathbb{H}}_\rho^{(n,b)}$ the processes defined in (4.12) calculated from $\{X_{i\Delta_n}^{(n)} \mid i = 0, \ldots, n\}$ and the b-th bootstrap multipliers $\xi_1^{(b)}, \ldots, \xi_n^{(b)}$. For a given level $\alpha \in (0, 1)$, we propose to reject $\mathbf{H}_0$ in favor of $\mathbf{H}_1^*$, if

$$(n\Delta_n)^{1/2} \mathbb{D}_{\rho,*}^{(n)}(1) \geq \hat{q}_{1-\alpha}^{(B)}\Big(\mathbb{H}_{\rho,*}^{(n)}(1)\Big), \tag{4.19}$$

where $\hat{q}_{1-\alpha}^{(B)}\big(\mathbb{H}_{\rho,*}^{(n)}(1)\big)$ denotes the $(1-\alpha)$-quantile of the sample $\hat{\mathbb{H}}_{\rho,*}^{(n,1)}(1), \ldots, \hat{\mathbb{H}}_{\rho,*}^{(n,B)}(1)$ with $\hat{\mathbb{H}}_{\rho,*}^{(n,b)}$ defined in (4.15). Similarly, for $t_0 \in \mathbb{R}$, the null hypothesis $\mathbf{H}_0$ is rejected in favor of $\mathbf{H}_1^*(t_0)$ if

$$R_{\rho,t_0}^{(n)} := (n\Delta_n)^{1/2} \sup_{(\zeta,\theta) \in C} \left|\mathbb{D}_\rho^{(n)}(\zeta, \theta, t_0)\right| \geq \hat{q}_{1-\alpha}^{(B)}\big(R_{\rho,t_0}^{(n)}\big), \tag{4.20}$$

where $\hat{q}_{1-\alpha}^{(B)}\big(R_{\rho,t_0}^{(n)}\big)$ denotes the $(1-\alpha)$-quantile of the sample $\hat{R}_{\rho,t_0}^{(n,1)},\dots,\hat{R}_{\rho,t_0}^{(n,B)}$, and

$$\hat{R}_{\rho,t_0}^{(n,b)} := \sup_{(\zeta,\theta)\in C} \big|\hat{\mathbb{H}}_\rho^{(n,b)}(\zeta,\theta,t_0)\big|.$$

In the following we show the behaviour of the aforementioned tests under $\mathbf{H}_0$, $\mathbf{H}_1^{(loc)}$ and the alternatves $\mathbf{H}_1^*$, $\mathbf{H}_1^*(t_0)$. To this end, recall the limit process $\mathbb{H}_{\rho,g_1} := \mathbb{H}_\rho + D_\rho^{(g_1)}$ in Theorem 4.4, where $D_\rho^{(g_1)}$ is defined in (4.1) and where the tight mean zero Gaussian process $\mathbb{H}_\rho$ in $\ell^\infty(C\times\mathbb{R})$ has the covariance function (4.11). Under the general Assumption 2.3 let $K_\rho : (\mathbb{R},\mathbb{B}) \to (\mathbb{R},\mathbb{B})$ be the c.d.f. of $\sup_{(\zeta,\theta,t)\in C\times\mathbb{R}} |\mathbb{H}_\rho(\zeta,\theta,t)|$ and let $K_\rho^{(t_0)} : (\mathbb{R},\mathbb{B}) \to (\mathbb{R},\mathbb{B})$ be the c.d.f. of $\sup_{(\zeta,\theta)\in C} |\mathbb{H}_\rho(\zeta,\theta,t_0)|$. Furthermore, let

$$\begin{aligned} H_{\rho,g_1} &:= \sup_{(\zeta,\theta,t)\in C\times\mathbb{R}} |\mathbb{H}_\rho(\zeta,\theta,t) + D_\rho^{(g_1)}(\zeta,\theta,t)|, \\ H_{\rho,g_1}^{(t_0)} &:= \sup_{(\zeta,\theta)\in C} |\mathbb{H}_\rho(\zeta,\theta,t_0) + D_\rho^{(g_1)}(\zeta,\theta,t_0)|. \end{aligned}$$

The proposition below shows the performance of the new tests under the local alternative $\mathbf{H}_1^{(loc)}$.

Proposition 4.13. *Under $\mathbf{H}_1^{(loc)}$ we have for each $\alpha\in(0,1)$*

$$\begin{aligned} &\mathbb{P}\big(K_\rho(H_{\rho,g_1}) > 1-\alpha\big) \le \liminf_{B\to\infty}\lim_{n\to\infty} \mathbb{P}\Big((n\Delta_n)^{1/2}\mathbb{D}_{\rho,*}^{(n)}(1) \ge \hat{q}_{1-\alpha}^{(B)}\big(\mathbb{H}_{\rho,*}^{(n)}(1)\big)\Big) \\ &\le \limsup_{B\to\infty}\lim_{n\to\infty} \mathbb{P}\Big((n\Delta_n)^{1/2}\mathbb{D}_{\rho,*}^{(n)}(1) \ge \hat{q}_{1-\alpha}^{(B)}\big(\mathbb{H}_{\rho,*}^{(n)}(1)\big)\Big) \le \mathbb{P}\big(K_\rho(H_{\rho,g_1}) \ge 1-\alpha\big), \end{aligned}$$

if there exist $\bar{t}\in\mathbb{R}$, $\bar{\zeta}\in(0,1)$ with $N_{\rho^2}(g_0;\bar{\zeta},\bar{t}) > 0$, and furthermore

$$\begin{aligned} &\mathbb{P}\big(K_\rho^{(t_0)}(H_{\rho,g_1}^{(t_0)}) > 1-\alpha\big) \le \liminf_{B\to\infty}\lim_{n\to\infty} \mathbb{P}\Big(R_{\rho,t_0}^{(n)} \ge \hat{q}_{1-\alpha}^{(B)}\big(R_{\rho,t_0}^{(n)}\big)\Big) \\ &\qquad \le \limsup_{B\to\infty}\lim_{n\to\infty} \mathbb{P}\Big(R_{\rho,t_0}^{(n)} \ge \hat{q}_{1-\alpha}^{(B)}\big(R_{\rho,t_0}^{(n)}\big)\Big) \le \mathbb{P}\big(K_\rho^{(t_0)}(H_{\rho,g_1}^{(t_0)}) \ge 1-\alpha\big) \end{aligned}$$

holds for each $\alpha\in(0,1)$, if there exists a $\bar{\zeta}\in(0,1)$ with $N_{\rho^2}(g_0;\bar{\zeta},t_0) > 0$.

With the result above and an inspection of the limiting probability $\mathbb{P}\big(K_\rho(H_{\rho,g_1}) \ge 1-\alpha\big)$, which is beyond the scope of this paper, one can show for which direction g_1 it is more difficult to distinguish the null hypothesis from the alternative. An immediate consequence of Proposition 4.13 is that the tests (4.19) and (4.20) hold the level α asymptotically.

Corollary 4.14. *The tests (4.19) and (4.20) are asymptotic level α tests in the following sense: Under $\mathbf{H}_0$ with $\nu_0 \neq 0$ we have for each $\alpha\in(0,1)$*

$$\lim_{B\to\infty}\lim_{n\to\infty} \mathbb{P}\Big((n\Delta_n)^{1/2}\mathbb{D}_{\rho,*}^{(n)}(1) \ge \hat{q}_{1-\alpha}^{(B)}\big(\mathbb{H}_{\rho,*}^{(n)}(1)\big)\Big) = \alpha$$

and moreover

$$\lim_{B\to\infty}\lim_{n\to\infty} \mathbb{P}\Big(R_{\rho,t_0}^{(n)} \ge \hat{q}_{1-\alpha}^{(B)}\big(R_{\rho,t_0}^{(n)}\big)\Big) = \alpha,$$

holds for all $\alpha\in(0,1)$, if $N_{\rho^2}(\nu_0;t_0) > 0$.

The tests (4.19) and (4.20) are also consistent under the fixed alternatives $\mathbf{H}_1^*$, $\mathbf{H}_1^*(t_0)$ in the sense of the following proposition.

Proposition 4.15. *Under $\mathbf{H}_1^*$, we have for all $B \in \mathbb{N}$*

$$\lim_{n\to\infty} \mathbb{P}\Big((n\Delta_n)^{1/2} \mathbb{D}_{\rho,*}^{(n)}(1) \geq \hat{q}_{1-\alpha}^{(B)}\big(\mathbb{H}_{\rho,*}^{(n)}(1)\big)\Big) = 1.$$

Under $\mathbf{H}_1^*(t_0)$, we have for all $B \in \mathbb{N}$

$$\lim_{n\to\infty} \mathbb{P}\Big(R_{\rho,t_0}^{(n)} \geq \hat{q}_{1-\alpha}^{(B)}\big(R_{\rho,t_0}^{(n)}\big)\Big) = 1.$$

5. Finite sample properties

In this section we present the results of a simulation study assessing the finite sample properties of the new statistical procedures. We divide this study into two parts: In Section 5.1 we investigate the performance of the new tests and estimators by means of a simulation study. Finally, we apply the new methods to high-frequency stock exchange prices in Section 5.2.

5.1. Monte Carlo experiments

This section is dedicated to a Monte Carlo simulation study. The design of this study is as follows:

(i) We apply our estimators and test statistics to n data points $\{X_{\Delta_n}, \ldots, X_{n\Delta_n}\}$ as realizations of an Itō semimartingale $(X_t)_{t\in\mathbb{R}_+}$ with characteristics (b, σ, ν_s). For the sample size we choose either $n = 10000$ or $n = 22500$, where for the effective sample size we consider the choices $k_n := n\Delta_n = 50, 100, 200$ in the case $n = 10000$ resulting in frequencies $\Delta_n^{-1} = 200, 100, 50$ and in the case $n = 22500$ we consider $k_n = n\Delta_n = 50, 75, 100, 150, 250$ resulting in $\Delta_n^{-1} = 450, 300, 225, 150, 90$.

(ii) Corresponding to our basic rescaling assumption (2.1) the jump characteristic satisfies

$$\nu_s(dz) = g\Big(\frac{s}{n\Delta_n}, dz\Big),$$

where the transition kernel $g(y, dz)$ is given by

$$g(y, [z,\infty)) = \begin{cases} \Big(\frac{\eta(y)}{\pi z}\Big)^{1/2} - \Big(\frac{1}{\pi 10^6}\Big)^{1/2}, & \text{if } 0 < z \leq \eta(y) 10^6, \\ 0, & \text{otherwise,} \end{cases} \tag{5.1}$$

and $g(y, (-\infty, z]) = 0$ for all $z < 0$.

(iii) In order to simulate data points $\{X_{\Delta_n}, \ldots, X_{n\Delta_n}\}$ including an abrupt change we choose

$$\eta(y) = \begin{cases} 1, & \text{if } y \le \theta_0, \\ \psi, & \text{if } y > \theta_0, \end{cases} \qquad (y \in [0,1]) \tag{5.2}$$

for $\theta_0 \in (0,1)$, $\psi \ge 1$ and we use a modification of Algorithm 6.13 in Cont and Tankov (2004) to simulate pure jump Itō semimartingales under $\mathbf{H}_0$, i.e. for $\psi = 1$. Under the alternative of an abrupt change, i.e. for $\psi > 1$, we merge two paths of independent semimartingales together.

(iv) A gradual change in the jump characteristic is realized by choosing

$$\eta(y) = \begin{cases} 1, & \text{if } y \le \theta_0, \\ (A(y-\theta_0)^w + 1)^2, & \text{if } y \ge \theta_0, \end{cases} \qquad (y \in [0,1]) \tag{5.3}$$

in (5.1) for some $\theta_0 \in [0,1]$, $A > 0$ and $w > 0$. In order to obtain pure jump Itō semimartingale data according to this model we sample 15 times more frequently, i.e. for $j \in \{1, \dots, 15n\}$ we use a modification of Algorithm 6.13 in Cont and Tankov (2004) to simulate an increment $Z_j = \tilde{X}^{(j)}_{j\Delta_n/15} - \tilde{X}^{(j)}_{(j-1)\Delta_n/15}$ of a 1/2-stable pure jump Lévy subordinator with characteristic exponent

$$\Phi^{(j)}(u) = \int (e^{iuz} - 1)\nu^{(j)}(dz),$$

where $\nu^{(j)}(dz) = g(j/(15n), dz)$. For the resulting data vector $\{X_{\Delta_n}, \dots, X_{n\Delta_n}\}$ we use

$$X_{k\Delta_n} = \sum_{j=1}^{15k} Z_j, \quad (k = 1, \dots, n).$$

(v) In order to investigate the performance of our truncation method we either use the plain pure jump data vector $\{X_{\Delta_n}, \dots, X_{n\Delta_n}\}$ as described above, resulting in the characteristics $b = \sigma = 0$ for the continuous part, or we use $\{X_{\Delta_n} + S_{\Delta_n}, \dots, X_{n\Delta_n} + S_{n\Delta_n}\}$, where $S_t = W_t + t$ with a Brownian motion $(W_t)_{t \in \mathbb{R}_+}$ resulting in $b = \sigma = 1$. In the graphics depicted below the results for pure jump data are presented on the left-hand side, while the results including a continuous component are always placed on the right-hand side.

(vi) For the truncation sequence $v_n = \gamma \Delta_n^{\overline{w}}$ we choose $\gamma = 1$ and $\overline{w} = 3/4$ in each run resulting in the parameter $\tau = 2/15$ in Assumption 2.3.

(vii) Due to computational reasons we approximate the supremum in $t \in \mathbb{R}$ by taking the maximum either over the finite grid $T_1 := \{0.1 \cdot j \mid j = 1, \dots, 30\}$ or the finite grid $T_2 := \{0.1 + j \cdot 0.3 \mid j = 0, 1, \dots, 9\}$.

(viii) For the function ρ we use $\rho_{L,p}$ from (2.8) in Example 2.5 with parameters $L = 1$ and $p = 2$.

(ix) Each combination of parameters we present below is run 500 times and if the statistical procedure includes a bootstrap method we always use $B = 200$ bootstrap replications. In order to illustrate the power of our test procedures we

Table 1
Simulated rejection probabilities of the tests (3.9), (3.10) and (3.11) under the null hypothesis. Upper part: pure jump subordinator data. Lower part: jump subordinator data plus a Brownian motion with drift.

k_n	Test (3.9)	Pointwise Tests	$t_0 = 0.5$	$t_0 = 1$	$t_0 = 1.5$	$t_0 = 2$	$t_0 = 2.5$	$t_0 = 3$
50	0.026	(3.10)	0.062	0.036	0.024	0.036	0.026	0.036
		(3.11)	0.060	0.042	0.030	0.030	0.016	0.020
75	0.052	(3.10)	0.058	0.048	0.046	0.040	0.046	0.050
		(3.11)	0.040	0.046	0.032	0.036	0.028	0.030
100	0.050	(3.10)	0.046	0.054	0.042	0.046	0.038	0.042
		(3.11)	0.038	0.038	0.036	0.040	0.028	0.032
150	0.068	(3.10)	0.038	0.054	0.054	0.054	0.058	0.066
		(3.11)	0.036	0.036	0.050	0.042	0.052	0.044
250	0.060	(3.10)	0.068	0.056	0.056	0.058	0.064	0.060
		(3.11)	0.046	0.034	0.034	0.032	0.044	0.052
50	0.040	(3.10)	0.038	0.042	0.036	0.054	0.034	0.036
		(3.11)	0.036	0.030	0.028	0.042	0.026	0.028
75	0.058	(3.10)	0.024	0.050	0.030	0.048	0.058	0.050
		(3.11)	0.030	0.032	0.020	0.042	0.046	0.036
100	0.050	(3.10)	0.044	0.050	0.040	0.046	0.048	0.052
		(3.11)	0.034	0.040	0.026	0.046	0.040	0.048
150	0.054	(3.10)	0.040	0.050	0.048	0.056	0.048	0.060
		(3.11)	0.040	0.032	0.038	0.038	0.030	0.038
250	0.060	(3.10)	0.046	0.058	0.036	0.056	0.062	0.058
		(3.11)	0.036	0.050	0.030	0.044	0.054	0.046

display simulated rejection probabilities, i.e. the mean of the 500 test results. Furthermore, we measure the performance of our estimators by mean absolute deviation, i.e. if $\Theta = \{\hat{\theta}_1, \ldots, \hat{\theta}_{500}\}$ is the set of obtained estimation results we depict

$$\ell^1(\Theta, \theta_0) = \frac{1}{500} \sum_{j=1}^{500} |\hat{\theta}_j - \theta_0|,$$

where θ_0 is the location of the change point.

5.1.1. Statistical inference for abrupt changes

To illustrate the finite sample performance of the procedures introduced in Section 3 we choose the sample size $n = 22500$ and the grid $T_1 = \{0.1 \cdot j \mid j = 1, \ldots, 30\}$ to approximate the supremum in $t \in \mathbb{R}$. The confidence level of the test procedures is $\alpha = 5\%$ in each run.

In Table 1 we display the rejection probabilities of the tests (3.9), (3.10) and (3.11) under the null hypothesis. We observe a reasonable approximation of the nominal level $\alpha = 0.05$. The test (3.11) appears to be slightly more conservative than the test (3.10). Note that the process investigated in the lower part of Table 1 includes a continuous component with $b = \sigma = 1$.

In the upper part of Figure 1 we depict the rejection probabilities of the test (3.9) for different effective sample sizes $k_n = n\Delta_n$. The factor of jump size corresponds to ψ in (5.2) and the dashed red line indicates the nominal level $\alpha = 5\%$. The change point is located at $\theta_0 = 0.5$. Large differences of the jump size before and after the change yield higher rejection probabilities. Moreover, the rejection probabilities increase with $k_n = n\Delta_n$. Notice also that

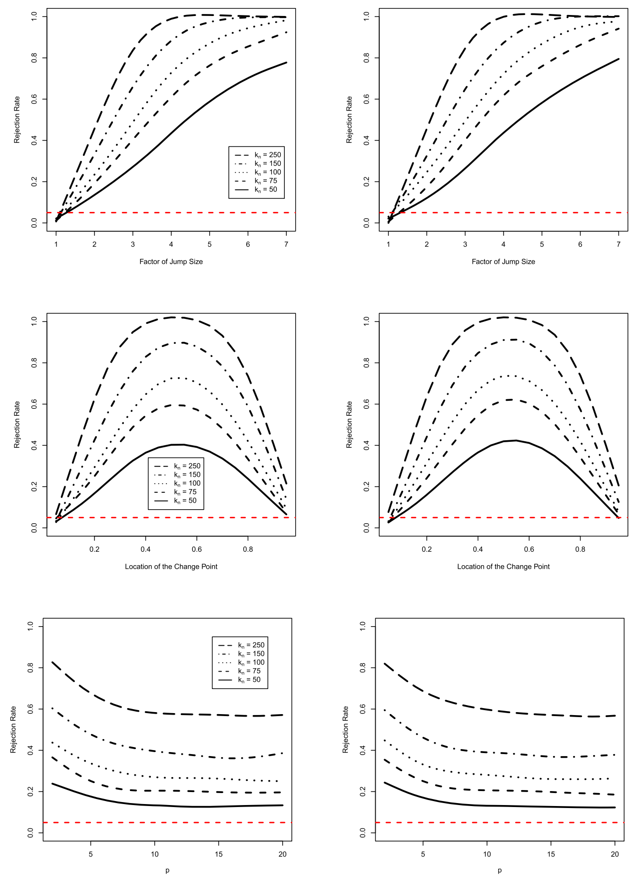

FIG 1. *Simulated rejection probabilities of the test* (3.9). *Upper part: different factors of jump size* ψ *in* (5.2) *(location of change point fixed at* $\theta_0 = 0.5$*). Middle part: different locations of the change point* θ_0 *(*$\psi = 4$ *fixed). Lower part: different values of the parameter* $p \geq 2$ *in the function* $\rho_{1,p}$ *(*$\theta_0 = 0.5$*,* $\psi = 3$ *fixed). Left panels: Pure jump data. Right panels: pure jump data plus a Brownian motion with drift. The dashed red line indicates* $\alpha = 5\%$*.*

the results for pure jump Itō semimartingales and for data including a continuous component are very similar. This fact indicates a reasonable performance of the proposed truncation technique for an ordinary sample size $n = 22500$. The middle part of Figure 1 shows the rejection probabilities for varying locations of the change point θ_0, where $\psi = 4$ in (5.2). Our results illustrate that an abrupt change can be detected best, if it is located close to $\theta_0 \approx 0.5$. Furthermore, in this case the power of the test is increasing with $k_n = n\Delta_n$ and the performance for data including a continuous component is nearly the same. In the lower part of Figure 1 we display the rejection probabilities for different values of the parameter $p \in [2, 20]$ of the function $\rho_{1,p}$ in (2.8), which is used to calculate the process $\mathbb{T}^{(n)}_{\rho_{1,p}}(\theta, t)$. Here the change point is located at $\theta_0 = 0.5$ and we

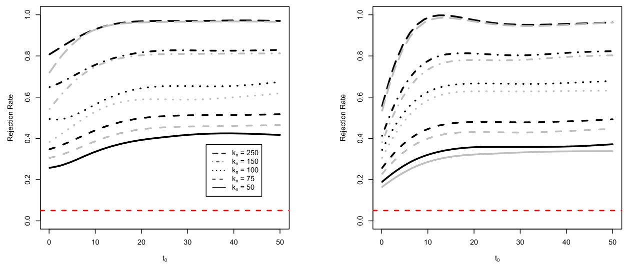

FIG 2. *Simulated rejection probabilities of the test* (3.10) *(black lines) and the test* (3.11) *(grey lines) for different values* t_0 *for pure jump data (left-hand side) and with an additional Brownian motion with drift (right-hand side). The dashed red line indicates the nominal level* $\alpha = 5\%$.

choose $\psi = 3$ in (5.2). The results suggest to use the lowest possible value of the parameter p in order to obtain the maximum power of the test. Again, the rejection probabilities of the test are nearly unaffected by the presence of a Brownian component.

In Figure 2 we depict rejection probabilities of the tests (3.10) and (3.11) for different values of $t_0 \in [0.1, 50]$. In the underlying model (5.1) we use $\eta(y)$ defined in (5.2) with $\theta_0 = 0.5$ and $\psi = 3$. We observe that the test (3.10) has slightly more power than the test (3.11) and the power of both tests is increasing for small values of t_0. The latter can be explained by the fact that less increments of the underlying Itō semimartingale which take values in the interval $(v_n, t_0]$ are used to calculate the test statistics. The effect is even more significant when a Brownian component is present (right panel). In this case it is more difficult to detect a change, because of the superposition of small increments with an i.i.d. sequence of random variables following a normal distribution with variance Δ_n (see also Figure 3 in Bücher et al. (2017)). Furthermore, one can show (see, for instance, Lemma 6.3 in Hoffmann et al. (2017)) that in the case of a pure jump Itō semimartingale the probability of the event that m increments exceed the value t_0 is bounded by $Kt_0^{-m/2}$. As a consequence, for large t_0 the power of both tests reaches a saturation, because only a negligible proportion of increments exceed t_0.

We conclude this section with a brief investigation of the argmax-estimator (3.17). In the upper part of Figure 3 we display mean absolute deviations of the estimator (3.17) for different values $\psi \in [1, 5]$ in (5.2) ($\theta_0 = 0.5$ fixed), and in the lower part we consider different locations of the change point $\theta_0 \in (0, 1)$ ($\psi = 3$ fixed). The results in the upper part correspond to Figure 1 in the sense that large values of ψ yield a better performance of the statistical procedure. Additionally, we also observe that due to the truncation approach the mean absolute deviation is nearly unaffected by the presence of a Brownian component. Similar to the middle part of Figure 1 the results in the lower part suggest that a change point can be detected best if it is located at $\theta_0 \approx 0.5$. Note also that the estimation error is decreasing with the effective sample size k_n.

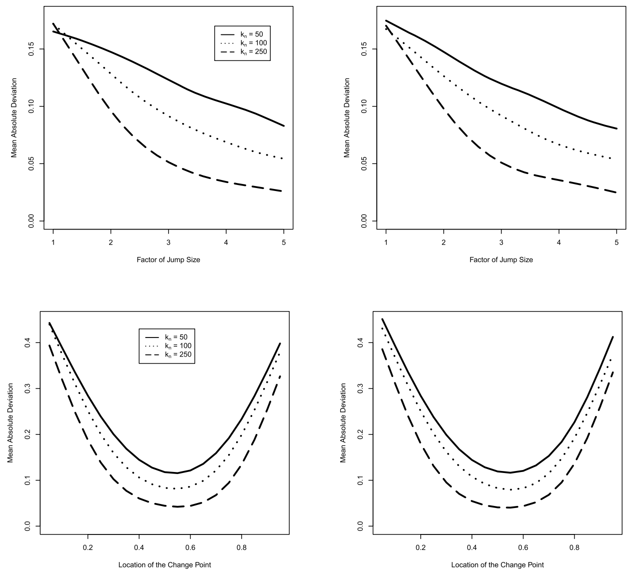

FIG 3. *Mean absolute deviation of the estimator* (3.17). *Upper part: different values of ψ in* (5.2)*, $\theta_0 = 0.5$ fixed. Lower part: different locations of $\theta_0 \in (0,1)$, $\psi = 3$ fixed. Left panels: pure jump data. Right panels: pure jump data plus an additional Brownian motion with drift.*

5.1.2. Statistical inference for gradual changes

In this section we investigate the finite sample performance of the statistical procedures introduced in Section 4.

In Table 2 we show the simulated rejection probabilities of the tests (4.19) and (4.20) under the null hypothesis, i.e. for $\psi = 1$ in (5.2). The sample size is $n = 22500$ and for the test (4.19) we approximate the supremum in $t \in \mathbb{R}$ by taking the maximum over the finite grid $T_1 = \{0.1 \cdot j \mid j = 1, \ldots, 30\}$. In both cases for pure jump Itō semimartingales ($b = \sigma = 0$) and for Itō semimartingales including a Brownian component ($b = \sigma = 1$) we observe a reasonable approximation of the nominal level $\alpha = 5\%$.

To save computational time the rejection probabilities of the tests (4.19) and (4.20) under the alternative are obtained for the sample size $n = 10000$ and effective sample size $k_n \in \{50, 100, 200\}$. The upper part of Figure 4 shows the simulated rejection probabilities of the test (4.19) for different degrees of smoothness of the change w in (5.3). The change is located at $\theta_0 = 0.4$ and A is chosen such that the characteristic quantity for a gradual change satisfies $\mathcal{D}_\rho^{(g)}(1) = 3$ in each scenario. As expected, it is more difficult to distinguish a very smooth change from the null hypothesis and therefore the rejection probability is decreasing in w. Similar to the CUSUM test investigated in Section 5.1.1 the power of the test is increasing with $k_n = n\Delta_n$. In the lower part of Figure 4 we depict the rejection rates of the test (4.19) for different locations of the change

TABLE 2

Simulated rejection probabilities of the tests (4.19) and (4.20) under the null hypothesis. Upper part: pure jump Itō semimartingale data. Lower part: pure jump Itō semimartingale data plus a Brownian motion with drift.

	Test (4.19)	Test (4.20)					
k_n	T_1	$t_0 = 0.5$	$t_0 = 1$	$t_0 = 1.5$	$t_0 = 2$	$t_0 = 2.5$	$t_0 = 3$
50	0.050	0.028	0.020	0.030	0.040	0.056	0.046
75	0.048	0.058	0.058	0.048	0.048	0.048	0.044
100	0.056	0.062	0.038	0.046	0.038	0.046	0.060
150	0.076	0.056	0.062	0.066	0.054	0.062	0.078
250	0.062	0.070	0.070	0.058	0.056	0.054	0.066
50	0.044	0.036	0.026	0.028	0.044	0.040	0.040
75	0.042	0.050	0.054	0.042	0.044	0.038	0.044
100	0.074	0.040	0.038	0.036	0.046	0.062	0.068
150	0.044	0.036	0.056	0.058	0.052	0.042	0.044
250	0.050	0.034	0.042	0.056	0.062	0.062	0.058

point $\theta_0 \in (0,1)$. We simulate a linear change, i.e. we have $w = 1$ in (5.3), and A is chosen such that $\mathcal{D}_\rho^{(g)}(1) = 0.3$ holds in each run. As before, the power of the test is increasing in the effective sample size $k_n = n\Delta_n$ and moreover it is decreasing in θ_0. The latter observation can be explained by the shape of model (5.3), because for larger values of θ_0 the jump characteristic is "closer" to the null hypothesis.

Note also that all results are very similar for pure jump processes and processes including a Brownian component. This indicates that our truncation approach also works in this setup.

We conclude this section with a study of the change point estimator $\hat{\theta}_\rho^{(n)}$ in (4.14). Following Hoffmann et al. (2017) we implement the estimator $\hat{\theta}_\rho^{(n)}$ in five steps as follows:

Step 1. Choose a preliminary estimate $\hat{\theta}^{(pr)} \in (0,1)$, a probability level $\alpha \in (0,1)$ and a parameter
$r \in (0,1]$.

Step 2. Initial choice of the tuning parameter $\varkappa_n$: Evaluate (4.16) for $\hat{\theta}^{(pr)}, \alpha$ and r (with $B = 200$
as described above and where the supremum in $t \in \mathbb{R}$ is approximated by the maximum
over $t \in T_2 = \{0.1 + j \cdot 0.3 \mid j = 0, 1, \ldots, 9\}$) and obtain $\hat{\varkappa}^{(in)}$.

Step 3. Intermediate estimate of the change point. Evaluate (4.14) for $\hat{\varkappa}^{(in)}$ and obtain $\hat{\theta}^{(in)}$.

Step 4. Final choice of the tuning parameter $\varkappa_n$: Evaluate (4.16) for $\hat{\theta}^{(in)}, \alpha, r$ and obtain $\hat{\varkappa}^{(fi)}$.

Step 5. Estimate θ_0. Evaluate (4.14) for $\hat{\varkappa}^{(fi)}$ and obtain the final estimate $\hat{\theta}$ of the change point.

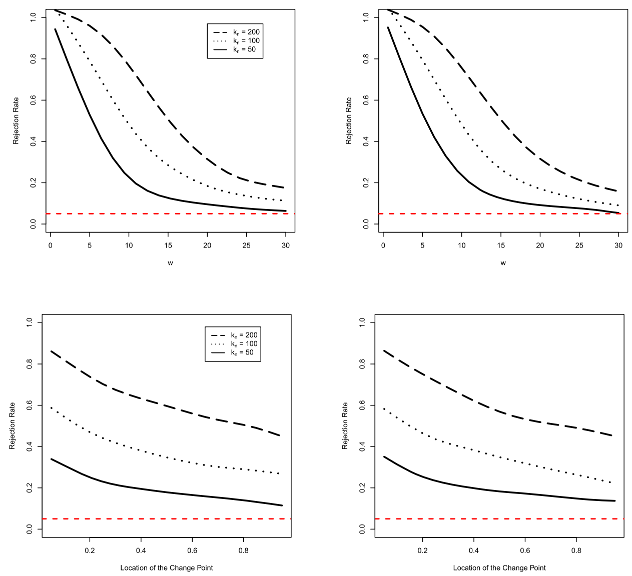

Fig 4. *Simulated rejection probabilities of the test* (4.19) *under the alternative. Upper part: different values* $w \in [0.6, 30]$ *in* (5.3), $\theta_0 = 0.4$ *fixed. Lower part: different locations of the change point* $\theta_0 \in (0,1)$, $w = 1$ *fixed. Left panels: pure jump Itō semimartingales; Right panels: pure jump Itō semimartingales plus a Brownian motion with drift. The dashed red line indicates the nominal level* $\alpha = 5\%$.

From the theoretical point of view as discussed in Section 4.3 we have to ensure that the preliminary estimate $\hat{\theta}^{(pr)}$ in Step 1 is consistent in order to guarantee consistency of the final estimate $\hat{\theta}$. If not mentioned otherwise, we always make the "arbitrary" choice $\hat{\theta}^{(pr)} = 0.1$ for two reasons: First, a simulation study which is not included in this paper, where the estimation procedure is started in Step 2 with the choice $\hat{\varkappa}^{(in)} = \sqrt[3]{n\Delta_n}$ (which yields consistency according to Theorem 4.7) has shown similar results as the ones depicted below. Secondly, with the small choice of $\hat{\theta}^{(pr)} = 0.1$ in Step 1 we obtain smaller values of the thresholds $\hat{\varkappa}^{(in)}$, $\hat{\varkappa}^{(fi)}$ and this reduces the calculation time. Furthermore, in the following simulation study we choose for the sample size $n = 22500$ and vary the effective sample size $k_n = n\Delta_n$ in $\{50, 100, 250\}$. For the evaluation of (4.16) we always use $\alpha = 10\%$ and for computational reasons suprema in $t \in \mathbb{R}$ are approximated by maxima over $t \in T_2 = \{0.1 + j \cdot 0.3 \mid j = 0, 1, \dots, 9\}$. If not mentioned otherwise, we simulate a linear change, i.e. $w = 1$ in (5.3), which is located at $\theta_0 = 0.4$. A is always chosen such that the characteristic quantity for a gradual change satisfies $\mathcal{D}_\rho^{(g)}(1) = 3$ in all scenarios.

The upper part of Figure 5 shows the mean absolute deviation of the estimator (4.14) for different choices of $r \in (0, 1]$ in Step 1. We observe that in all cases the mean absolute deviation for $r = 0.3$ is close to its overall minimum. Thus, we choose $r = 0.3$ in Step 1 in all following investigations. In the lower part we display the mean absolute deviation for different choices of the pre-

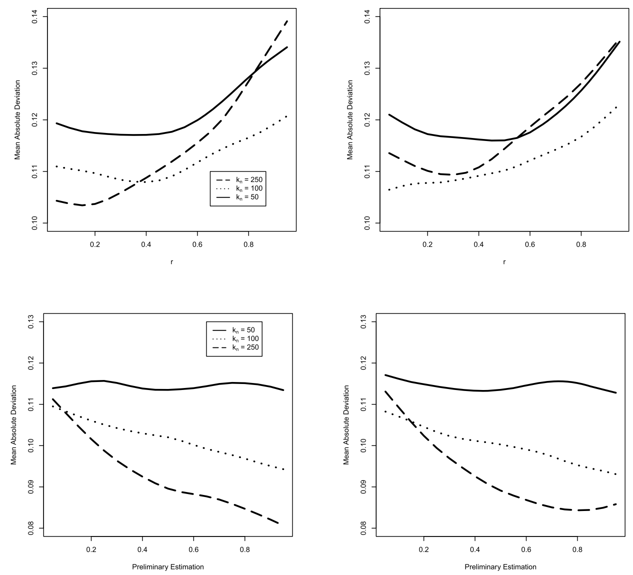

FIG 5. *Mean absolute deviation of the estimator* (4.14). *Upper part: different choices of* $r \in (0, 1]$ *in Step 1. Lower part: different choices of the preliminary estimate* $\hat{\theta}^{(pr)} \in (0, 1)$ *in Step 1. Left panels: pure jump Itō semimartingales. Right panels: pure jump Itō semimartingales plus an additional Brownian component.*

liminary estimate $\hat{\theta}^{(pr)} \in (0, 1)$ in Step 1. The smallest error is obtained, if the preliminary estimate is chosen close to 1. These findings were confirmed by a further simulation study which is not presented here and demonstrates that the procedure (4.14) tends to underestimate the change point. As a consequence, $\hat{\theta}^{(pr)}$ close to 1 induces larger values of the quantities $\hat{\varkappa}^{(in)}, \hat{\theta}^{(in)}, \hat{\varkappa}^{(fi)}$ in Steps 2-4 and prevents the underestimation error.

The upper part of Figure 6 shows the simulated mean absolute deviation of the estimator (4.14) for different degrees of smoothness of the change w in (5.3). The results correspond to the upper part of Figure 4 and confirm the intuitive idea that a smooth change is more difficult to detect. Moreover, larger effective sample sizes $k_n = n\Delta_n$ reduce the estimation error. In the lower part we display the simulated mean absolute deviation of the estimator $\hat{\theta}_\rho^{(n)}$ for different locations of the change point $\theta_0 \in (0, 1)$ in (5.3). The results correspond to lower part of Figure 4 and show that for small values of θ_0 the change point can be detected best. This is a consequence of model (5.3), where for larger values of $\theta_0 \in (0, 1)$ the jump behaviour is nearly constant.

5.2. Real data application

In this section we show the results of an application of the new methods to mid price data (in US dollar) of Apple shares between 09:30 and 16:00 on 21-

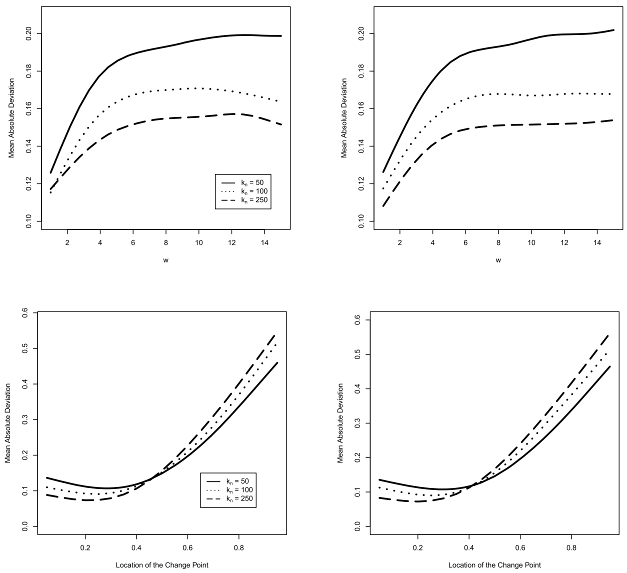

FIG 6. *Mean absolute deviation of the estimator* (4.14). *Upper part: different degrees of smoothness of the change w in* (5.3). *Lower part: different locations of the change point. Left panels: pure jump processes. Right panels: pure jump processes plus an additional additional Brownian motion with drift.*

06-2012, which is depicted in Figure 7 and consists of $n = 106626$ data points. We choose $k_n = n\Delta_n = 23400$, which corresponds to the number of seconds between 09:30 and 16:00. Furthermore, we use again the function $\rho_{L,p}$ from (2.8) in Example 2.5 with parameters $L = 1$ and $p = 2$. For the truncation sequence we choose $v_n = \gamma(k_n/n)^{3/4}$, where we use $\gamma = 0.005$ to address the fact that the increments $\Delta_i^n X$ in the data are very small, due to the extremely high frequency of sampling. For the same reason we approximate the supremum in $t \in \mathbb{R}$ in the methods from Section 3 by the maximum over the finite grid $\{j \cdot 0.0005 \mid j \in \{1, \dots, 80\}\}$, while the supremum in $t \in \mathbb{R}$ in the methods from Section 4 is approximated by the maximum over the finite grid $\{j \cdot 0.004 \mid j \in \{1, \dots, 10\}\}$. As in Section 5.1 we use 200 bootstrap replications whenever a procedure requires resampling. Due to the huge sample size we use $r = 0.8$ in order to reduce the calculation time.

The test (3.9) and the test (4.19) reject the null hypothesis of no change in the jump behaviour (level $\alpha = 5\%$). In order to locate the abrupt change point an application of the argmax-estimator (3.17) yields the value 38029 seconds after midnight for the time of an abrupt change. Similarly, the 5-step-procedure for the estimator (4.14) introduced in Section 5.1.2 gives 38131 seconds after midnight for the time of a gradual change (here we also choose $\hat{\theta}^{(pr)} = 0.1$ and $\alpha = 0.1$ in step 1). We also applied the standard binary segmentation algorithm to detect further change points [see Vostrikova (1981)], where we

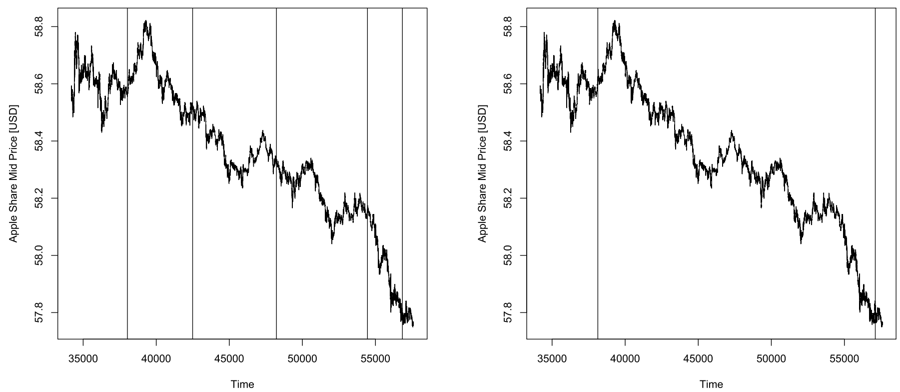

FIG 7. *Mid prices of Apple shares in US dollar between 09:30 and 16:00 on 21-06-2012. The time is measured on the x-axis by seconds after midnight. The solid vertical lines show the results of the argmax-estimator* (3.17) *(left panel) and the estimator* (4.14) *(right panel).*

restrict ourselves to a maximal number of 5 change points. The results are depicted by the solid vertical lines in Figure 7, where the left and right panel correspond to the detection of abrupt and gradual changes, respectively.

Remark 5.1. An important and extremely difficult problem in finite samples is to distinguish between abrupt and gradual changes. In the context of change point detection in a sequence of means it was demonstrated in Section 7.1 of Vogt and Dette (2015) that the procedure for detecting gradual changes is also able to detect abrupt changes in the mean. Similarly, the classical test for an abrupt change detects also a gradual one. However, a test applied under a wrong assumption might be less powerful, even if it is consistent. In the context for changes in the jump behaviour a similar observation can be made. Both tests detect abrupt and gradual changes, but a test constructed for a specific form of the change might be less efficient if the form of the change deviates from the postulated model.

On the other hand, even in the context of detecting a change in a sequence of means, a formal statistical test to distinguish between an abrupt and gradual change is to our best knowledge not available. Problems of this type are very challenging topics for future research.

6. Proofs

The proofs of the results in this paper are technically very demanding and we decompose the arguments in several parts. The main steps are given in this section. We begin stating general assumptions in Section 6.1 which are sufficient for all results presented in this paper and implied by the more readable assumptions made in Section 2. In Section 6.2 we state results regarding the weak convergence of two empirical processes, which are used in the definition of the statistics considered in Section 3 and 4. Proofs for the results in these sections can be found in Section 6.3 and 6.4. All arguments presented here rely on several technical auxiliary results, which can be found in Appendix A–E of the supplement (Hoffmann and Dette (2019)).

6.1. Alternative assumptions

All results in this paper also hold under the weaker assumptions given below. Here and throughout this section K or $K(\delta)$ denote generic constants depending in some cases on a quantity δ and may change from place to place.

Assumption 6.1. *At step $n \in \mathbb{N}$ we observe an Itō semimartingale $X^{(n)}$ adapted to the filtration of some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P})$ with characteristics $(b_s^{(n)}, \sigma_s^{(n)}, \nu_s^{(n)})$ at the equidistant time points $\{i\Delta_n \mid i = 0, 1, \dots, n\}$. Furthermore, the following assumptions are satisfied:*

(a) Assumptions on the jump characteristic and the function ρ:
For each $n \in \mathbb{N}$ and $s \in [0, n\Delta_n]$ we have

$$\nu_s^{(n)}(dz) = g^{(n)}\Big(\frac{s}{n\Delta_n}, dz\Big), \tag{6.1}$$

where there exist transition kernels g_0, g_1, g_2 from $([0,1], \mathbb{B}([0,1]))$ into $(\mathbb{R}, \mathbb{B})$ such that for each $y \in [0,1]$

$$g^{(n)}(y, dz) = g_0(y, dz) + \frac{1}{\sqrt{n\Delta_n}} g_1(y, dz) + \mathcal{R}_n(y, dz) \tag{6.2}$$

and for each $y \in [0,1]$, $B \in \mathbb{B}$ and $n \in \mathbb{N}$ the kernel $\mathcal{R}_n$ satisfies $\mathcal{R}_n(y, B) \leq a_n g_2(y, B)$ for a sequence $a_n = o((n\Delta_n)^{-1/2})$ of non-negative real numbers. Furthermore, we have

(1) There exists $\beta \in [0,2]$ with

$$\max_{i=0,1,2} \Big(\lambda_1 - \text{ess sup}_{y\in[0,1]} \Big(\int \big(1 \wedge |z|^{(\beta+\delta)\wedge 2}\big) g_i(y, dz)\Big)\Big) \leq K(\delta) < \infty$$

for each $\delta > 0$.

(2) $\rho\colon \mathbb{R} \to \mathbb{R}$ is a bounded $\mathcal{C}^1$-function with $\rho(0) = 0$. Furthermore, there exists some $p > \beta + (\beta \vee 1)$ such that the derivative satisfies $|\rho'(z)| \leq K|z|^{p-1}$ for all $z \in \mathbb{R}$ and some $K > 0$.

(3) For $\overline{p} = (p-1) \vee 1$ with p from (a2) we have

$$\max_{i=0,1,2} \Big(\lambda_1 - \text{ess sup}_{y\in[0,1]} \Big(\int |z|^{\overline{p}} \mathbb{1}_{\{|z|\geq 1\}} g_i(y, dz)\Big)\Big) < \infty.$$

(4) *(I) There exist $\overline{r} > \overline{v} > 0$, $\alpha_0 > 0$, $q > 0$ and $K > 0$ such that for every choice $m_1, m_2 \in \{g_0, g_1, g_2\}$*

$$\lambda_2 - \text{ess sup}_{y_1,y_2 \in [0,1]} \Big(\int \int \mathbb{1}_{\{|x-z| \le \Delta_n^{\overline{r}}\}} \mathbb{1}_{\{\Delta_n^{\overline{v}}/2 < |x| \le \alpha_0\}} \times \\ \times \mathbb{1}_{\{\Delta_n^{\overline{v}}/2 < |z| \le \alpha_0\}} m_1(y_1, dx) m_2(y_2, dz) \Big) \le K \Delta_n^q,$$

holds for $n \in \mathbb{N}$ sufficiently large, where λ_2 denotes the restriction of the two-dimensional Lebesgue measure to the measure space $([0,1]^2, [0,1]^2 \cap \mathcal{L}_2)$ with the two-dimensional Lebesgue σ-algebra $\mathcal{L}_2$ on $\mathbb{R}^2$.

(II) For each $\alpha > 0$ there is a $K(\alpha) > 0$ such that for every choice $m_1, m_2 \in \{g_0, g_1, g_2\}$ we have

$$\lambda_2 - \text{ess sup}_{y_1,y_2 \in [0,1]} \Big(\int \int \mathbb{1}_{\{|x-z| \le \Delta_n^{\overline{r}}\}} \mathbb{1}_{\{|x| > \alpha\}} \times \\ \times \mathbb{1}_{\{|z| > \alpha\}} m_1(y_1, dx) m_2(y_2, dz) \Big) \le K(\alpha) \Delta_n^q,$$

for $n \in \mathbb{N}$ large enough with the constants from (a(4)I).

(b) Assumptions on the truncation sequence v_n and the observation scheme: We have $v_n = \gamma \Delta_n^{\overline{w}}$ for some $\gamma > 0$ and $\overline{w}$ satisfying $\frac{1}{2(p-\beta)} < \overline{w} < \frac{1}{2} \wedge \frac{1}{2\beta}$. Furthermore, the observation scheme satisfies with the constants from the previous assumptions:

(1) $\Delta_n \to 0$,

(2) $n\Delta_n \to \infty$,

(3) $n\Delta_n^{1+q/2} \to 0$,

(4) $n\Delta_n^{1+2\overline{w}} \to 0$,

(5) $n\Delta_n^{1+2\overline{v}(p-\beta-\delta)} \to 0$ *for some* $\delta > 0$,

(6) $n\Delta_n^{2(1-\beta\overline{w}(1+\epsilon))} \to 0$ *for some* $\epsilon > 0$,

(7) $n\Delta_n^{((1+2(\overline{r}-\overline{w}))\vee 1)+\delta} \to \infty$ *for some* $\delta > 0$.

(c) Assumptions on the drift and the diffusion coefficient: For $m_b = \frac{1+2\overline{w}}{1-\overline{w}} \le 4$ and $m_\sigma = \frac{1+2\overline{w}}{1/2-\overline{w}}$, we have

$$\sup_{n \in \mathbb{N}} \sup_{s \in \mathbb{R}_+} \left\{ \mathbb{E}\big|b_s^{(n)}\big|^{m_b} \vee \mathbb{E}\big|\sigma_s^{(n)}\big|^{m_\sigma} \right\} < \infty.$$

Throughout the following proofs we will work with Assumption 6.1 without further mention. This is due to the following result which proves that Assumption 2.3 implies the set of conditions above.

Proposition 6.2. *Assumption 2.3 is sufficient for Assumption 6.1.*

Proof. Let $0 < \beta < 2$, $0 < \tau < (1/5 \wedge \frac{2-\beta}{2+5\beta})$ and $p > \beta + ((\frac{1}{2} + \frac{3}{2}\beta) \vee \frac{2}{1+5\tau})$ and suppose that Assumption 2.3 is satisfied for these constants. In order to verify Assumption 6.1 define the following quantities:

$$\overline{r} := 3\tau, \quad \overline{v} := \frac{\tau}{1+3\beta}, \quad q := \overline{r} - (1+3\beta)\overline{v} = 2\tau, \tag{6.3}$$

and recall that $\overline{w} = (1+5\tau)/4$.

ρ is suitable for Assumption 6.1(a2), as in particular $p > \beta + (\beta \vee 1)$ is satisfied due to $(1+3\beta)/2 > \beta$ and $2/(1+5\tau) > 1$. Assumption 6.1(b) is established, since $1/(2(p-\beta)) < \overline{w} = (1+5\tau)/4$ is equivalent to $p > \beta + (2/(1+5\tau))$ and $\overline{w} = (1+5\tau)/4 < 1/2 \wedge 1/(2\beta)$ holds due to $\tau < (1/5 \wedge \frac{2-\beta}{2+5\beta})$. Furthermore, simple calculations show

$$\begin{aligned}(1+2\overline{r}-2\overline{w}) \vee 1 = t_2^{-1} < 1+\tau = t_1^{-1} = (1+\frac{q}{2})\\ = 2(1-\beta\overline{w}(1+\epsilon)) < (1+2\overline{v}(p-\beta)) \wedge (1+2\overline{w})\end{aligned} \tag{6.4}$$

with $\epsilon = \frac{2-2\tau-\beta(1+5\tau)}{\beta(1+5\tau)} > 0$, since $\tau < \frac{2-\beta}{2+5\beta}$ and $(p-\beta) > (1+3\beta)/2$. Therefore, all conditions on the observation scheme are satisfied.

Additionally, if $\eta, M > 0$ and a Lebesgue null set $L \in [0,1] \cap \mathcal{L}_1$ are chosen such that the requirements of Definition 2.1 hold, we have $h_y^{(i)}(z)|z|^{(\beta+\delta)\wedge 2} \leq K|z|^{(-1+\delta)\wedge(1-\beta)}$ for each $\delta > 0$ and all $y \in [0,1] \setminus L$, $z \in (-\eta, \eta)$, $i \in \{0,1,2\}$, where $h_y^{(i)}$ denotes a density for the kernel g_i. Therefore, and due to Definition 2.1(2) and (3), we obtain $\lambda_1 - \text{ess sup}\big(\int \big(1 \wedge |z|^{(\beta+\delta)\wedge 2}\big) g_i(y,dz)\big) \leq K(\delta) < \infty$ for every $\delta > 0$ and all $i \in \{0,1,2\}$. Moreover, due to Definition 2.1(3) we have

$$h_y^{(i)}(z)|z|^{\overline{p}} \leq K|z|^{-1-\epsilon}, \tag{6.5}$$

for all $|z| \geq M$, $y \in [0,1] \setminus L$, $i \in \{0,1,2\}$ and some $K > 0$. So together with Definition 2.1(2) we obtain $\lambda_1 - \text{ess sup}_{y\in[0,1]}\big(\int |z|^{\overline{p}} \mathbb{1}_{\{|z|\geq 1\}} g_i(y,dz)\big) < \infty$ for each $i \in \{0,1,2\}$ which is Assumption 6.1(a3).

Furthermore it follows that $\frac{1+2\overline{w}}{1-\overline{w}} = \frac{6+10\tau}{3-5\tau}$ and $\frac{1+2\overline{w}}{1/2-\overline{w}} = \frac{6+10\tau}{1-5\tau}$, and as consequence Assumption 2.3(c) implies Assumption 6.1(c).

We are thus left with proving Assumption 6.1(a(4)I) and (a(4)II). Obviously, $0 < \overline{v} < \overline{r}$ holds with the choice in (6.3). First, we verify Assumption 6.1(a(4)I). To this end, we choose $\eta > 0$ and a Lebesgue null set $L \in [0,1] \cap \mathcal{L}_1$ such that $h_y^{(i)}(z) \leq K|z|^{-(1+\beta)}$ holds for all $z \in (-\eta,\eta) \setminus \{0\}$, $y \in [0,1] \setminus L$, $i \in \{0,1,2\}$ according to Definition 2.1(1) and we set $\alpha_0 := \eta/2$. Then for any choice $m_1, m_2 \in \{g_0, g_1, g_2\}$ we get

$$\begin{aligned}&\int\int \mathbb{1}_{\{|x-z|\leq\Delta_n^{\overline{r}}\}} \mathbb{1}_{\{\Delta_n^{\overline{v}}/2<|x|\leq\alpha_0\}} \mathbb{1}_{\{\Delta_n^{\overline{v}}/2<|z|\leq\alpha_0\}} m_1(y_1,dx) m_2(y_2,dz)\\ &\leq K \int\int \mathbb{1}_{\{|x-z|\leq\Delta_n^{\overline{r}}\}} \mathbb{1}_{\{\Delta_n^{\overline{v}}/2<|x|\leq\alpha_0\}} \mathbb{1}_{\{\Delta_n^{\overline{v}}/2<|z|\leq\alpha_0\}} |x|^{-(1+\beta)}|z|^{-(1+\beta)} dxdz\\ &\leq 2K \int_0^\infty \int_0^\infty \mathbb{1}_{\{|x-z|\leq\Delta_n^{\overline{r}}\}} \mathbb{1}_{\{\Delta_n^{\overline{v}}/2<x\leq\alpha_0\}} \mathbb{1}_{\{\Delta_n^{\overline{v}}/2<z\leq\alpha_0\}} x^{-(1+\beta)} z^{-(1+\beta)} dxdz,\end{aligned}$$

for all $(y_1, y_2) \in ([0,1]\setminus L)\times([0,1]\setminus L)$ and $n \in \mathbb{N}$ large enough. For the second inequality we have used symmetry of the integrand as well as $\Delta_n^{\overline{r}} < \Delta_n^{\overline{v}}/2$. In the following, we ignore the extra condition on x. Evaluation of the integral with respect to x plus a Taylor expansion give the further upper bounds

$$\begin{aligned} &K\int_0^\infty \frac{|(z-\Delta_n^{\overline{r}})^\beta - (z+\Delta_n^{\overline{r}})^\beta|}{|z^2-\Delta_n^{2\overline{r}}|^\beta} z^{-(1+\beta)} \mathbb{1}_{\{\Delta_n^{\overline{v}}/2 < z \le \alpha_0\}} dz \\ &\le K\Delta_n^{\overline{r}} \int_0^\infty \frac{\xi(z)^{\beta-1}}{|z^2-\Delta_n^{2\overline{r}}|^\beta} z^{-(1+\beta)} \mathbb{1}_{\{\Delta_n^{\overline{v}}/2 < z \le \alpha_0\}} dz \end{aligned}$$

for some $\xi(z) \in [z-\Delta_n^{\overline{r}}, z+\Delta_n^{\overline{r}}]$. Finally, we distinguish the cases $\beta < 1$ and $\beta \ge 1$ for which the numerator has to be treated differently, depending on whether it is bounded or not. The denominator is always smallest if we plug in $\Delta_n^{\overline{v}}/2$ for z. Overall,

$$\begin{aligned} &\int\int \mathbb{1}_{\{|x-z|\le\Delta_n^{\overline{r}}\}} \mathbb{1}_{\{\Delta_n^{\overline{v}}/2<|x|\le\alpha_0\}} \mathbb{1}_{\{\Delta_n^{\overline{v}}/2<|z|\le\alpha_0\}} m_1(y_1,dx) m_2(y_2,dz) \\ &\le \begin{cases} K\Delta_n^{\overline{r}}\Delta_n^{-(1+\beta)\overline{v}} \int_{\Delta_n^{\overline{v}}/2}^{\alpha_0} z^{-(1+\beta)} dz, & \text{if } \beta<1 \\ K\Delta_n^{\overline{r}}\Delta_n^{-2\beta\overline{v}} \int_{\Delta_n^{\overline{v}}/2}^{\alpha_0} z^{-(1+\beta)} dz, & \text{if } \beta\ge 1 \end{cases} \\ &\le K\Delta_n^{\overline{r}-(1+3\beta)\overline{v}} = K\Delta_n^q \end{aligned}$$

for all $m_1, m_2 \in \{0,1,2\}$ and $(y_1,y_2) \in [0,1]^2 \setminus L^2$. Finally, we consider Assumption 6.1(a(4)II), for which we proceed similarly with $n \in \mathbb{N}$ large enough, $\alpha > 0$ and $(y_1,y_2) \in [0,1]^2 \setminus L^2$, as well as $m_1, m_2 \in \{g_0, g_1, g_2\}$ arbitrary:

$$\begin{aligned} &\int\int \mathbb{1}_{\{|x-z|\le\Delta_n^{\overline{r}}\}} \mathbb{1}_{\{|x|>\alpha\}} \mathbb{1}_{\{|z|>\alpha\}} m_1(y_1,dx) m_2(y_2,dz) \\ &\le O(\Delta_n^{\overline{r}}) + 2K\int_{M'}^\infty \int_{M'}^\infty \mathbb{1}_{\{|x-z|\le\Delta_n^{\overline{r}}\}} \mathbb{1}_{\{x>\alpha\}} \mathbb{1}_{\{z>\alpha\}} x^{-2} z^{-2} dx dz. \end{aligned}$$

This inequality holds with a suitable $M' > 0$ due to Definition 2.1 (2) and (3), as we have $h_y^{(i)}(z) \le K|z|^{-2}$ for $y \in [0,1]\setminus L$, $i \in \{0,1,2\}$ and large $|z|$. Therefore,

$$\begin{aligned} &\int\int \mathbb{1}_{\{|x-z|\le\Delta_n^{\overline{r}}\}} \mathbb{1}_{\{|x|>\alpha\}} \mathbb{1}_{\{|z|>\alpha\}} m_1(y_1,dx) m_2(y_2,dz) \\ &\le O(\Delta_n^{\overline{r}}) + K\Delta_n^{\overline{r}} \int_{M'}^\infty \frac{1}{|z^2-\Delta_n^{2\overline{r}}|} z^{-2} \mathbb{1}_{\{z>\alpha\}} dz = o(\Delta_n^q) \end{aligned} \tag{6.6}$$

for $(y_1,y_2) \in [0,1]^2 \setminus L^2$ and any choice $m_1, m_2 \in \{g_0, g_1, g_2\}$. The final bound in (6.6) holds since the last integral is finite. □

6.2. Weak convergence of the empirical truncated Lévy distribution function

The proofs of the statements in Section 3 and Section 4 rely on two deep results about the weak convergence of empirical processes which are the basic blocks in the statistics considered there. We begin with a central limit theorem for the process

$$G_\rho^{(n)}(\theta,t) = \sqrt{n\Delta_n}\big(N_\rho^{(n)}(\theta,t) - N_\rho(g^{(n)};\theta,t)\big),$$

where $N_\rho(\cdot,\cdot)$ and $N_\rho(g;\cdot,\cdot)$ are defined in (2.2) and (2.3), respectively. The following result is a generalization of Theorem 3.1 in Hoffmann and Vetter (2017) which can be obtained by the choice $g_0(y,dz) = \nu(dz)$ for a Lévy measure ν and $g_1 = g_2 = 0$. The proof is given in Section A of the supplement.

Theorem 6.3. *Let Assumption 2.3 be satisfied. Then we have weak convergence* $G_\rho^{(n)} \rightsquigarrow \mathbb{G}_\rho$ *in* $\ell^\infty([0,1]\times\mathbb{R})$*, where* $\mathbb{G}_\rho$ *is a tight mean zero Gaussian process in* $\ell^\infty([0,1]\times\mathbb{R})$ *with covariance function*

$$H_\rho((\theta_1,t_1);(\theta_2,t_2)) := \int_0^{\theta_1\wedge\theta_2}\int_{-\infty}^{t_1\wedge t_2} \rho^2(z) g_0(y,dz)dy. \tag{6.7}$$

Additionally, the sample paths of $\mathbb{G}_\rho$ *are almost surely uniformly continuous with respect to the semimetric*

$$\begin{aligned} d_\rho((\theta_1,t_1);(\theta_2,t_2)) = \Big\{ \int_0^{\theta_1}\int_{t_1\wedge t_2}^{t_1\vee t_2} \rho^2(z) g_0(y,dz)dy \\ + \int_{\theta_1}^{\theta_2}\int_{-\infty}^{t_2} \rho^2(z) g_0(y,dz)dy \Big\}^{1/2} \end{aligned} \tag{6.8}$$

for $\theta_1 \le \theta_2$.

We also need a result regarding the weak convergence of a bootstrapped version of $G_\rho^{(n)}$. The corresponding process is defined by

$$\hat{G}_\rho^{(n)}(\theta,t) = \frac{1}{\sqrt{n\Delta_n}} \sum_{i=1}^{\lfloor n\theta\rfloor} \xi_i \rho\big(\Delta_i^n X^{(n)}\big) \mathbb{1}_{(-\infty,t]}\big(\Delta_i^n X^{(n)}\big) \mathbb{1}_{\{|\Delta_i^n X^{(n)}|>v_n\}}, \tag{6.9}$$

for $(\theta,t)\in[0,1]\times\mathbb{R}$, where the sequence of multipliers $(\xi_i)_{i\in\mathbb{N}}$ satisfies Assumption 3.6. The proof is given in Section B of the supplement.

Theorem 6.4. *If Assumption 2.3 holds and the multipliers* $(\xi_i)_{i\in\mathbb{N}}$ *satisfy Assumption 3.6, we have* $\hat{G}_\rho^{(n)} \rightsquigarrow_\xi \mathbb{G}_\rho$ *in* $\ell^\infty([0,1]\times\mathbb{R})$*, where the process* $\mathbb{G}_\rho$ *is defined in Theorem 6.3.*

6.3. Proofs of the results in Section 3

Proof of Theorem 3.1. For each $(\theta,t)\in[0,1]\times\mathbb{R}$, $n\in\mathbb{N}$ we have under $\mathbf{H}_1^{(loc)}$

$$\begin{aligned} \mathbb{T}_\rho^{(n)}(\theta,t) &= h_n\big(G_\rho^{(n)}\big)(\theta,t) + \sqrt{n\Delta_n}\Big(N_\rho(g^{(n)};\theta,t) - \frac{\lfloor n\theta\rfloor}{n} N_\rho(g^{(n)};1,t)\Big) \\ &= h_n\big(G_\rho^{(n)}\big)(\theta,t) + \sqrt{n\Delta_n}\Big(\theta - \frac{\lfloor n\theta\rfloor}{n}\Big)\int_{-\infty}^t \rho(z)\nu_0(dz) \end{aligned}$$

$$+ \big(N_\rho(g_1;\theta,t) - \frac{\lfloor n\theta \rfloor}{n} N_\rho(g_1;1,t)\big)$$
$$+ \sqrt{n\Delta_n}\big(N_\rho(\mathcal{R}_n;\theta,t) - \frac{\lfloor n\theta \rfloor}{n} N_\rho(\mathcal{R}_n;1,t)\big),$$

with the mappings $h_n : \ell^\infty([0,1]\times\mathbb{R}) \to \ell^\infty([0,1]\times\mathbb{R})$ defined by

$$h_n(f)(\theta,t) = f(\theta,t) - \frac{\lfloor n\theta \rfloor}{n} f(1,t), \quad (n\in\mathbb{N}), \quad h_0(f)(\theta,t) = f(\theta,t) - \theta f(1,t). \tag{6.10}$$

Thus, by Assumption 2.3(a) we obtain $\mathbb{T}_\rho^{(n)}(\theta,t) = h_n\big(G_\rho^{(n)}\big)(\theta,t) + \mathbb{T}_{\rho,g_1}(\theta,t) + o(1)$, where the o-term is deterministic. By the same reasoning as in the proof of Theorem 2.6 in Bücher et al. (2017) it can be seen that $h_n\big(G_\rho^{(n)}\big) \rightsquigarrow h_0\big(\mathbb{G}_\rho\big) = \mathbb{T}_\rho$ in $\ell^\infty([0,1]\times\mathbb{R})$. As a consequence, Slutsky's lemma (Example 1.4.7 in Van der Vaart and Wellner (1996)) yields the assertion, since the tight process $\mathbb{T}_\rho$ is separable (see Lemma 1.3.2 in the previously mentioned reference). □

Proof of Proposition 3.3. (6.5) in the proof of Proposition 6.2 shows that Assumption 6.1(a3) is also valid for $2p$ instead of p. Thus, Theorem 6.3 also holds with the function ρ replaced by ρ^2. As a consequence, we have $N_{\rho^2}^{(n)}(1,t_0) - N_{\rho^2}(g^{(n)};1,t_0) = o_{\mathbb{P}}(1)$. By (2.7) we obtain

$$N_{\rho^2}\big(g^{(n)};1,t_0\big) = \int_{-\infty}^{t_0} \rho^2(z)\nu_0(dz) + \frac{1}{\sqrt{n\Delta_n}} \int_0^1 \int_{-\infty}^{t_0} \rho^2(z) g_1(y,dz)dy +$$
$$+ \int_0^1 \int_{-\infty}^{t_0} \rho^2(z)\mathcal{R}_n(y,dz)dy = \int_{-\infty}^{t_0} \rho^2(z)\nu_0(dz) + o(1).$$

Finally, $(N_{\rho^2}^{(n)}(1,t_0))^{-1/2}\mathbb{1}_{\{N_{\rho^2}^{(n)}(1,t_0)>0\}} = (\int_{-\infty}^{t_0}\rho^2(z)\nu_0(dz))^{-1/2} + o_{\mathbb{P}}(1)$ follows due to $\int_{-\infty}^{t_0}\rho^2(z)\,\nu_0(dz) > 0$. Thus, Theorem 3.1, the continuous mapping theorem and Slutsky's lemma (Example 1.4.7 in Van der Vaart and Wellner (1996)) yield

$$\mathbb{V}_{\rho,t_0}^{(n)}(\theta) \rightsquigarrow \Big(\int_{-\infty}^{t_0} \rho^2(z)\nu_0(dz)\Big)^{-1/2} \big(\mathbb{T}_\rho(\theta,t_0) + \mathbb{T}_{\rho,g_1}(\theta,t_0)\big) = \mathbb{K}(\theta) + \bar{\mathbb{V}}_{\rho,t_0}^{(g_1)}(\theta),$$

in $\ell^\infty([0,1])$, because the process $(\int_{-\infty}^{t_0}\rho^2(z)\nu_0(dz))^{-1/2}\mathbb{T}_\rho(\cdot,t_0)$ is a tight mean zero Gaussian process with covariance function $K(\theta_1,\theta_2) = (\theta_1\wedge\theta_2) - \theta_1\theta_2$. □

Proof of Theorem 3.7. Recall the Lipschitz continuous functions $h_n : \ell^\infty([0,1]\times\mathbb{R}) \to \ell^\infty([0,1]\times\mathbb{R})$, $(n\in\mathbb{N}_0)$ defined in (6.10). Then we have $\hat{\mathbb{T}}_\rho^{(n)} = h_n(\hat{G}_\rho^{(n)})$ and Proposition 10.7 in Kosorok (2008) together with Theorem 6.4 give $h_0(\hat{G}_\rho^{(n)}) \rightsquigarrow_\xi h_0(\mathbb{G}_\rho)$ in $\ell^\infty([0,1]\times\mathbb{R})$. Moreover, we have

$$\sup_{(\theta,t)\in[0,1]\times\mathbb{R}} \big|h_n(\hat{G}_\rho^{(n)})(\theta,t) - h_0(\hat{G}_\rho^{(n)})(\theta,t)\big| =$$
$$= \sup_{(\theta,t)\in[0,1]\times\mathbb{R}} \big|\big(\theta - \frac{\lfloor n\theta\rfloor}{n}\big)\hat{G}_\rho^{(n)}(1,t)\big| = o(1)\times O_{\mathbb{P}}(1) = o_{\mathbb{P}}(1)$$

and thus Lemma E.1 yields $\hat{\mathbb{T}}_\rho^{(n)} \rightsquigarrow_\xi h_0(\mathbb{G}_\rho)$. The covariance structure (3.8) of $h_0(\mathbb{G}_\rho) = \mathbb{T}_\rho$ can be obtained using (6.7). $\square$

Proof of Proposition 3.10. First, we show (3.12) with a reasoning which is similar to the proof of Proposition F.1 in the supplement to Bücher and Kojadinovic (2016).

Fix $\alpha \in (0,1) \setminus \mathbb{Q}$. According to Proposition E.2 and the continuous mapping theorem we have $(T_\rho^{(n)}, \hat{T}^{(n)}_{\rho,\xi^{(1)}}, \dots, \hat{T}^{(n)}_{\rho,\xi^{(B)}}) \rightsquigarrow (T_{\rho,g_1}, T_{\rho,(1)}, \dots, T_{\rho,(B)})$ in $(\mathbb{R}^{B+1}, \mathbb{B}^{B+1})$ for fixed $B \in \mathbb{N}$, where $T_{\rho,(1)}, \dots, T_{\rho,(B)}$ are independent copies of the limit T_ρ in Corollary 3.9. Furthermore, let $L_{n,B}$ be the empirical c.d.f. based on the observations $\hat{T}^{(n)}_{\rho,\xi^{(1)}}, \dots, \hat{T}^{(n)}_{\rho,\xi^{(B)}}$ and let L_B be the empirical c.d.f. calculated from $T_{\rho,(1)}, \dots, T_{\rho,(B)}$. Due to the right continuity of $L_{n,B}$ we have

$$\mathbb{P}\big(T_\rho^{(n)} \geq \hat{q}^{(B)}_{1-\alpha}\big(T_\rho^{(n)}\big)\big) = \mathbb{P}\big(L_{n,B}(T_\rho^{(n)}) \geq 1-\alpha\big).$$

Moreover, using Corollary 1.3 and Remark 4.1 in Gaenssler et al. (2007) as well as Assumption 2.3(a3) and the covariance structure (3.3) of the Gaussian process $\mathbb{T}_\rho$ in Theorem 3.1 it follows that T_ρ has a continuous c.d.f. Thus, the function $\Psi_{(B)} : \mathbb{R}^{B+1} \to \mathbb{R}$ given by $\Psi_{(B)}(x_0, x_1, \dots, x_B) = B^{-1} \sum_{i=1}^B \mathbb{1}(x_i \leq x_0)$ is almost surely continuous with respect to the image measure $(T_{\rho,g_1}, T_{\rho,(1)}, \dots, T_{\rho,(B)})(\mathbb{P})$. As a consequence, we have $L_{n,B}(T_\rho^{(n)}) \rightsquigarrow L_B(T_{\rho,g_1})$ as $n \to \infty$ and with the Portmanteau theorem we obtain

$$\lim_{n\to\infty} \mathbb{P}\big(T_\rho^{(n)} \geq \hat{q}^{(B)}_{1-\alpha}\big(T_\rho^{(n)}\big)\big) = \mathbb{P}\big(L_B(T_{\rho,g_1}) \geq 1-\alpha\big),$$

because $1-\alpha \notin \{0, \frac{1}{B}, \dots, \frac{B-1}{B}, 1\}$. By the Glivenko-Cantelli theorem for every $\varepsilon \in (0, 1-\alpha)$ we can choose $B_0(\varepsilon) \in \mathbb{N}$ such that

$$\mathbb{P}\big(\sup_{x\in\mathbb{R}} |L_B(x) - L_\rho(x)| \geq \varepsilon\big) \leq \varepsilon, \tag{6.11}$$

for all $B \geq B_0(\varepsilon)$, since $T_{\rho,(1)}, \dots, T_{\rho,(B)}$ are i.i.d. with distribution function L_ρ. Thus, for every such $B \in \mathbb{N}$ we have

$$\begin{aligned}
\mathbb{P}\big(L_B(T_{\rho,g_1}) \geq 1-\alpha\big) &= \mathbb{P}\big(L_B(T_{\rho,g_1}) - L_\rho(T_{\rho,g_1}) + L_\rho(T_{\rho,g_1}) \geq 1-\alpha\big) \\
&\leq \mathbb{P}\big(L_B(T_{\rho,g_1}) - L_\rho(T_{\rho,g_1}) \geq \varepsilon\big) + \mathbb{P}\big(L_\rho(T_{\rho,g_1}) \geq 1-\alpha-\varepsilon\big) \\
&\leq \varepsilon + \mathbb{P}\big(L_\rho(T_{\rho,g_1}) \geq 1-\alpha-\varepsilon\big) \overset{\varepsilon\downarrow 0}{\to} \mathbb{P}\big(L_\rho(T_{\rho,g_1}) \geq 1-\alpha\big)
\end{aligned}$$

and we obtain

$$\limsup_{B\to\infty} \mathbb{P}\big(L_B(T_{\rho,g_1}) \geq 1-\alpha\big) \leq \mathbb{P}\big(L_\rho(T_{\rho,g_1}) \geq 1-\alpha\big). \tag{6.12}$$

The terms on both sides of inequality (6.12) are increasing in α and the right-hand side is right continuous in α. As a consequence, (6.12) is also valid for each $\alpha \in (0,1) \cap \mathbb{Q}$. Furthermore, we have

$$\liminf_{B\to\infty} \mathbb{P}\big(L_B(T_{\rho,g_1}) \geq 1-\alpha\big) \geq \mathbb{P}\big(L_\rho(T_{\rho,g_1}) > 1-\alpha\big), \tag{6.13}$$

because according to (6.11)

$$\mathbb{P}\big(L_B(T_{\rho,g_1}) \geq 1-\alpha\big) = \mathbb{P}\big(L_B(T_{\rho,g_1}) - L_\rho(T_{\rho,g_1}) + L_\rho(T_{\rho,g_1}) \geq 1-\alpha\big)$$
$$\geq \mathbb{P}\big(L_\rho(T_{\rho,g_1}) \geq 1-\alpha+\varepsilon\big) - \varepsilon \overset{\varepsilon\downarrow 0}{\to} \mathbb{P}\big(L_\rho(T_{\rho,g_1}) > 1-\alpha\big)$$

holds. Both sides of (6.13) are increasing in α and the right-hand side is left continuous in α. Thus, (6.13) is also true for $\alpha \in (0,1)\cap\mathbb{Q}$. Finally, (3.14) can be shown by exactly the same steps as above and (3.13) is an immediate consequence of the Portmanteau theorem. □

Proof of Corollary 3.12. Under $\mathbf{H}_0$ we have $\mathbb{T}_{\rho,g_1} = 0$ and $T_{\rho,g_1} = T_\rho$ is distributed acccording to L_ρ. Due to $\nu_0 \neq 0$, Assumption 2.3(a3) and the covariance structure (3.3) of $\mathbb{T}_\rho$ the c.d.f. L_ρ is continuous in virtue of Corollary 1.3 and Remark 4.1 in Gaenssler et al. (2007). As a consequence, $L_\rho(T_{\rho,g_1}) = L_\rho(T_\rho)$ is uniformly distributed on $(0,1)$ and we have $\mathbb{P}\big(L_\rho(T_\rho) > 1-\alpha\big) = \mathbb{P}\big(L_\rho(T_\rho) \geq 1-\alpha\big) = \alpha$ for all $\alpha \in (0,1)$. Hence, (3.15) follows from (3.12) and the claim (3.16) can be obtained by a similar reasoning using (3.13) as well as (3.14). □

Proof of Proposition 3.14. As in the proof of Proposition 3.3 we obtain

$$\sup_{(\theta,t)\in[0,1]\times\mathbb{R}} \big|N_\rho^{(n)}(\theta,t) - N_\rho(g_0;\theta,t)\big| = o_{\mathbb{P}}(1).$$

$(n\Delta_n)^{-1/2}\mathbb{T}_\rho^{(n)}(\theta,t)$ is given by $N_\rho^{(n)}(\theta,t) - \frac{\lfloor n\theta\rfloor}{n} N_\rho^{(n)}(1,t)$ according to (3.2). Consequently, a simple calculation shows

$$(n\Delta_n)^{-1/2}\mathbb{T}_\rho^{(n)}(\theta,t) = N_\rho(g_0;\theta,t) - \theta N_\rho(g_0;1,t) + o_{\mathbb{P}}(1) = T_{(1)}^\rho(\theta,t) + o_{\mathbb{P}}(1)$$

under $\mathbf{H}_1$, where the o-term is uniform in $(\theta,t)\in[0,1]\times\mathbb{R}$. □

Proof of Proposition 3.13. By the continuous mapping theorem, Theorem 3.7 and Remark 3.5(ii) we have $\hat{T}_{\rho,\xi^{(b)}}^{(n)} = O_{\mathbb{P}}(1)$ and $\hat{W}_{\rho,\xi^{(b)}}^{(n,t_0)} = O_{\mathbb{P}}(1)$ for all $b \in \{1,\dots,B\}$. Therefore, it suffices to show $\mathbb{P}(V_{\rho,t_0}^{(n)} \geq K) \to 1$ and $\mathbb{P}(W_\rho^{(n,t_0)} \geq K) \to 1$ for every $K > 0$ under $\mathbf{H}_1^{(\rho,t_0)}$ and $\mathbb{P}(T_\rho^{(n)} \geq K) \to 1$ for each $K>0$ under $\mathbf{H}_1$.

According to the proof of Proposition 3.3 and Proposition 3.14 the quantities $(n\Delta_n)^{-1/2}V_{\rho,t_0}^{(n)}$ and $(n\Delta_n)^{-1/2}W_\rho^{(n,t_0)}$ converge to a constant in $(0,\infty)$ in outer probability under $\mathbf{H}_1^{(\rho,t_0)}$, because $|T_{(1)}^\rho(\theta_0,t_0)| > 0$ in this case. Furthermore, due to Assumption 2.3(a3) we have $\sup_{(\theta,t)\in[0,1]\times\mathbb{R}} |T_{(1)}^\rho(\theta,t)| > 0$ under $\mathbf{H}_1$ and $(n\Delta_n)^{-1/2}T_\rho^{(n)} = \sup_{(\theta,t)\in[0,1]\times\mathbb{R}} |T_{(1)}^\rho(\theta,t)| + o_{\mathbb{P}}(1)$, because of Proposition 3.14. Thus, the assertion follows from $n\Delta_n \to \infty$. □

Proof of Proposition 3.15. According to Proposition 3.14 the random functions $\theta \mapsto \sup_{t\in\mathbb{R}} |(n\Delta_n)^{-1/2}\mathbb{T}_\rho^{(n)}(\theta,t)|$ converges weakly in $\ell^\infty([0,1])$ to the continuous function $\theta \mapsto \sup_{t\in\mathbb{R}} |T_{(1)}^\rho(\theta,\, t)|$, which due to Assumption 2.3(a3) attains a unique maximum at θ_0 under $\mathbf{H}_1$. Therefore, the claim for $\mathbf{H}_1$ follows from the argmax-continuous mapping theorem (Theorem 2.7 in Kim and Pollard (1990)). The assertion regarding $\mathbf{H}_1^{(\rho,t_0)}$ follows with a similar reasoning. □

6.4. Proofs of the results in Section 4

Proof of Lemma 4.1. If the kernel $g_0(\cdot, dz)$ is Lebesgue almost everywhere constant on $[0,\theta]$, we have $D_\rho^{(g_0)}(\zeta,\theta,t) = 0$ for all $0 \leq \zeta \leq \theta$ and $t \in \mathbb{R}$, since $\zeta^{-1}\int_0^\zeta \int_{-\infty}^t \rho(z) g_0(y,dz)dy$ is constant on $(0,\theta]$.

If on the other hand $D_\rho^{(g_0)}(\zeta,\theta,t) = 0$ for all $\zeta \in [0,\theta]$ and $t \in \mathbb{R}$ we have

$$\int_0^\zeta \int_{-\infty}^t \rho(z) g_0(y,dz)dy = \zeta\Big(\frac{1}{\theta}\int_0^\theta \int_{-\infty}^t \rho(z) g_0(y,dz)dy\Big) =: \zeta A_\theta(t)$$

for each $\zeta \in [0,\theta]$ and $t \in \mathbb{R}$. Therefore, $\int_{-\infty}^t \rho(z) g_0(y,dz) = A_\theta(t)$ holds for each fixed $t \in \mathbb{R}$ and every $y \in [0,\theta] \setminus M_{(t)}$ by Assumption 2.3(a4) and the fundamental theorem of calculus. Consequently,

$$\int_{-\infty}^t \rho(z) g_0(y,dz) = A_\theta(t) \tag{6.14}$$

holds for every $t \in \mathbb{Q}$ and each $y \in [0,\theta]$ outside the Lebesgue null set $\bigcup_{t\in\mathbb{Q}} M_{(t)}$. According to Assumption 2.3 the function $y \mapsto \int (1 \wedge |z|^p) g_0(y,dz)$ is bounded on $[0,1]$. Hence, by Lebesgue's dominated convergence theorem and the assumptions on ρ the quantities on both sides of (6.14) are right-continuous in $t \in \mathbb{R}$. As a consequence, (6.14) holds for every $t \in \mathbb{R}$ and each $y \in [0,\theta]$ outside the Lebesgue null set $\bigcup_{t\in\mathbb{Q}} M_{(t)}$. Thus, by the uniqueness theorem for measures the kernel $\rho(z) g_0(y,dz)$ is Lebesgue almost everywhere on $[0,\theta]$ equal to the finite signed measure η_θ with measure generating function $t \mapsto A_\theta(t)$ of bounded variation. Now, recall that $g_0(y,dz)$ does not charge $\{0\}$, so by Assumption 2.3(a3) the kernel $g_0(y,dz)$ is Lebesgue almost everywhere on $[0,\theta]$ equal to the measure with density $(1/\rho(z))\mathbb{1}_{\{\rho(z)\neq 0\}}\eta_\theta(dz)$. □

Proof of Theorem 4.4. We consider the functional $\Lambda\colon \ell^\infty([0,1]\times\mathbb{R}) \to \ell^\infty(C \times \mathbb{R})$ defined by

$$\Lambda(f)(\zeta,\theta,t) := f(\zeta,t) - \frac{\zeta}{\theta} f(\theta,t). \tag{6.15}$$

As $\|\Lambda(f_1) - \Lambda(f_2)\|_{C\times\mathbb{R}} \leq 2\|f_1 - f_2\|_{[0,1]\times\mathbb{R}}$ the mapping Λ is Lipschitz continuous. Thus, by Theorem 6.3 and the continuous mapping theorem $\Lambda(G_\rho^{(n)})$ converges weakly in $\ell^\infty(C\times\mathbb{R})$ to the tight mean zero Gaussian process $\mathbb{H}_\rho := \Lambda(\mathbb{G}_\rho)$ with covariance structure (4.11). Furthermore, we have $\mathbb{H}_\rho^{(n)} = \Lambda(G_\rho^{(n)}) + D_\rho^{(g_1)} + \sqrt{n\Delta_n} D_\rho^{(\mathcal{R}_n)} = \Lambda(G_\rho^{(n)}) + D_\rho^{(g_1)} + o(1)$, where the o-term is deterministic and uniform in $(\zeta,\theta,t) \in C \times \mathbb{R}$ by Assumption 2.3. Finally, the desired weak convergence follows using Slutsky's lemma (Example 1.4.7 in Van der Vaart and Wellner (1996)) and the fact that $\mathbb{H}_\rho$ is separable as it is tight (see Lemma 1.3.2 in the previously mentioned reference). □

Proof of Theorem 4.5. We have $\hat{\mathbb{H}}_\rho^{(n)} = \Lambda(\hat{G}_\rho^{(n)})$ and $\mathbb{H}_\rho = \Lambda(\mathbb{G}_\rho)$ with the Lipschitz continuous mapping Λ defined in (6.15). Thus, the assertion follows from Proposition 10.7 in Kosorok (2008). □

Proof of Theorem 4.7. 4.8 and 4.9. The assertions follow by a similar reasoning as given in the proof of Theorem 4.2, 4.3 and 4.4 in Hoffmann et al. (2017), respectively. □

Proof of Theorem 4.10. We start with a proof of $\varphi_n^* \xrightarrow{\mathbb{P}} 0$ which is equivalent to $\hat{\varkappa}_{n,B_n}^{(\alpha_n,\rho)}(r)/\sqrt{n\Delta_n} \xrightarrow{\mathbb{P}} 0$. Therefore, we have to show

$$\mathbb{P}\big(\hat{\varkappa}_{n,B_n}^{(\alpha_n,\rho)}(r)/\sqrt{n\Delta_n} \le x\big) = \mathbb{P}\Big(\frac{1}{B_n}\sum_{i=1}^{B_n} \mathbb{1}_{\{\hat{\mathbb{H}}_{\rho,*}^{(n,i)}(\hat{\theta}_n^*) \le (\sqrt{n\Delta_n}x)^{1/r}\}} \ge 1-\alpha_n\Big) \to 1, \tag{6.16}$$

for arbitrary $x>0$, by the definition of $\hat{\varkappa}_{n,B_n}^{(\alpha_n,\rho)}(r)$ in (4.16). Since the

$$\mathbb{1}_{\{\hat{\mathbb{H}}_{\rho,*}^{(n,i)}(\hat{\theta}_n^*) \le (\sqrt{n\Delta_n}x)^{1/r}\}} - \mathbb{P}_\xi\Big(\hat{\mathbb{H}}_{\rho,*}^{(n)}(\hat{\theta}_n^*) \le (\sqrt{n\Delta_n}x)^{1/r}\Big), \quad i=1,\dots,B_n,$$

are pairwise uncorrelated with mean zero and bounded by 1, we have

$$\mathbb{P}\Big(\Big|\frac{1}{B_n}\sum_{i=1}^{B_n} \mathbb{1}_{\{\hat{\mathbb{H}}_{\rho,*}^{(n,i)}(\hat{\theta}_n^*) \le (\sqrt{n\Delta_n}x)^{1/r}\}} - \mathbb{P}_\xi\Big(\hat{\mathbb{H}}_{\rho,*}^{(n)}(\hat{\theta}_n^*) \le (\sqrt{n\Delta_n}x)^{1/r}\Big)\Big| > \alpha_n/2\Big) \le 4\alpha_n^{-2}B_n^{-1} \to 0.$$

Therefore, in order to prove (6.16), it suffices to verify

$$\begin{aligned}\mathbb{P}\Big(&\mathbb{P}_\xi\Big(\hat{\mathbb{H}}_{\rho,*}^{(n)}(\hat{\theta}_n^*) \le (\sqrt{n\Delta_n}x)^{1/r}\Big) < 1-\alpha_n/2\Big)\\ &\le \frac{2}{\alpha_n}\mathbb{P}\Big(\hat{\mathbb{H}}_{\rho,*}^{(n)}(\hat{\theta}_n^*) > (\sqrt{n\Delta_n}x)^{1/r}\Big)\\ &\le 2\alpha_n^{-1}\mathbb{P}(Q_n^C)\\ &\qquad + 2\alpha_n^{-1}\mathbb{P}\Big(\Big\{2\sup_{t\in\mathbb{R}}\sup_{\theta\in[0,1]}|\hat{G}_\rho^{(n)}(\theta,t)| > (\sqrt{n\Delta_n}x)^{1/r}\Big\}\cap Q_n\Big) \to 0,\end{aligned} \tag{6.17}$$

with Q_n the set defined in (A.16). The first inequality in the above display follows with the Markov inequality and the last inequality in (6.17) is a consequence of the fact that $\hat{\mathbb{H}}_{\rho,*}^{(n)}(\hat{\theta}_n^*) \le \hat{\mathbb{H}}_{\rho,*}^{(n)}(1) \le 2\sup_{t\in\mathbb{R}}\sup_{\theta\in[0,1]}|\hat{G}_\rho^{(n)}(\theta,t)|$. Due to Lemma C.5 in part C of the supplement we obtain $\mathbb{P}\big(Q_n^C\big) \le Kn\Delta_n^{1+\tau}$ and consequently $\alpha_n^{-1}\mathbb{P}(Q_n^C) \to 0$. For the second summand on the right-hand side of (6.17) the definition of $\hat{G}_\rho^{(n)}$ in (6.9) gives

$$\mathbb{E}\Big\{\sup_{t\in\mathbb{R}}\sup_{\theta\in[0,1]}|\hat{G}_\rho^{(n)}(\theta,t)|\mathbb{1}_{Q_n}\Big\}$$

$$\leq \frac{1}{\sqrt{n\Delta_n}} \sum_{i=1}^{n} \mathbb{E}\big(|\xi_i||\rho(\Delta_i^n X^{(n)})|\mathbb{1}_{\{|\Delta_i^n X^{(n)}|>v_n\}}\mathbb{1}_{Q_n}\big) \leq K\sqrt{n\Delta_n}.$$

The final estimate above follows using Lemma C.21 in part C of the supplement, $\mathbb{E}|\xi_i| \leq 1$ for every $i = 1, \dots, n$ and independence of the multipliers and the other involved quantities. Therefore, with the Markov inequality we obtain

$$\alpha_n^{-1}\mathbb{P}\Big(\Big\{2\sup_{t\in\mathbb{R}}\sup_{\theta\in[0,1]} |\hat{G}_\rho^{(n)}(\theta,t)| > (\sqrt{n\Delta_n}x)^{1/r}\Big\} \cap Q_n\Big) \leq K\Big((n\Delta_n)^{\frac{1-r}{2r}}\alpha_n\Big)^{-1} \to 0,$$

by the assumptions on the involved sequences. Thus, we conclude $\beta_n^* \xrightarrow{\mathbb{P}} 0$.

Next we show $\hat{\varkappa}_{n,B_n}^{(\alpha_n,\rho)}(r) \xrightarrow{\mathbb{P}} \infty$, which is equivalent to

$$\mathbb{P}\big(\hat{\varkappa}_{n,B_n}^{(\alpha_n,\rho)}(r) \leq x\big) = \mathbb{P}\Big(\frac{1}{B_n}\sum_{i=1}^{B_n} \mathbb{1}_{\{\hat{\mathbb{H}}_{\rho,*}^{(n,i)}(\hat{\theta}_n^*) \leq x^{1/r}\}} \geq 1-\alpha_n\Big) \to 0,$$

for each $x > 0$. By the same considerations as in the previous paragraph it is sufficient to show

$$\mathbb{P}\Big(\mathbb{P}_\xi\Big(\hat{\mathbb{H}}_{\rho,*}^{(n)}(\hat{\theta}_n^*) > x^{1/r}\Big) \leq 2\alpha_n\Big) \to 0.$$

Let $t_0 \in \mathbb{R}$ with $N_{\rho^2}(\theta_0, t_0) > 0$. By continuity of the function $\zeta \mapsto N_{\rho^2}(\zeta, t_0)$ we can find $0 < \bar{\zeta} < \bar{\theta} < \theta_0$ with

$$N_{\rho^2}(\bar{\zeta}, t_0) > 0. \tag{6.18}$$

As $\hat{\mathbb{H}}_\rho^{(n)}(\bar{\zeta},\bar{\theta},t_0) \leq \hat{\mathbb{H}}_{\rho,*}^{(n)}(\hat{\theta}_n^*) \implies \mathbb{P}_\xi\big(\hat{\mathbb{H}}_\rho^{(n)}(\bar{\zeta},\bar{\theta},t_0) > x^{1/r}\big) \leq \mathbb{P}_\xi\big(\hat{\mathbb{H}}_{\rho,*}^{(n)}(\hat{\theta}_n^*) > x^{1/r}\big)$ on the set $\{\bar{\theta} < \hat{\theta}_n^*\}$ and consistency of the preliminary estimate it further suffices to prove

$$\mathbb{P}\Big(\mathbb{P}_\xi\Big(\hat{\mathbb{H}}_{\rho,*}^{(n)}(\hat{\theta}_n^*) > x^{1/r}\Big) \leq 2\alpha_n,\ \bar{\theta} < \hat{\theta}_n^*\Big) \leq \mathbb{P}\Big(\mathbb{P}_\xi\Big(\hat{\mathbb{H}}_\rho^{(n)}(\bar{\zeta},\bar{\theta},t_0) > x^{1/r}\Big) \leq 2\alpha_n\Big) \to 0. \tag{6.19}$$

For a proof (6.19) we use a Berry-Esseen type result. Recall the notation $\hat{\mathbb{H}}_\rho^{(n)}(\bar{\zeta},\bar{\theta},t_0) = \frac{1}{\sqrt{n\Delta_n}}\sum_{j=1}^n \hat{B}_j^n \xi_j$ from (4.12) with $\hat{B}_j^n = (\mathbb{1}_{\{j \leq \lfloor n\bar{\zeta}\rfloor\}} - \frac{\bar{\zeta}}{\bar{\theta}}\mathbb{1}_{\{j\leq\lfloor n\bar{\theta}\rfloor\}})$ $\times \hat{A}_j^n$, where

$$\hat{A}_j^n = \rho(\Delta_j^n X^{(n)})\mathbb{1}_{(-\infty,t_0]}(\Delta_j^n X^{(n)})\mathbb{1}_{\{|\Delta_j^n X^{(n)}|>v_n\}}, \quad j = 1, \dots, n.$$

It is easy to see that $\hat{W}_n^2 := \mathbb{E}_\xi(\hat{\mathbb{H}}_\rho^{(n)}(\bar{\zeta},\bar{\theta},t_0))^2 = \frac{1}{n\Delta_n}\sum_{j=1}^n (\hat{B}_j^n)^2$. Thus, Theorem 2.1 in Chen and Shao (2001) yields

$$\sup_{x\in\mathbb{R}}\Big|\mathbb{P}_\xi\Big(\hat{\mathbb{H}}^{(n)}_\rho(\bar\zeta,\bar\theta,t_0)>x\Big)-(1-\Phi(x/\hat W_n))\Big|$$
$$\le K\Big\{\sum_{i=1}^n \mathbb{E}_\xi \hat U_{i,n}^2 \mathbb{1}_{\{|\hat U_{i,n}|>1\}}+\sum_{i=1}^n \mathbb{E}_\xi|\hat U_{i,n}|^3\mathbb{1}_{\{|\hat U_{i,n}|\le 1\}}\Big\}, \tag{6.20}$$

if $\hat W_n>0$ with $\hat U_{i,n}=\frac{\hat B_i^n\xi_i}{\sqrt{n\Delta_n}\hat W_n}$ and where Φ denotes the standard normal distribution function. Before we proceed further in the proof of (6.19), we first show

$$\frac{1}{\hat W_n^2}=\frac{n\Delta_n}{\sum\limits_{j=1}^n(\hat B_j^n)^2}=O_{\mathbb{P}}(1), \tag{6.21}$$

that is $\lim\limits_{M\to\infty}\limsup\limits_{n\to\infty}\mathbb{P}\big(n\Delta_n>M\sum\limits_{j=1}^n(\hat B_j^n)^2\big)=0$. Let $M>0$. Then a straightforward calculation gives

$$\mathbb{P}\Big(n\Delta_n>M\sum_{j=1}^n(\hat B_j^n)^2\Big)\le\mathbb{P}\Big(n\Delta_n>M'\sum_{j=1}^{\lfloor n\bar\zeta\rfloor}(\hat A_j^n)^2\Big)=\mathbb{P}\Big(1/M'>N^{(n)}_{\rho^2}(\bar\zeta,t_0)\Big),$$

with $M'=M(1-\bar\zeta/\bar\theta)^2$. Consequently, with (6.18) we obtain (6.21) due to

$$N^{(n)}_{\rho^2}(\bar\zeta,t_0)=N_{\rho^2}(g^{(n)};\bar\zeta,t_0)+o_{\mathbb{P}}(1)=N_{\rho^2}(g_0;\bar\zeta,t_0)+o_{\mathbb{P}}(1),$$

because Theorem 6.3 also holds for ρ^2 since Assumption 6.1 is also valid for $2p$ instead of p (cf. (6.5) in the proof of Proposition 6.2). Recall that our main objective is to show (6.19) and thus we consider the Berry-Esseen bound on the right-hand side of (6.20). For the first summand we distinguish two cases according to the assumptions on the multiplier sequence.

Let us discuss the case of bounded multipliers first. For $M>0$ we have $|\hat U_{i,n}|\le\frac{\sqrt{M}K}{\sqrt{n\Delta_n}}$ for all $i=1,\dots,n$ on the set $\{1/\hat W_n^2\le M\}$, since $|\hat B_i^n|$ is bounded. As a consequence,

$$\sum_{i=1}^n\mathbb{E}_\xi\hat U_{i,n}^2\mathbb{1}_{\{|\hat U_{i,n}|>1\}}=0 \tag{6.22}$$

for large $n\in\mathbb{N}$ on the set $\{1/\hat W_n^2\le M\}$.

In the situation of normal multipliers, recall that there exist constants K_1, $K_2>0$ such that for $\xi\sim\mathcal{N}(0,1)$ and $y>0$ large enough we have

$$\mathbb{E}_\xi\xi^2\mathbb{1}_{\{|\xi|>y\}}=\frac{2}{\sqrt{2\pi}}\int_y^\infty z^2e^{-z^2/2}dz\le K\mathbb{P}(\mathcal{N}(0,2)>y)\le K_1\exp(-K_2y^2). \tag{6.23}$$

Thus, we can calculate for $n \in \mathbb{N}$ large enough on the set $\{1/\hat{W}_n^2 \leq M\}$

$$\begin{aligned}\sum_{i=1}^n \mathbb{E}_\xi \hat{U}_{i,n}^2 \mathbb{1}_{\{|\hat{U}_{i,n}|>1\}} &= \sum_{i=1}^n \Big(\sum_{j=1}^n (\hat{B}_j^n)^2 \Big)^{-1} (\hat{B}_i^n)^2 \mathbb{E}_\xi \xi_i^2 \mathbb{1}_{\{|\xi_i| > (\sum_{j=1}^n (\hat{B}_j^n)^2)^{1/2}/|\hat{B}_i^n|\}} \\ &\leq K \sum_{i=1}^n \Big(\sum_{j=1}^n (\hat{B}_j^n)^2 \Big)^{-1} \mathbb{E}_\xi \xi_i^2 \mathbb{1}_{\{|\xi_i| > (\sum_{j=1}^n (\hat{B}_j^n)^2)^{1/2}/K\}} \\ &\leq \frac{KM}{n\Delta_n} \sum_{i=1}^n \mathbb{E}_\xi \xi_i^2 \mathbb{1}_{\{|\xi_i| > (n\Delta_n/M)^{1/2}/K\}} \\ &\leq \frac{K_1}{\Delta_n} \exp(-K_2 n\Delta_n),\end{aligned}$$

where K_1 and K_2 depend on M. The first inequality in the display above uses boundedness of $|\hat{B}_i^n|$ again and the last one follows with (6.23). Now, according to Assumption 2.3(b) let $0 < t_2 \leq 1$ and $\delta > 0$ with $n^{-t_2+\delta} = o(\Delta_n)$. Furthermore, define $\bar{\delta} > 0$ via $1 + \bar{\delta} = 1/(t_2 - \delta)$ and $\bar{q} := 1/\bar{\delta}$. Then we have $n\Delta_n^{1+\bar{\delta}} \to \infty$ and for $n \geq N(M) \in \mathbb{N}$ on the set $\{1/\hat{W}_n^2 \leq M\}$, using $\exp(-K_2 n\Delta_n) \leq (n\Delta_n)^{-\bar{q}}$, we conclude

$$\sum_{i=1}^n \mathbb{E}_\xi \hat{U}_{i,n}^2 \mathbb{1}_{\{|\hat{U}_{i,n}|>1\}} \leq K_1 \Delta_n^{-1} (n\Delta_n)^{-\bar{q}} = K_1 \big(n\Delta_n^{1+\bar{\delta}}\big)^{-\bar{q}}. \tag{6.24}$$

We now consider the second term on the right-hand side of (6.20), for which

$$\begin{aligned}\sum_{i=1}^n \mathbb{E}_\xi |\hat{U}_{i,n}|^3 \mathbb{1}_{\{|\hat{U}_{i,n}|\leq 1\}} &\leq \sum_{i=1}^n \Big(\sum_{j=1}^n (\hat{B}_j^n)^2 \Big)^{-3/2} |\hat{B}_i^n|^3 \mathbb{E}_\xi |\xi_i|^3 \\ &\leq \frac{K}{(n\Delta_n)^{3/2}} \sum_{i=1}^n |\hat{B}_i^n|\end{aligned}$$

holds on $\{1/\hat{W}_n^2 \leq M\}$, using boundedness of $|\hat{B}_i^n|$ again. With Lemma C.21 we see that

$$\mathbb{E}\Big(\sum_{i=1}^n |\hat{B}_i^n| \mathbb{1}_{Q_n} \Big) \leq 2\mathbb{E}\Big(\sum_{i=1}^n |\hat{A}_i^n| \mathbb{1}_{Q_n} \Big) \leq K n\Delta_n.$$

Consequently,

$$\begin{aligned}&\mathbb{P}\Big(\Big\{ 1/\hat{W}_n^2 \leq M \text{ and } K \sum_{i=1}^n \mathbb{E}_\xi |\hat{U}_{i,n}|^3 \mathbb{1}_{\{|\hat{U}_{i,n}|\leq 1\}} > (n\Delta_n)^{-1/4} \Big\} \cap Q_n \Big) \\ &\leq \mathbb{P}\Big(\Big\{ \frac{K}{(n\Delta_n)^{3/2}} \sum_{i=1}^n |\hat{B}_i^n| > (n\Delta_n)^{-1/4} \Big\} \cap Q_n \Big) \leq K (n\Delta_n)^{-1/4}\end{aligned} \tag{6.25}$$

follows. Thus, from (6.22), (6.24) and (6.25) we see that with $K > 0$ from (6.20) for each $M > 0$ there exists a $K_3 > 0$ such that

$$\mathbb{P}\Big(1/\hat{W}_n^2 \le M \text{ and } K\Big\{\sum_{i=1}^n \mathbb{E}_\xi \hat{U}_{i,n}^2 \mathbb{1}_{\{|\hat{U}_{i,n}|>1\}} + \sum_{i=1}^n \mathbb{E}_\xi |\hat{U}_{i,n}|^3 \mathbb{1}_{\{|\hat{U}_{i,n}|\le 1\}}\Big\} > K_3((n\Delta_n)^{-1/4} + (n\Delta_n^{1+\bar{\delta}})^{-\bar{q}})\Big) \to 0. \quad (6.26)$$

Now we can show (6.19). Let $\eta > 0$ and according to (6.21) choose an $M > 0$ with $\mathbb{P}(1/\hat{W}_n^2 > M) < \eta/2$ for all $n \in \mathbb{N}$. For this $M > 0$ choose a $K_3 > 0$ such that the probability in (6.26) is smaller than $\eta/2$ for large n. Then for $n \in \mathbb{N}$ large enough we have

$$\mathbb{P}\Big(\mathbb{P}_\xi\Big(\hat{\mathbb{H}}_\rho^{(n)}(\bar{\zeta}, \bar{\theta}, t_0)) > x^{1/r}\Big) \le 2\alpha_n\Big) < \mathbb{P}\Big((1 - \Phi(x^{1/r}/\hat{W}_n)) \le 2\alpha_n + K_3((n\Delta_n)^{-1/4} + (n\Delta_n^{1+\bar{\delta}})^{-\bar{q}}) \text{ and } 1/\hat{W}_n^2 \le M\Big) + \eta = \eta,$$

using (6.20) and the fact, that if $1/\hat{W}_n^2 \le M$ there exists a $c' > 0$ with $(1 - \Phi(x^{1/r}/\hat{W}_n)) > c'$.

Thus, we have shown $\hat{\varkappa}_{n,B_n}^{(\alpha_n,\rho)}(r) \xrightarrow{\mathbb{P}} \infty$ and it remains to prove (4.17). Let $K = ((1+\varepsilon)/c)^{1/\varpi} > (1/c)^{1/\varpi}$ for some $\varepsilon > 0$. Then

$$\begin{aligned}\mathbb{P}\Big(\hat{\theta}_\rho^{(n)}(\hat{\varkappa}_{n,B_n}^{(\alpha_n,\rho)}(r)) > \theta_0 + K\varphi_n^*\Big) &\le \mathbb{P}\Big(\sqrt{n\Delta_n}\mathbb{D}_{\rho,*}^{(n)}(\theta) \le \hat{\varkappa}_{n,B_n}^{(\alpha_n,\rho)}(r) \text{ for some } \theta > \theta_0 + K\varphi_n^*\Big) \\ &\le \mathbb{P}\Big(\sqrt{n\Delta_n}\mathcal{D}_\rho(\theta) - \mathbb{H}_{\rho,*}^{(n)}(1) \le \hat{\varkappa}_{n,B_n}^{(\alpha_n,\rho)}(r) \text{ for some } \theta > \theta_0 + K\varphi_n^*\Big).\end{aligned}$$

By (4.13) there exists a $y_0 > 0$ with

$$\inf_{\theta\in[\theta_0+Ky_1,1]} \mathcal{D}_\rho(\theta) = \mathcal{D}_\rho(\theta_0 + Ky_1) \ge (c/(1+\varepsilon/2))(Ky_1)^\varpi$$

for all $0 \le y_1 \le y_0$. Distinguishing the cases $\{\varphi_n^* > y_0\}$ and $\{\varphi_n^* \le y_0\}$ we get due to $\varphi_n^* \xrightarrow{\mathbb{P}} 0$

$$\begin{aligned}\mathbb{P}\Big(\hat{\theta}_\rho^{(n)}(\hat{\varkappa}_{n,B_n}^{(\alpha_n,\rho)}(r)) > \theta_0 + K\varphi_n^*\Big) &\le \mathbb{P}\Big(\sqrt{n\Delta_n}(c/(1+\varepsilon/2))(K\varepsilon_n^*)^\varpi - \mathbb{H}_{\rho,*}^{(n)}(1) \le \hat{\varkappa}_{n,B_n}^{(\alpha_n,\rho)}(r)\Big) + o(1) \\ &\le P_n^{(1)} + P_n^{(2)} + o(1)\end{aligned}$$

with

$$P_n^{(1)} = \mathbb{P}\Big(\sqrt{n\Delta_n}(c/(1+\varepsilon/2))(K\varphi_n^*)^\varpi - \mathbb{H}_{\rho,*}^{(n)}(1) \le \hat{\varkappa}_{n,B_n}^{(\alpha_n,\rho)}(r) \text{ and } \mathbb{H}_{\rho,*}^{(n)}(1) \le b_n\Big),$$

$$P_n^{(2)} = \mathbb{P}\Big(\mathbb{H}_{\rho,*}^{(n)}(1) > b_n\Big),$$

where $b_n := \sqrt{\hat{\varkappa}_{n,B_n}^{(\alpha_n,\rho)}(r)}$. Due to the choice $K = \Big((1+\varepsilon)/c\Big)^{1/\varpi}$ and the definition of φ_n^* it is clear that $P_n^{(1)} = o(1)$, because $\hat{\varkappa}_{n,B_n}^{(\alpha_n,\rho)}(r) \xrightarrow{\mathbb{P}} \infty$.

Concerning $P_n^{(2)}$ let F_n be the distribution function of $\mathbb{H}_{\rho,*}^{(n)}(1)$ and let F be the distribution function of $\mathbb{H}_{\rho,*}(1)$. Then according to Corollary 1.3 and Remark 4.1 in Gaenssler et al. (2007) the function F is continuous, because $N_{\rho^2}(\theta_0, t_0) > 0$ for some $t_0 \in \mathbb{R}$. As a consequence, by Theorem 4.4 and the continuous mapping theorem F_n converges pointwise to F. Thus, for $\eta > 0$ choose an $x > 0$ with $1 - F(x) < \eta/2$ and conclude

$$P_n^{(2)} \leq \mathbb{P}(b_n \leq x) + 1 - F_n(x) \leq \mathbb{P}(b_n \leq x) + 1 - F(x) + |F_n(x) - F(x)| < \eta,$$

for $n \in \mathbb{N}$ large enough, because of $\hat{\varkappa}_{n,B_n}^{(\alpha_n,\rho)}(r) \xrightarrow{\mathbb{P}} \infty$. □

Proof of Proposition 4.13, Corollary 4.14 and Proposition 4.15. The assertions can be obtained by a similar reasoning as in the proofs of Proposition 3.10, Corollary 3.12 and Proposition 3.13 and we omit the details. □

Proof of the results in Example 2.5, Example 4.3 and Example 4.11(2).

(1) First we show that a transition kernel of the form (2.9) belongs to $\mathcal{G}(\hat{\beta}, \hat{p})$ and the function $\rho_{L,\hat{p}}$ satisfies Assumption 2.3(a2) and (a3) for $p = \hat{p}$. Let $\hat{A}$ denote a bound for $A : [0,1] \to (0,\infty)$, then for $z \in (-1,1) \setminus \{0\}$ we obtain

$$\sup_{y\in[0,1]} A(y) h_{\beta(y),p(y)}(z) \leq \hat{A} \sup_{y\in[0,1]} |z|^{-(1+\beta(y))} \leq \hat{A}|z|^{-(1+\hat{\beta})},$$

so Definition 2.1(1) is satisfied. Furthermore, for $n \in \mathbb{N}$ we have

$$\sup_{z\in C_n} \sup_{y\in[0,1]} A(y) h_{\beta(y),p(y)}(z) \leq \hat{A} \sup_{y\in[0,1]} n^{1+\beta(y)} \leq \hat{A} n^{1+\hat{\beta}},$$

which yields Definition 2.1(2). Definition 2.1(3) also holds, because for $|z| > 2$ we obtain

$$\sup_{y\in[0,1]} A(y) h_{\beta(y),p(y)}(z) \leq \hat{A} \sup_{y\in[0,1]} |z|^{-p(y)} \leq \hat{A}|z|^{-(2\hat{p}\vee 2)-\varepsilon},$$

since $\hat{p} > 1$. Obviously, $\rho_{L,\hat{p}} : \mathbb{R} \to \mathbb{R}$ is a bounded function with $\rho_{L,\hat{p}}(0) = 0$ and with the continuous derivative

$$\rho'_{L,\hat{p}}(z) = L \operatorname{sign}(z) \times \begin{cases} 2\hat{p}|z|^{\hat{p}-1}, & \text{for } 0 \leq |z| \leq 1, \\ 2\hat{p}(2 - |z|), & \text{for } 1 \leq |z| \leq 2, \\ 0, & \text{for } |z| > 2. \end{cases}$$

Consequently, there exists a $K > 0$ such that $|\rho'_{L,\hat{p}}(z)| \leq K|z|^{\hat{p}-1}$ holds for each $z \in \mathbb{R}$ and Assumption 2.3(a2) is satisfied. Moreover, Assumption 2.3(a3) is valid as well, since $\rho_{L,\hat{p}}(1) > 0$ and $\rho'_{L,\hat{p}}(z) \geq 0$ on $[1,2]$.

(2) Now we show that if additionally (4.7) and (4.8) are satisfied, $k_0 < \infty$ holds and $N_k(t)$ is a bounded function on $\mathbb{R}$ for each $k \in \mathbb{N}_0$ as stated in Example 4.11(2). To this end, elementary calculations show that the function $\bar{N}$ is given by $\bar{N}(y,t) = \Upsilon_{L,\hat{p}}(\bar{A}(y), \bar{\beta}(y), \bar{p}(y), t)$ with

$$\Upsilon_{L,\hat{p}}(a,\beta,p,t) =$$
$$La \times \begin{cases} \frac{2+\hat{p}}{p-1}|t|^{1-p}, & \text{for } t \le -2 \\ \frac{2+\hat{p}}{p-1}2^{1-p} + 4 - \frac{2\hat{p}}{3} - 2\hat{p}t^2 - \frac{\hat{p}}{3}t^3 + (2-3\hat{p})t, & \text{for } -2 \le t \le -1 \\ \frac{2+\hat{p}}{p-1}2^{1-p} + 2 + \frac{2\hat{p}}{3} + \frac{2}{\hat{p}-\beta}(1 + \operatorname{sign}(t)|t|^{\hat{p}-\beta}), & \text{for } -1 \le t \le 1 \\ \frac{2+\hat{p}}{p-1}2^{1-p} + 2\hat{p} + \frac{4}{\hat{p}-\beta} + 2\hat{p}t^2 - \frac{\hat{p}}{3}t^3 + 2t - 3\hat{p}t, & \text{for } 1 \le t \le 2 \\ \frac{4+2\hat{p}}{p-1}2^{1-p} + \frac{4}{\hat{p}-\beta} + \frac{4\hat{p}}{3} + 4 + \frac{2+\hat{p}}{1-p}t^{1-p}, & \text{for } t \ge 2. \end{cases} \tag{6.27}$$

Furthermore, it is well known from complex analysis that there is a domain $U \subset U^* \subset \mathbb{C}$ with holomorphic functions $A^* : U^* \to \mathbb{C}$, $\beta^* : U^* \to \mathbb{C}^{\hat{p}-} := \{u \in \mathbb{C} \mid \mathrm{Re}(u) < \hat{p}\}$ and $p^* : U^* \to \mathbb{C}^{1+} := \{u \in \mathbb{C} \mid \mathrm{Re}(u) > 1\}$ such that $\bar{A}$, $\bar{\beta}$ and $\bar{p}$ are the restrictions of A^*, β^* and p^* to U. Moreover, it can be seen from (6.27) that for fixed $t \in \mathbb{R}$ the mapping $(a,\beta,p) \mapsto \Upsilon_{L,\hat{p}}(a,\beta,p,t)$ is partially holomorphic on $\mathbb{C} \times \mathbb{C}^{\hat{p}-} \times \mathbb{C}^{1+}$, that is it is holomorphic in each of the variables a, β and p when the remaining variables are fixed. By a deep result of complex analysis in several variables which dates back to Hartogs (1906) this implies that $(a,\beta,p) \mapsto \Upsilon_{L,\hat{p}}(a,\beta,p,t)$ is holomorphic on $\mathbb{C} \times \mathbb{C}^{\hat{p}-} \times \mathbb{C}^{1+}$ for fixed $t \in \mathbb{R}$ (see also Remark 1.2.28 in Scheidemann (2005)). Additionally, by Proposition 1.2.2(5) in Scheidemann (2005) the function $\Xi : U^* \to \mathbb{C} \times \mathbb{C}^{\hat{p}-} \times \mathbb{C}^{1+}$ with $\Xi(y) := (A^*(y), \beta^*(y), p^*(y))$ is holomorphic and thus for each fixed $t \in \mathbb{R}$ the mapping $y \mapsto \bar{N}(y,t)$ is real analytic, because it is the restriction of the holomorphic function $y \mapsto \Upsilon_{L,\hat{p}}(\Xi(y), t)$ to U. Consequently, by shrinking the set U if necessary, we have the power series expansion

$$\bar{N}(y,t) = \sum_{k=0}^{\infty} \frac{N_k(t)}{k!}(y-\theta_0)^k, \tag{6.28}$$

for every $y \in U$ and $t \in \mathbb{R}$. If $k_0 = \infty$, then for any $k \in \mathbb{N}$ and $t \in \mathbb{R}$ we have $N_k(t) = 0$. Thus, we obtain for some constant $K > 0$

$$\Psi_{L,\hat{p}}(y) + K\frac{\bar{A}(y)}{1-\bar{p}(y)}t^{1-\bar{p}(y)} = N_0(t) \tag{6.29}$$

for each $t \ge 2$ and $y \in U$, where

$$\Psi_{L,\hat{p}}(y) = L\bar{A}(y)\Big(\frac{4+2\hat{p}}{\bar{p}(y)-1}2^{1-\bar{p}(y)} + \frac{4}{\hat{p}-\bar{\beta}(y)} + \frac{4\hat{p}}{3} + 4\Big). \tag{6.30}$$

Taking the derivative with respect to $y \in U$ on both sides of (6.29) yields

$$\Psi'_{L,\hat{p}}(y) + K\frac{\bar{A}'(y)(1-\bar{p}(y)) + \bar{A}(y)\bar{p}'(y)}{(1-\bar{p}(y))^2}t^{1-\bar{p}(y)} - \bar{p}'(y)\frac{K\bar{A}(y)}{1-\bar{p}(y)}\log(t)t^{1-\bar{p}(y)} = 0, \quad (6.31)$$

for each $y \in U$ and $t \geq 2$. Hence, $\bar{p}'(y)$ is equal to zero for each $y \in U$, because otherwise the display above is not valid for each $t \geq 2$. This fact together with (6.31) gives

$$\Psi'_{L,\hat{p}}(y) + K\frac{\bar{A}'(y)}{1-\bar{p}(y)}t^{1-\bar{p}(y)} = 0,$$

for all $y \in U$ and $t \geq 2$. Consequently, $\bar{A}'(y) = 0$ holds for every $y \in U$ and with (6.30) we obtain

$$\Psi'_{L,\hat{p}}(y) = 4L\bar{A}(\theta_0)\bar{\beta}'(y)(\hat{p} - \bar{\beta}(y))^{-2} = 0, \quad (y \in U)$$

which implies $\bar{\beta}'(y) = 0$ for all $y \in U$. Thus, $k_0 = \infty$ contradicts the assumption that at least one of the functions $\bar{A}$, $\bar{\beta}$ and $\bar{p}$ is non-constant.
The following consideration will be helpful in order to show that $N_k(t)$ is bounded in $t \in \mathbb{R}$ for each $k \in \mathbb{N}_0$. Let $f_1, f_2 : U \times \mathbb{R} \to \mathbb{R}$ be functions, which are arbitrarily often differentiable with respect to $y \in U$ for fixed $t \in \mathbb{R}$ such that for each $\ell \in \mathbb{N}_0$ the ℓ-th derivatives with respect to y satisfy

$$\sup_{t\in\mathbb{R}}\{|f_1^{(\ell)}(\theta_0,t)| \vee |f_2^{(\ell)}(\theta_0,t)|\} \leq K(K\ell)^\ell$$

for some constant $K > 0$ which does not depend on ℓ. (Here we set $0^0 := 1$.) Then by the product formula for higher derivatives we obtain for the ℓ-th derivative with respect to y of the product of f_1 and f_2

$$\begin{aligned}\sup_{t\in\mathbb{R}}|(f_1f_2)^{(\ell)}(\theta_0,t)| &= \sup_{t\in\mathbb{R}}\Big|\sum_{j=0}^{\ell}\binom{\ell}{j}f_1^{(j)}(\theta_0,t)f_2^{(\ell-j)}(\theta_0,t)\Big| \\ &\leq K^2(K\ell)^\ell\sum_{j=0}^{\ell}\binom{\ell}{j} \leq K(K\ell)^\ell.\end{aligned}$$

Observing (6.27) now yields a constant $K > 0$ such that

$$\sup_{t\in\mathbb{R}}|N_\ell(t)| \leq K(K\ell)^\ell \quad (6.32)$$

for each $\ell \in \mathbb{N}_0$ as soon as we can show that there exists a $K > 0$ such that for every $\ell \in \mathbb{N}_0$ the following bounds for the derivatives hold

$$|\bar{A}^{(\ell)}(\theta_0)| \leq K(K\ell)^\ell, \quad (6.33)$$

$$\Big|\Big(\frac{1}{\bar{p}(y)-1}\Big)^{(\ell)}(\theta_0)\Big| \leq K(K\ell)^\ell, \quad (6.34)$$

$$\Big|\Big(\frac{1}{\hat{p}-\bar{\beta}(y)}\Big)^{(\ell)}(\theta_0)\Big| \le K(K\ell)^\ell, \tag{6.35}$$

$$\sup_{t\ge 2}\Big|\Big(t^{1-\bar{p}(y)}\Big)^{(\ell)}(\theta_0)\Big| \le K(K\ell)^\ell, \tag{6.36}$$

$$\sup_{t\in[0,1]}\Big|\Big(t^{\hat{p}-\bar{\beta}(y)}\Big)^{(\ell)}(\theta_0)\Big| \le K(K\ell)^\ell. \tag{6.37}$$

Let $\bar{A}(y) = \sum_{\ell=0}^{\infty} A_\ell (y-\theta_0)^\ell$ be the power series expansion of the real analytic function $\bar{A}$ around θ_0. By the definition of real analytic functions this power series has a positive radius of convergence and due to the Cauchy-Hadamard formula this is equivalent to the existence of a constant $K>0$ with $|A_\ell| \le K^{\ell+1}$ for each $\ell\in\mathbb{N}_0$. Thus, because of $\bar{A}^{(\ell)}(\theta_0) = \ell! A_\ell$ for each $\ell\in\mathbb{N}_0$, (6.33) follows. By assumption in Example 4.3 we have $\bar{\beta}(y) \le \hat{\beta} \le 1 \vee \hat{\beta} < \hat{p} < \bar{p}(y)$ for each $y\in U$. As a consequence, the functions $y\mapsto \frac{1}{\bar{p}(y)-1}$ and $y \mapsto \frac{1}{\hat{p}-\bar{\beta}(y)}$ are real analytic on U as compositions of real analytic functions. So the same reasoning as above yields (6.34) and (6.35). Let the affine linear functions $\bar{\beta}$ and $\bar{p}$ be given by $\bar{\beta}(y) = \beta_0 + \beta_1(y-\theta_0)$ and $\bar{p}(y) = p_0 + p_1(y-\theta_0)$. Then for $\ell\in\mathbb{N}_0$, $t>0$ we have

$$\Big(t^{1-\bar{p}(y)}\Big)^{(\ell)}(\theta_0) = t^{1-p_0}(-p_1\log(t))^\ell,$$

and for $\ell\in\mathbb{N}_0$ let $h_\ell^{(1)} : (0,\infty)\to\mathbb{R}$ be defined by $h_\ell^{(1)}(t) = t^{1-p_0}(\log(t))^\ell$. $h_0^{(1)}$ is clearly bounded in $t\ge 2$ due to $p_0>1$ and for $\ell\in\mathbb{N}$ the only possible roots of the derivative of $h_\ell^{(1)}$ in $t\in(0,\infty)$ are $t=1$ and $t=\exp\{\ell/(p_0-1)\}$. Thus, we obtain for the supremum in (6.36)

$$\sup_{t\ge 2}\Big|\Big(t^{1-p(y)}\Big)^{(\ell)}(\theta_0)\Big| \le |p_1|^\ell \max\Big\{2^{1-p_0}\log(2)^\ell, \Big(\frac{\ell}{p_0-1}\Big)^\ell e^{-\ell}\Big\} \le K(K\ell)^\ell$$

for each $\ell\in\mathbb{N}_0$, because $\lim_{t\to\infty} h_\ell^{(1)}(t) = 0$. Similarly, we have for $\ell\in\mathbb{N}_0$, $t>0$

$$\Big(t^{\hat{p}-\bar{\beta}(y)}\Big)^{(\ell)}(\theta_0) = t^{\hat{p}-\beta_0}(-\beta_1\log(t))^\ell$$

and for $\ell\in\mathbb{N}_0$ let $h_\ell^{(2)} : (0,1]\to\mathbb{R}$ be defined by $h_\ell^{(2)}(t) = t^{\hat{p}-\beta_0}(\log(t))^\ell$. For $\ell\in\mathbb{N}$ the only possible roots in $(0,1]$ of the derivative of $h_\ell^{(2)}$ are $t=1$ and $t=\exp\{-\ell/(\hat{p}-\beta_0)\}$. As a consequence, we obtain for each $\ell\in\mathbb{N}_0$ for the supremum in (6.37)

$$\sup_{t\in[0,1]}\Big|\Big(t^{\hat{p}-\bar{\beta}(y)}\Big)^{(\ell)}(\theta_0)\Big| \le |\beta_1|^\ell\Big(\frac{\ell}{\hat{p}-\beta_0}\Big)^\ell e^{-\ell} \le K(K\ell)^\ell,$$

because $\lim_{t\to 0} h_\ell^{(2)}(t) = 0$. Notice that for $t=0$ the function $y\mapsto t^{\hat{p}-\bar{\beta}(y)}$ is zero constant and for $\ell=0$ the function $t\mapsto t^{\hat{p}-\beta_0}$ is bounded by 1 on $[0,1]$ due to $\hat{p}>\beta_0$.

(3) The expansion (4.18) can be deduced along the same lines as in step (3) of the proof of the results in Example 2.3 and Example 4.6(2) in Hoffmann et al. (2017) by using (6.28) and (6.32) instead of their equations (6.58) and (6.61). Furthermore, due to expansion (4.18) the quantity defined in (4.3) is clearly given by θ_0. □

Supplementary Material

Supplement: Proofs and technical details
(doi: 10.1214/19-EJS1610SUPP; .pdf).

References

Aït-Sahalia, Y. and Jacod, J. (2009a). Estimating the degree of activity of jumps in high frequency data. *The Annals of Statistics*, 37(5):2202–2244. MR2543690

Aït-Sahalia, Y. and Jacod, J. (2009b). Testing for jumps in a discretely observed process. *The Annals of Statistics*, 37(1):184–222. MR2488349

Aït-Sahalia, Y. and Jacod, J. (2010). Is Brownian motion necessary to model high-frequency data? *The Annals of Statistics*, 38:3093–3128. MR2722465

Aït-Sahalia, Y. and Jacod, J. (2014). *High-Frequency Financial Econometrics*. Princeton University Press.

Andreou, E. and Ghysels, E. (2009). Structural breaks in financial time series. In Mikosch, T., Kreiß, J.-P., Davis, R. A., and Andersen, T. G., editors, *Handbook of Financial Time Series*, pages 839–870. Springer Berlin Heidelberg.

Aue, A. and Horváth, L. (2013). Structural breaks in time series. *Journal of Time Series Analysis*, 34(1):1–16. MR3008012

Aue, A. and Steinebach, J. (2002). A note on estimating the change-point of a gradually changing stochastic process. *Statistics & Probability Letters*, 56:177–191. MR1881172

Bissell, A. F. (1984). The performance of control charts and cusums under linear trend. *Applied Statistics*, 33:145–151.

Bücher, A., Hoffmann, M., Vetter, M., and Dette, H. (2017). Nonparametric tests for detecting breaks in the jump behaviour of a time-continuous process. *Bernoulli*, 23(2):1335–1364. DOI: 10.3150/15-BEJ780. MR3606768

Bücher, A. and Kojadinovic, I. (2016). A dependent multiplier bootstrap for the sequential empirical copula process under strong mixing. *Bernoulli*, 22(2):927–968. MR3449804

Chen, L. and Shao, Q.-M. (2001). A non-uniform Berry-Esseen bound via Stein's method. *Probability Theory and Related Fields*, 120:236–254. MR1841329

Cont, R. and Tankov, P. (2004). *Financial Modelling with Jump Processes*. Chapman and Hall/CRC. ISBN: 1-58488-413-4. MR2042661

Delbaen, F. and Schachermayer, W. (1994). A general version of the fundamental theorem of asset pricing. *Mathematische Annalen*, 300:463–520. MR1304434

Gaenssler, P., Molnár, P., and Rost, D. (2007). On continuity and strict increase of the CDF for the sup-functional of a Gaussian process with applications to statistics. *Results in Mathematics*, 51:51–60. MR2375638

Gan, F. F. (1991). Ewma control chart under linear drift. *Journal of Statistical Computation and Simulation*, 38:181–200.

Hartogs, F. (1906). Zur Theorie der analytischen Funktionen mehrerer unabhängiger Veränderlichen, insbesondere über die Darstellung derselben durch Reihen, welche nach Potenzen einer Veränderlichen fortschreiten. *Mathematische Annalen*, 62:1–88. MR1511365

Hoffmann, M. and Vetter, M. (2017). Weak convergence of the empirical truncated distribution function of the Lévy measure of an Itō semimartingale. *Stochastic Processes and their Applications*, 127(5):1517–1543. MR3630234

Hoffmann, M., Vetter, M., and Dette, H. (2017). Nonparametric inference of gradual changes in the jump behaviour of time-continuous processes. *to appear: Stochastic Processes and their Applications*. arXiv:1704.04040. MR3860007

Hoffmann, M., and Dette, H. (2019). Supplement to "On detecting changes in the jumps of arbitrary size of a time-continuous stochastic process." DOI: 10.1214/19-EJS1610SUPP.

Hušková, M. (1999). Gradual changes versus abrupt changes. *Journal of Statistical Planning and Inference*, 76:109–125. MR1673343

Hušková, M. and Steinebach, J. (2002). Asymptotic tests for gradual changes. *Statistics & Decisions*, 20:137–151. MR1924608

Inoue, A. (2001). Testing for distributional change in time series. *Econometric Theory*, 17(1):156–187. MR1863569

Jacod, J. and Protter, P. (2012). *Discretization of Processes*. Springer. ISBN: 978-3-642-24126-0. MR2859096

Jandhyala, V., Fotopoulos, S., MacNeill, I., and Liu, P. (2013). Inference for single and multiple change-points in time series. *Journal of Time Series Analysis*, 34(4):423–446. doi: 10.1111/jtsa.12035. MR3070866

Kim, J. and Pollard, D. (1990). Cube root asymptotics. *The Annals of Statistics*, 18(1):191–219. MR1041391

Kosorok, M. (2008). *Introduction to Empirical Processes and Semiparametric Inference*. Springer Series in Statistics. Springer. ISBN: 978-0-387-74977-8. MR2724368

Madan, D. B., Carr, P. P., and Chang, E. C. (1998). The variance gamma process and option pricing. *European Finance Review*, 2:79–105.

Mallik, A., Banerjee, M., and Sen, B. (2013). Asymptotics for p-value based threshold estimation in regression settings. *Electronic Journal of Statistics*, 7:2477–2515. MR3117104

Mancini, C. (2009). Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scandinavian Journal of Statistics*, 36:270–296. MR2528985

Nickl, R. and Reiß, M. (2012). A Donsker theorem for Lévy measures. *Journal of Functional Analysis*, 263:3306–3332. MR2973342

Nickl, R., Reiß, M., Söhl, J., and Trabs, M. (2016). High-frequency Donsker theorems for Lévy measures. *Probability Theory and Related Fields*, 164:61–108. MR3449386

Page, E. (1954). Continuous inspection schemes. *Biometrika*, 41(1-2):100–115.

MR0088850

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

Perron, P. (2006). Dealing with structural breaks. In Patterson, K. and Mills, T., editors, *Palgrave Handbook of Econometrics*, volume 1, pages 278–352. Palgrave Macmillan. MR3184412

Reeves, J., Chen, J., Wang, X., Lund, R., and Lu, Q. (2007). A review and comparison of changepoint detection techniques for climate data. *Journal of Applied Meteorology and Climatology*, 46:900–915.

Scheidemann, V. (2005). *Introduction to Complex Analysis in Several Variables*. Birkhäuser. ISBN: 3-7643-7490-X. MR2176976

Siegmund, D. O. and Zhang, H. (1994). Confidence regions in broken line regression. In Carlstein, E., Müller, H.-G., and Siegmund, D., editors, *Change-point problems*, volume 23, pages 292–316. Institute of Mathematical Statistics. MR1477932

Stoumbos, Z., Reynolds, J. M., Ryan, T., and Woodall, W. (2000). The state of statistical process control as we proceed into the 21st century. *Journal of the American Statistical Association*, 95:992–998.

Van der Vaart, A. and Wellner, J. (1996). *Weak Convergence and Empirical Processes*. Springer. ISBN: 0-387-94640-3. MR1385671

van Kampen, N. G. (2007). *Stochastic Processes in Physics and Chemistry*. Elsevier, 3 edition. ISBN-10: 0-444-52965-9. MR0648937

Vogt, M. and Dette, H. (2015). Detecting gradual changes in locally stationary processes. *The Annals of Statistics*, 43(2):713–740. MR3319141

Vostrikova, L. (1981). Detecting disorder in multidimensional random processes. *Soviet Mathematics Doklady*, 24:55–59.