Bernoulli **23**(2), 2017, 1335–1364
DOI: 10.3150/15-BEJ780

Nonparametric tests for detecting breaks in the jump behaviour of a time-continuous process

AXEL BÜCHER[1,*], MICHAEL HOFFMANN[1,**], MATHIAS VETTER[2] and HOLGER DETTE[1,†]

[1] *Ruhr-Universität Bochum, Fakultät für Mathematik, 44780 Bochum, Germany.*
E-mail: [*] *axel.buecher@rub.de;* [**] *michael.hoffmann@rub.de;* [†] *holger.dette@rub.de*
[2] *Christian-Albrechts-Universität zu Kiel, Mathematisches Seminar, Ludewig-Meyn-Str. 4, 24118 Kiel, Germany. E-mail:* *vetter@math.uni-kiel.de*

This paper is concerned with tests for changes in the jump behaviour of a time-continuous process. Based on results on weak convergence of a sequential empirical tail integral process, asymptotics of certain test statistics for breaks in the jump measure of an Itô semimartingale are constructed. Whenever limiting distributions depend in a complicated way on the unknown jump measure, empirical quantiles are obtained using a multiplier bootstrap scheme. An extensive simulation study shows a good performance of our tests in finite samples.

Keywords: change points; Lévy measure; multiplier bootstrap; sequential empirical processes; weak convergence

1. Introduction

Recent years have witnessed a growing interest in statistical tools for high-frequency observations of time-continuous processes. With a view on finance, the seminal paper by [9] suggests to model such a process using an Itô semimartingale, say X, which is why most research has focused on the estimation of (or on tests concerned with) its characteristics. Particular interest has been paid to integrated volatility or the entire quadratic variation, mostly adapting parametric procedures based on normal distributions, as the continuous martingale part of an Itô semimartingale is nothing but a time-changed Brownian motion. For an overview on methods in this field see the recent monographs by [14] and [3].

Still less popular is inference on the jump behaviour only, even though empirical research shows a strong evidence supporting the presence of a jump component within X; see, for example, [2] or [1]. In this work, we will address the question whether the jump behaviour of X is time-invariant. Corresponding tests, commonly referred to as change point tests, are well known in the framework of discrete time series, but have recently also been extended to time-continuous processes; see, for example, [19] on changes in the drift or [12] on changes in the volatility function of X. However, to the best of our knowledge, no procedures are available for detecting breaks in the jump component.

1350-7265

Suppose that we observe an Itô semimartingale X which admits a decomposition of the form

$$X_t = X_0 + \int_0^t b_s\,ds + \int_0^t \sigma_s\,dW_s + \int_0^t \int_{\mathbb{R}} u 1_{\{|u|\le 1\}}(\mu - \bar{\mu})(ds, du) \\ + \int_0^t \int_{\mathbb{R}} u 1_{\{|u|>1\}} \mu(du, dz), \tag{1.1}$$

where W is a standard Brownian motion, μ is a Poisson random measure on $\mathbb{R}^+ \times \mathbb{R}$, and the predictable compensator $\bar{\mu}$ satisfies $\bar{\mu}(ds, du) = ds\,\nu_s(du)$. It is well known from standard literature (see, for instance, [15]) that in this setup

$$\int_0^t \int_{\mathbb{R}} \bigl(1 \wedge |u|^2\bigr)\nu_s(du)\,ds < \infty$$

holds for each $t \ge 0$.

Now, we assume that we have data from the process in a high-frequency setup. Precisely, at stage $n \in \mathbb{N}$, we are able to observe realizations of the process X at the equidistant times $i\Delta_n$ for $i = 1, \dots, n$, where the mesh $\Delta_n \to 0$, while $n\Delta_n \to \infty$. In this situation we want to test the null hypothesis that the jump behaviour of the process is the same for all n observations, that is, there exists some measure ν such that $\nu_t(dz) = \nu(dz)$ for all t, against alternatives involving the non-constancy of ν_t. For instance, one might consider an alternative consisting of one break point, that is, there exists some $\theta_0 \in (0, 1)$ and two Lévy measures ν_1, ν_2 such that the process giving the first $\lfloor n\theta_0 \rfloor$ observations has Lévy measure ν_1 and the remaining $n - \lfloor n\theta_0 \rfloor$ observations are coming from a process with Lévy measure ν_2.

For $z > 0$, set $\mathcal{I}(z) := [z, \infty)$, whereas for $z < 0$ set $\mathcal{I}(z) := (-\infty, z]$. Let $U(z) := \nu(\mathcal{I}(z))$ denote the tail integral (or spectral measure; see [23]) associated with ν, which determines the jump measure uniquely. For $\ell_1, \ell_2 \in \{1, \dots, n\}$ such that $\ell_1 < \ell_2$, define

$$U_{\ell_1:\ell_2}(z) := \frac{1}{(\ell_2 - \ell_1 + 1)\Delta_n} \sum_{j=\ell_1}^{\ell_2} 1_{\{\Delta_j^n X \in \mathcal{I}(z)\}} \qquad \bigl(z \in \mathbb{R} \setminus \{0\}\bigr),$$

with $\Delta_j^n X := X_{j\Delta_n} - X_{(j-1)\Delta_n}$, which serves as an empirical tail integral based on the increments $\Delta_{\ell_1}^n X, \dots, \Delta_{\ell_2}^n X$. If X is a Lévy process with a Lévy measure ν not changing in time, [10] illustrated that $U_{1:n}(z)$ is a suitable estimator for the tail integral $U(z)$ in the sense that, under regularity conditions, $U_{1:n}(z)$ is L^2-consistent for $U(z)$. Such a result can be shown to hold for processes with time-varying drift and volatility as well. Hence, following the approach in [13], it is likely that we can base tests for $\mathbf{H}_0$ on suitable functionals of the process

$$D_n(\theta, z) := U_{1:\lfloor n\theta \rfloor}(z) - U_{(\lfloor n\theta \rfloor + 1):n}(z),$$

where $\theta \in [0, 1]$ and $z \in \mathbb{R} \setminus \{0\}$. Under the null hypothesis, this expression can be expected to converge to 0 for all $\theta \in [0, 1]$ and $z \in \mathbb{R} \setminus \{0\}$, whereas under alternatives, for instance, those involving a change at θ_0 as described before, $D_n(\theta_0, z)$ should converge to an expression which is non-zero.

More precisely, we will consider the following standardized version of D_n, namely

$$\mathbb{T}_n(\theta, z) := \sqrt{n\Delta_n}\lambda_n(\theta)\big\{U_{1:\lfloor n\theta\rfloor}(z) - U_{(\lfloor n\theta\rfloor+1):n}(z)\big\} \tag{1.2}$$

for $\theta \in [0, 1]$ and $z \in \mathbb{R} \setminus \{0\}$, where $\lambda_n(\theta) = \frac{\lfloor n\theta\rfloor}{n}\frac{n-\lfloor n\theta\rfloor}{n}$. An appropriate functional allowing to test the hypothesis of a constant Lévy measure is for instance given by a Kolmogorov–Smirnov statistic of the form

$$T_n^{(\varepsilon)} := \sup_{\theta\in[0,1]}\sup_{|z|\geq\varepsilon}\big|\mathbb{T}_n(\theta, z)\big|, \qquad \varepsilon > 0. \tag{1.3}$$

The null hypothesis of no change in the Lévy measure is rejected for large values of $T_n^{(\varepsilon)}$. The restriction to jumps larger than ε is important, since there might be infinitely many of arbitrary small size.

The limiting distribution of the previously mentioned test statistic will turn out to depend in a complicated way on the unknown Lévy measure ν. Therefore, corresponding quantiles are not easily accessible and must be obtained by suitable bootstrap approximations. Following related ideas for detecting breaks within multivariate empirical distribution functions [13], we opt for using empirical counterparts based on a multiplier bootstrap scheme, frequently also referred to as *wild* or *weighted* bootstrap. The approach essentially consists of multiplying each indicator within the respective empirical tail integrals with an additional, independent and standardized multiplier. The underlying empirical process theory is for instance summarized in the monograph [18].

The remaining part of this paper is organized as follows: the derivation of a functional weak convergence result for the process $\mathbb{T}_n$ under the null hypothesis is the content of Section 2. The asymptotic properties of $T_n^{(\varepsilon)}$ can then easily be derived from the continuous mapping theorem. Section 3 is concerned with the approximation of the limiting distribution using the previously described multiplier bootstrap scheme. In Section 4, we discuss the formal derivation of several tests for a time-homogeneous jump behaviour, whereas an extensive simulation study is presented in Section 5. All proofs are deferred to the Appendix.

2. Functional weak convergence of the sequential empirical tail integral

In this section, we derive a functional weak convergence result for the process $\mathbb{T}_n$ defined in (1.2) under the null hypothesis. More precisely, for any fixed $\varepsilon > 0$, we will show weak convergence in the metric space $\ell^\infty(A_\varepsilon)$ of all real-valued bounded functions on A_ε equipped with the sup-norm, where $A_\varepsilon := [0, 1] \times M_\varepsilon$ with $M_\varepsilon := (-\infty, -\varepsilon] \cup [\varepsilon, \infty)$. Throughout this work, we denote by $\|f\|_M$ the sup-norm of a real-valued function f defined on a set M.

The following regularity conditions will be imposed on the underlying Itô semimartingale X with representation (1.1) and on the sampling scheme, respectively.

Condition 2.1. (a) *The drift b_t and the volatility σ_t are predictable processes and there exists a non-negative random variable S on the underlying probability space with $\mathbb{E}S^p < \infty$ for some*

$p > 2$ such that, for $\mathbb{P}$-almost every $\omega \in \Omega$,

$$\sup_{t \geq 0} \{ |b_t(\omega)| + |\sigma_t(\omega)| \} \leq S(\omega).$$

(b) *There exists some Lévy measure ν such that $\nu_t \equiv \nu$ for all $t \geq 0$.*

(c) *ν is absolutely continuous with respect to the Lebesgue measure. Its density $h = d\nu/d\lambda$, called Lévy density, is continuously differentiable in every point $z \in \mathbb{R} \setminus \{0\}$ with derivative h' and satisfies, for any $\varepsilon > 0$,*

$$\sup_{|z| \geq \varepsilon} \{ |h(z)| + |h'(z)| \} < \infty.$$

(d) *The observation scheme satisfies*

$$\Delta_n \to 0, \qquad n\Delta_n \to \infty \quad \text{and} \quad n\Delta_n^{1+\tau} \to 0, \tag{2.1}$$

where $\tau = (p-2)/(p+1) \in (0,1)$ with $p > 2$ from Condition 2.1(a).

Remark 2.2 (Lévy processes and processes with independent increments). If b_t and σ_t are deterministic and bounded functions, the process X has independent increments and condition (2.1) can be weakened to

$$\Delta_n \to 0, \qquad n\Delta_n \to \infty \quad \text{and} \quad n\Delta_n^3 \to 0. \tag{2.2}$$

The details are not worked out for the sake of brevity but they can be found in a former version of this article on arXiv; see [5]. In particular, condition (2.2) is sufficient in the important case of Lévy processes, where drift and volatility are constant.

The limiting behaviour of the process $\mathbb{T}_n$ can mainly be deduced from the next theorem, which is a result for weak convergence of a sequential empirical tail integral process. For $\theta \in [0,1]$ and $z \in \mathbb{R} \setminus \{0\}$ set

$$U_n(\theta, z) := \frac{\lfloor n\theta \rfloor}{n} U_{1:\lfloor n\theta \rfloor}(z) = \frac{1}{k_n} \sum_{j=1}^{\lfloor n\theta \rfloor} 1_{\{\Delta_j^n X \in \mathcal{I}(z)\}}, \tag{2.3}$$

where $k_n := n\Delta_n$, and denote its standardized version by

$$\mathbb{G}_n(\theta, z) := \sqrt{k_n} \{ U_n(\theta, z) - \mathbb{E} U_n(\theta, z) \}. \tag{2.4}$$

Obviously, the sample paths of $U_n(\theta, z)$ are elements of $\ell^\infty(A_\varepsilon)$.

Theorem 2.3. *Suppose the assumptions of Condition 2.1 are satisfied. Then, for any $\varepsilon > 0$, $\mathbb{G}_n \rightsquigarrow \mathbb{G}$ in $(\ell^\infty(A_\varepsilon), \|\cdot\|_{A_\varepsilon})$, where $\mathbb{G}$ is a tight mean zero Gaussian process with covariance*

$$H(\theta_1, z_1; \theta_2, z_2) := \mathbb{E}[\mathbb{G}(\theta_1, z_1)\mathbb{G}(\theta_2, z_2)] = (\theta_1 \wedge \theta_2) \times \nu(\mathcal{I}(z_1) \cap \mathcal{I}(z_2))$$

for $(\theta_1, z_1), (\theta_2, z_2) \in A_\varepsilon$. *The sample paths of* $\mathbb{G}$ *are almost surely uniformly continuous on* A_ε *with respect to the semimetric*

$$\rho(\theta_1, z_1; \theta_2, z_2) := \left\{(\theta_1 \wedge \theta_2)\nu\big(\mathcal{I}(z_1)\triangle\mathcal{I}(z_2)\big) + |\theta_1 - \theta_2|\nu\big(\mathcal{I}(z_{L(\theta_1,\theta_2)})\big)\right\}^{1/2}$$

with $L(\theta_1, \theta_2) := 1 + \mathbb{1}_{\{\theta_1 \le \theta_2\}}$ *and where* $\triangle$ *denotes the symmetric difference of two sets. Moreover* (A_ε, ρ) *is totally bounded.*

Note that we have centered $U_n(\theta, z)$ around its expectation in (2.4). In most applications, however, we are interested in estimating functionals of the jump measure, for which the next lemma is essential. By a standard approximation argument, it is sufficient for our purposes to have the lemma for Lévy processes. In that case, similar statements can be found in [11], with slightly stronger assumptions on h, and in [7] in the bivariate case. As the proof is essentially the same up to minor modifications, it is not carried out explicitly for the sake of brevity.

Lemma 2.4. *Let* X *be a Lévy process with characteristic triplet* (b, σ, ν) *and with the jump measure* ν *satisfying Condition* 2.1(c). *Let further* $\delta > 0$ *be fixed. Then there exists a constant* $K = K(\delta) > 0$ *such that, for all* $|z| \ge \delta$ *and all* $t \ge 0$,

$$\left|\mathbb{P}\big(X_t \in \mathcal{I}(z)\big) - t\nu\big(\mathcal{I}(z)\big)\right| \le Kt^2.$$

We are now in a position to consider the process

$$\tilde{\mathbb{G}}_n(\theta, z) := \sqrt{k_n}\left\{U_n(\theta, z) - \theta\nu\big(\mathcal{I}(z)\big)\right\}.$$

As an immediate consequence of the previous two results, we obtain the following sequential generalization of Theorem 4.2 of [7].

Corollary 2.5. *Suppose the assumptions of Condition* 2.1 *are satisfied. Then, for any* $\varepsilon > 0$, $\tilde{\mathbb{G}}_n \rightsquigarrow \mathbb{G}$ *in* $(\ell^\infty(A_\varepsilon), \|\cdot\|_{A_\varepsilon})$, *where* $\mathbb{G}$ *denotes the Gaussian process from Theorem* 2.3.

A further consequence of Theorem 2.3 is the desired weak convergence of the process $\mathbb{T}_n$, which was defined in (1.2), under the null hypothesis.

Theorem 2.6. *Suppose the assumptions of Condition* 2.1 *are satisfied. Then, for any* $\varepsilon > 0$, *the process* $\mathbb{T}_n$ *defined in* (1.2) *converges weakly to* $\mathbb{T}$ *in* $(\ell^\infty(A_\varepsilon), \|\cdot\|_{A_\varepsilon})$, *where*

$$\mathbb{T}(\theta, z) = \mathbb{G}(\theta, z) - \theta\mathbb{G}(1, z)$$

for $(\theta, z) \in A_\varepsilon$. $\mathbb{T}$ *is a tight mean zero Gaussian process with covariance function*

$$\hat{H}(\theta_1, z_1; \theta_2, z_2) := \mathbb{E}\left\{\mathbb{T}(\theta_1, z_1)\mathbb{T}(\theta_2, z_2)\right\} = \left\{(\theta_1 \wedge \theta_2) - \theta_1\theta_2\right\}\nu\big(\mathcal{I}(z_1) \cap \mathcal{I}(z_2)\big).$$

Using the continuous mapping theorem, we are now able to derive the weak convergence of various statistics allowing for the detection of breaks in the jump behaviour. The following corollary treats the statistic $T_n^{(\varepsilon)}$ defined in (1.3).

Corollary 2.7. *Under Condition 2.1 we have, for each $\varepsilon > 0$,*

$$T_n^{(\varepsilon)} \rightsquigarrow T^{(\varepsilon)} := \sup_{0\le\theta\le 1} \sup_{|z|\ge\varepsilon} \big|\mathbb{T}(\theta, z)\big|.$$

The covariance function of the limiting process in Theorem 2.6 depends on the Lévy measure of the underlying process, which is usually unknown in applications. If one only wants to detect changes in the tail integral of the Lévy measure at a fixed point $z_0 \in \mathbb{R} \setminus \{0\}$, the following proposition deals with the simple transformation

$$\mathbb{V}_n^{(z_0)}(\theta) := \frac{\mathbb{T}_n(\theta, z_0)}{\sqrt{U_{1:n}(z_0)}} 1_{\{U_{1:n}(z_0)>0\}}$$

of $\mathbb{T}_n$ which yields a pivotal limiting distribution.

Proposition 2.8. *Suppose Condition 2.1 is satisfied and let $z_0 \in \mathbb{R} \setminus \{0\}$ be a real number with $\nu(\mathcal{I}(z_0)) > 0$. Then, $\mathbb{V}_n^{(z_0)} \rightsquigarrow \mathbb{B}$ in $\ell^\infty([0, 1])$, where $\mathbb{B}$ denotes a standard Brownian bridge. As a consequence,*

$$V_n^{(z_0)} := \sup_{\theta\in[0,1]} \big|\mathbb{V}_n^{(z_0)}(\theta)\big| \rightsquigarrow \sup_{\theta\in[0,1]} \big|\mathbb{B}(\theta)\big|.$$

Remark 2.9. We have derived the previous results under somewhat simplified assumptions on the observation scheme in order to keep the presentation rather simple. A more realistic setting could involve additional microstructure noise effects or might rely on non-equidistant data. In both cases, standard techniques still yield similar results.

For example, in case of noisy observations, [25] has shown that a particular de-noising technique allows for virtually the same results on weak convergence as for the plain $U_n(\theta, z)$ in the case without noise. For non-equidistant data, the limiting covariance functions H and $\hat{H}$ in general depend on the sampling scheme. The latter effect is well known from high-frequency statistics in the case of volatility estimation; see, for example, [21].

3. Bootstrap approximations for the sequential empirical tail integral

We have seen in Theorem 2.6 that the distribution of the limit $\mathbb{T}$ of the process $\mathbb{T}_n$ depends in a complicated way on the unknown Lévy measure of the underlying process. However, we need the quantiles of $\mathbb{T}$ or at least good approximations for them to obtain a feasible test procedure. Typically, one uses resampling methods to solve this problem.

Probably the most natural way to do so is to use $U_{1:n}(z)$ in order to obtain an estimator $\hat{\nu}_n$ for the Lévy measure first, and to draw a large number of independent samples of an Itô semimartingale with Lévy measure $\hat{\nu}_n$ then, possibly with estimates for drift and volatility as well. Based on each sample, one might then compute the test statistic $\mathbb{T}_n$, and by doing so one obtains empirical quantiles for $\mathbb{T}$.

However, from a computational side, such a method is computationally expensive since one has to generate independent Itô semimartingales for each stage within the bootstrap algorithm.

Therefore, we have decided to work with an alternative bootstrap method based on multipliers, where one only needs to generate n i.i.d. random variables with mean zero and variance one (see also [13], who used a similar approach in the context of empirical processes).

Precisely, the situation now is as follows: The bootstrapped processes, say $\hat{Y}_n = \hat{Y}_n(X_1, \dots, X_n, \xi_1, \dots, \xi_n)$, will depend on some random variables $X_1, \dots, X_n$ and on some random weights $\xi_1, \dots, \xi_n$. The $X_1, \dots, X_n$, that we consider as collected data, are defined on a probability space $(\Omega_X, \mathcal{A}_X, \mathbb{P}_X)$. The random weights $\xi_1, \dots, \xi_n$ are defined on a distinct probability space $(\Omega_\xi, \mathcal{A}_\xi, \mathbb{P}_\xi)$. Thus, the bootstrapped processes live on the product space $(\Omega, \mathcal{A}, \mathbb{P}) := (\Omega_X, \mathcal{A}_X, \mathbb{P}_X) \otimes (\Omega_\xi, \mathcal{A}_\xi, \mathbb{P}_\xi)$. The following notion of conditional weak convergence will be essential. It can be found in [18] on pages 19–20.

Definition 3.1. *Let $\hat{Y}_n = \hat{Y}_n(X_1, \dots, X_n; \xi_1, \dots, \xi_n): (\Omega, \mathcal{A}, \mathbb{P}) \to \mathbb{D}$ be a (bootstrapped) element 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 map into $\mathbb{D}$. Then $\hat{Y}_n$ converges weakly to Y conditional on the data $X_1, X_2, \dots$ in probability, notationally $\hat{Y}_n \rightsquigarrow_\xi Y$, if and only if:*

(a) $\sup_{f \in \mathrm{BL}_1(\mathbb{D})} |\mathbb{E}_\xi f(\hat{Y}_n) - \mathbb{E} f(Y)| \xrightarrow{\mathbb{P}^*} 0$,

(b) $\mathbb{E}_\xi f(\hat{Y}_n)^* - \mathbb{E}_\xi f(\hat{Y}_n)_* \xrightarrow{\mathbb{P}} 0$ *for all* $f \in \mathrm{BL}_1(\mathbb{D})$.

Here, $\mathbb{E}_\xi$ denotes the conditional expectation over the weights ξ given the data $X_1, \dots, X_n$, whereas $\mathrm{BL}_1(\mathbb{D})$ is the space of all real-valued Lipschitz continuous functions f on $\mathbb{D}$ with sup-norm $\|f\|_\infty \le 1$ and Lipschitz constant 1. Moreover, $f(\hat{Y}_n)^*$ and $f(\hat{Y}_n)_*$ denote a minimal measurable majorant and a maximal measurable minorant with respect to the joint data (including the weights ξ), respectively.

Remark 3.2. (i) Note that we do not use a measurable majorant or minorant in item (a) of the definition. This is justified through the fact that, in 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.

(ii) Note that the implication "(ii) ⇒ (i)" in the proof of Theorem 2.9.6 in [24] shows that, in general, conditional weak convergence $\rightsquigarrow_\xi$ implies unconditional weak convergence $\rightsquigarrow$ with respect to the product measure $\mathbb{P}$.

Throughout this paper, we denote by

$$\hat{\mathbb{G}}_n = \hat{\mathbb{G}}_n(\theta, z) = \hat{\mathbb{G}}_n(X_{\Delta_n}, \dots, X_{n\Delta_n}, \xi_1, \dots, \xi_n; \theta, z)$$

the bootstrap approximation which is defined by

$$\begin{aligned}\hat{\mathbb{G}}_n(\theta, z) &:= \frac{1}{n\sqrt{k_n}} \sum_{j=1}^{\lfloor n\theta \rfloor} \sum_{i=1}^{n} \xi_j \{ 1_{\{\Delta_j^n X \in \mathcal{I}(z)\}} - 1_{\{\Delta_i^n X \in \mathcal{I}(z)\}} \} \\ &= \frac{1}{\sqrt{k_n}} \sum_{j=1}^{\lfloor n\theta \rfloor} \xi_j \big\{ 1_{\{\Delta_j^n X \in \mathcal{I}(z)\}} - \eta_n(z) \big\},\end{aligned}$$

where $\eta_n(z) = n^{-1} \sum_{i=1}^{n} \mathbb{1}_{\{\Delta_i^n X \in \mathcal{I}(z)\}}$. The following theorem establishes conditional weak convergence of this bootstrap approximation for the sequential empirical tail integral process $\mathbb{G}_n$.

Theorem 3.3. *Let Condition 2.1 be satisfied and suppose that $(\xi_j)_{j\in\mathbb{N}}$ are independent and identically distributed random variables with mean 0 and variance 1, defined on a distinct probability space as described above. Then, for any $\varepsilon > 0$,*

$$\hat{\mathbb{G}}_n \rightsquigarrow_\xi \mathbb{G}$$

in $(\ell^\infty(A_\varepsilon), \|\cdot\|_{A_\varepsilon})$, where $\mathbb{G}$ denotes the limiting process of Theorem 2.3.

Theorem 3.3 suggests to define the following bootstrapped counterparts of the process $\mathbb{T}_n$ defined in equation (1.2):

$$\begin{aligned}
\hat{\mathbb{T}}_n(\theta, z) &:= \hat{\mathbb{T}}_n(X_{\Delta_n}, \dots, X_{n\Delta_n}; \xi_1, \dots, \xi_n; \theta, z) := \hat{\mathbb{G}}_n(\theta, z) - \frac{\lfloor n\theta \rfloor}{n} \hat{\mathbb{G}}_n(1, z) \\
&= \sqrt{n\Delta_n} \frac{\lfloor n\theta \rfloor}{n} \frac{n - \lfloor n\theta \rfloor}{n} \Bigg[\frac{1}{\lfloor n\theta \rfloor \Delta_n} \sum_{j=1}^{\lfloor n\theta \rfloor} \xi_j \big\{ \mathbb{1}_{\{\Delta_j^n X \in \mathcal{I}(z)\}} - \eta_n(z) \big\} \\
&\qquad - \frac{1}{(n - \lfloor n\theta \rfloor)\Delta_n} \sum_{j=\lfloor n\theta \rfloor + 1}^{n} \xi_j \big\{ \mathbb{1}_{\{\Delta_j^n X \in \mathcal{I}(z)\}} - \eta_n(z) \big\} \Bigg],
\end{aligned}$$

The following result establishes consistency of $\mathbb{T}_n$ in the sense of Definition 3.1.

Theorem 3.4. *Under Condition 2.1, for any $\varepsilon > 0$, we have*

$$\hat{\mathbb{T}}_n \rightsquigarrow_\xi \mathbb{T}$$

in $(\ell^\infty(A_\varepsilon), \|\cdot\|_{A_\varepsilon})$, with $\mathbb{T}$ defined in Theorem 2.6.

The distribution of the limit of the Kolmogorov–Smirnov-type test statistic $T_n^{(\varepsilon)}$ defined in (1.3) can be approximated by the bootstrap statistics investigated in the following corollary, which is a simple consequence of Proposition 10.7 in [18].

Corollary 3.5. *Under Condition 2.1 we have, for each $\varepsilon > 0$,*

$$\hat{T}_n^{(\varepsilon)} := \sup_{0\le\theta\le1} \sup_{|z|\ge\varepsilon} \big| \hat{\mathbb{T}}_n(\theta, z) \big| \rightsquigarrow_\xi \sup_{0\le\theta\le1} \sup_{|z|\ge\varepsilon} \big| \mathbb{T}(\theta, z) \big| =: T^{(\varepsilon)}.$$

4. The testing procedures

4.1. Hypotheses

In order to derive a test procedure which utilizes the results on weak convergence from the previous two sections, we have to formulate our hypotheses first. Under the null hypothesis, the jump behaviour of the process is constant. More precisely, this means the following:

$\mathbf{H}_0$: We observe an Itô semimartingale as in equation (1.1) with a characteristic triplet (b_t, σ_t, ν) that satisfies Condition 2.1.

We want to test this hypothesis versus the alternative that there is exactly one change in the jump behaviour. This means in detail:

$\mathbf{H}_1$: There exists some $\theta_0 \in (0,1)$ and two Lévy measures $\nu_1 \neq \nu_2$ satisfying Condition 2.1(c) such that, at stage n, we observe an Itô semimartingale $X = X(n)$ with characteristic triplet $(b_t^{(n)}, \sigma_t^{(n)}, \nu_t^{(n)})$ such that

$$\nu_t^{(n)} = 1_{\{t < \lfloor n\theta_0 \rfloor \Delta_n\}} \nu_1 + 1_{\{t \geq \lfloor n\theta_0 \rfloor \Delta_n\}} \nu_2.$$

Furthermore, $b_t^{(n)}$ and $\sigma_t^{(n)}$ satisfy Condition 2.1(a) with a bound S which is uniform in $n \in \mathbb{N}$ and $t \geq 0$. Moreover, the observation scheme satisfies Condition 2.1(d).

The corresponding alternative for a fixed $z_0 \in \mathbb{R} \setminus \{0\}$ is then given through:

$\mathbf{H}_1^{(z_0)}$: We have the situation from $\mathbf{H}_1$, but with $\nu_1(\mathcal{I}(z_0)) \neq \nu_2(\mathcal{I}(z_0))$.

4.2. The tests and their asymptotic properties

In the sequel, let $B \in \mathbb{N}$ be some large number and let $(\xi^{(b)})_{b=1,\dots,B}$ denote independent vectors of i.i.d. random variables, $\xi^{(b)} := (\xi_j^{(b)})_{j=1,\dots,n}$, with mean zero and variance one. As before, we assume that these random variables are generated independently from the original data. We denote by $\hat{\mathbb{T}}_{n,\xi^{(b)}}$ or $\hat{T}_{n,\xi^{(b)}}^{(\varepsilon)}$ the particular statistics calculated with respect to the data and the bth bootstrap multipliers $\xi_1^{(b)}, \dots, \xi_n^{(b)}$. For a given level $\alpha \in (0,1)$, we consider the following test procedures:

KSCP-Test 1. Reject $\mathbf{H}_0$ in favor of $\mathbf{H}_1^{(z_0)}$ if $V_n^{(z_0)} \geq q_{1-\alpha}^K$, where $V_n^{(z_0)}$ is defined in Proposition 2.8 and where $q_{1-\alpha}^K$ denotes the $1-\alpha$ quantile of the Kolmogorov–Smirnov-(KS-)distribution, that is the distribution of $K = \sup_{s \in [0,1]} |\mathbb{B}(s)|$ with a standard Brownian bridge $\mathbb{B}$.

KSCP-Test 2. Reject $\mathbf{H}_0$ in favor of $\mathbf{H}_1^{(z_0)}$ if

$$W_n^{(z_0)} := \sup_{\theta \in [0,1]} \left| \mathbb{T}_n(\theta, z_0) \right| \geq \hat{q}_{1-\alpha}^{(B)}\left(W_n^{(z_0)}\right),$$

where $\hat{q}_{1-\alpha}^{(B)}(W_n^{(z_0)})$ denotes the $(1-\alpha)$-sample quantile of $\hat{W}_{n,\xi^{(1)}}^{(z_0)}, \dots, \hat{W}_{n,\xi^{(B)}}^{(z_0)}$, and where $\hat{W}_{n,\xi^{(b)}}^{(z_0)} := \sup_{\theta \in [0,1]} |\hat{\mathbb{T}}_{n,\xi^{(b)}}(\theta, z_0)|$.

CP-Test. Choose an appropriate small $\varepsilon > 0$ and reject $\mathbf{H}_0$ in favor of $\mathbf{H}_1$, if

$$T_n^{(\varepsilon)} \geq \hat{q}_{1-\alpha}^{(B)}\big(T_n^{(\varepsilon)}\big),$$

where $\hat{q}_{1-\alpha}^{(B)}(T_n^{(\varepsilon)})$ denotes the $(1-\alpha)$-sample quantile of $\hat{T}_{n,\xi^{(1)}}^{(\varepsilon)}, \dots, \hat{T}_{n,\xi^{(B)}}^{(\varepsilon)}$.

Since $\varepsilon > 0$ has to be chosen prior to an application of the CP-Test, we can only detect changes in the jumps larger than ε. From a theoretical point of view this is not entirely satisfactory, since one is interested in distinguishing arbitrary jump measures. On the other hand, in most applications only the larger jumps are of particular interest, and at least the size of Δ_n provides a natural bound to disentangle jumps from volatility. Thus, a practitioner can choose a minimum jump size ε first, and use the CP-Test to decide whether there is a change in the jumps larger than ε.

The following proposition shows that the three aforementioned tests keep the asymptotic level α under the null hypothesis.

Proposition 4.1. *Under* $\mathbf{H}_0$, *KSCP-Test* 1, *KSCP-Test* 2 *and CP-Test have asymptotic level* α *in the sense that, for all* $\alpha \in (0, 1)$,

$$\lim_{n\to\infty} \mathbb{P}\big(V_n^{(z_0)} \geq q_{1-\alpha}^K\big) = \alpha, \qquad \lim_{B\to\infty} \lim_{n\to\infty} \mathbb{P}\big\{W_n^{(z_0)} \geq \hat{q}_{1-\alpha}^{(B)}\big(W_n^{(z_0)}\big)\big\} = \alpha,$$

for all $z_0 \in \mathbb{R} \setminus \{0\}$ *with* $\nu(\mathcal{I}(z_0)) > 0$, *and*

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

for all $\varepsilon > 0$ *such that* $\nu((-\infty, -\varepsilon] \cup [\varepsilon, \infty)) > 0$.

The next proposition shows that the preceding tests are consistent under the fixed alternatives defined in Section 4.1. For the sake of brevity, we only consider alternatives involving one change point, even though such a result can be extended to a known number of multiple breaks by essentially the same proofs. We also suspect that continuous changes can be detected, but the theory becomes substantially more complicated then.

Proposition 4.2. *KSCP-Test* 1, *KSCP-Test* 2 *and CP-Test are consistent in the following sense: under* $\mathbf{H}_1^{(z_0)}$, *for all* $\alpha \in (0, 1)$ *and all* $B \in \mathbb{N}$, *we have*

$$\lim_{n\to\infty} \mathbb{P}\big(V_n^{(z_0)} \geq q_{1-\alpha}^K\big) = 1 \quad \text{and} \quad \lim_{n\to\infty} \mathbb{P}\big(W_n^{(z_0)} \geq \hat{q}_{1-\alpha}^{(B)}\big(W_n^{(z_0)}\big)\big) = 1.$$

Under $\mathbf{H}_1$, *there exists an* $\varepsilon > 0$ *such that, for all* $\alpha \in (0, 1)$ *and all* $B \in \mathbb{N}$,

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

4.3. Locating the change point

Let us finally discuss how to construct suitable estimators for the location of the change point. Again, we concentrate on the detection of a single change point. Multiple change points can be detected using a standard binary segmentation algorithm dating back to [26]. We begin with a useful proposition.

Proposition 4.3. *Fix $\varepsilon > 0$. Then, under $\mathbf{H}_1$, $(\theta, z) \mapsto k_n^{-1/2}\mathbb{T}_n(\theta, z)$ converges in $(\ell^\infty(A_\varepsilon), \|\cdot\|_{A_\varepsilon})$ to the function*

$$T(\theta, z) := \begin{cases} \theta(1-\theta_0)\big\{\nu_1(z) - \nu_2(z)\big\}, & \text{if } \theta \le \theta_0, \\ \theta_0(1-\theta)\big\{\nu_1(z) - \nu_2(z)\big\}, & \text{if } \theta \ge \theta_0, \end{cases}$$

in outer probability, with $\nu_1(z) := \nu_1(\mathcal{I}(z))$ and $\nu_2(z) := \nu_2(\mathcal{I}(z))$.

Since $\theta \mapsto T(\theta, z)$ attains its maximum in θ_0, natural estimators for the position of the change point are therefore given by

$$\hat{\theta}_n^{(\varepsilon)} := \operatorname*{arg\,max}_{\theta\in[0,1]} \sup_{|z|\ge\varepsilon} \big|\mathbb{T}_n(\theta, z)\big|$$

for the test problem $\mathbf{H}_0$ versus $\mathbf{H}_1$ and by

$$\tilde{\theta}_n^{(z_0)} := \operatorname*{arg\,max}_{\theta\in[0,1]} \big|\mathbb{T}_n(\theta, z_0)\big|$$

in the setup $\mathbf{H}_0$ versus $\mathbf{H}_1^{(z_0)}$. Both estimators are consistent.

Proposition 4.4. *If $\mathbf{H}_1$ is true, there exists an $\varepsilon > 0$ such that $\hat{\theta}_n^{(\varepsilon)} = \theta_0 + o_{\mathbb{P}}(1)$ as $n \to \infty$. In the special case of $\mathbf{H}_1^{(z_0)}$, we have $\tilde{\theta}_n^{(z_0)} = \theta_0 + o_{\mathbb{P}}(1)$.*

5. Finite-sample performance

In this section, we present results of a large scale Monte Carlo simulation study, assessing the finite-sample performance of the proposed test statistics for detecting breaks in the Lévy measure. Moreover, under the alternative of one single break, we show results on the performance of the estimator for the break point from Section 4.3.

The experimental design of the study is as follows.

- We consider five different choices for the *number of trading days*, namely $k_n = 50, 75, 100, 150, 250$, and corresponding *frequencies* $\Delta_n^{-1} = 450, 300, 225, 150, 90$. Note that $n = k_n\Delta_n^{-1} = 22,500$ for any of these choices.
- We consider two different models for the *drift* and the *volatility*: either, we set $b_t = \sigma_t \equiv 0$ or $b_t = \sigma_t \equiv 1$, resulting in a pure jump process and a process including a continuous component, respectively.

- We consider a one parametric model for the *tail integral*, namely

$$U_\beta(z) = \nu_\beta\big(\mathcal{I}(z)\big) = \begin{cases} \left(\dfrac{\beta}{\pi z}\right)^{1/2}, & \text{if } z > 0, \\ 0, & \text{if } z < 0, \end{cases} \qquad \beta > 0 \tag{5.1}$$

  (which yields a 1/2-stable subordinator in the case of $b_t = \sigma_t \equiv 0$). For the parameter β, we consider 51 different choices, that is $\beta = 1 + 2j/25$, with $j \in 0, \dots, 50$, ranging from $\beta = 1$ to $\beta = 5$.
- We consider models with one single break in the tail integral at 50 different *break points*, ranging from $\theta_0 = 0$ to $\theta_0 = 0.98$ (note that $\theta_0 = 0$ corresponds to the null hypothesis). The tail integrals before and after the break point are chosen from the previous parametric model.

The target values of our study are, on the one hand, the empirical rejection level of the tests and, on the other hand, the empirical distribution of the estimators for the change point θ_0. To assess these target values, any combination of the previously described settings was run 1000 times, with the bootstrap tests being based on $B = 250$ bootstrap replications. The Itô semimartingales were simulated by a straight-forward modification of Algorithm 6.13 in [8], where, under alternatives involving one break point, we simply merged two paths of independent semimartingales together.

The simulation results under these settings are partially reported in Tables 1 and 2 (for the null hypothesis) and in Figures 1–4 (for various alternatives). More precisely, Tables 1 and 2 contain simulated rejection rates under the null hypothesis for various values of k_n and z_0 in the KSCP-Tests, for the pure jump subordinator (Table 1) and for the process involving a continuous component (Table 2). For the CP-Tests, the suprema over $z \in (-\infty, -\varepsilon] \cup [\varepsilon, \infty)$ were approximated by taking a maximum over a finite grid M of positive numbers, since the simulated

Table 1. Test procedures under $\mathbf{H}_0$. Simulated relative frequency of rejections in the application of the KSCP-Test 1, the KSCP-Test 2 and the CP-Test, using 1000 pure jump subordinator data vectors under the null hypothesis

k_n	CP-Test	Pointwise tests	$z_0 = 0.1$	$z_0 = 0.15$	$z_0 = 0.25$	$z_0 = 1$	$z_0 = 2$
50	0.06	KSCP-Test 1	0.048	0.056	0.047	0.035	0.033
		KSCP-Test 2	0.060	0.067	0.060	0.050	0.048
75	0.054	KSCP-Test 1	0.034	0.044	0.045	0.041	0.046
		KSCP-Test 2	0.045	0.059	0.061	0.058	0.060
100	0.06	KSCP-Test 1	0.047	0.044	0.042	0.044	0.042
		KSCP-Test 2	0.060	0.056	0.058	0.062	0.056
150	0.06	KSCP-Test 1	0.049	0.056	0.049	0.040	0.042
		KSCP-Test 2	0.065	0.064	0.065	0.059	0.061
250	0.07	KSCP-Test 1	0.046	0.042	0.046	0.055	0.050
		KSCP-Test 2	0.054	0.048	0.059	0.072	0.060

Table 2. Test procedures under $\mathbf{H}_0$. Simulated relative frequency of rejections in the application of the KSCP-Test 1, the KSCP-Test 2 and the CP-Test, using 1000 subordinator data vectors plus a drift $b = 1$ and plus a Brownian motion under $\mathbf{H}_0$

k_n	CP-Test	Pointwise tests	$z_0 = 2\sqrt{\Delta_n}$	$z_0 = 3.5\sqrt{\Delta_n}$	$z_0 = 6.5\sqrt{\Delta_n}$	$z_0 = 7\sqrt{\Delta_n}$
50	0.049	KSCP-Test 1	0.032	0.036	0.035	0.031
		KSCP-Test 2	0.049	0.051	0.049	0.050
75	0.050	KSCP-Test 1	0.042	0.039	0.039	0.032
		KSCP-Test 2	0.050	0.057	0.051	0.053
100	0.051	KSCP-Test 1	0.039	0.040	0.037	0.038
		KSCP-Test 2	0.051	0.054	0.049	0.057
150	0.057	KSCP-Test 1	0.038	0.045	0.034	0.039
		KSCP-Test 2	0.057	0.057	0.053	0.052
250	0.049	KSCP-Test 1	0.031	0.035	0.042	0.030
		KSCP-Test 2	0.049	0.048	0.053	0.042

processes had only positive jumps (see (5.1)): we used the grids $M = \{j \cdot 0.05 \mid j = 1, \dots, 200\}$ in the pure jump case, resulting in $\varepsilon = 0.05$, and $M = \{(2 + j \cdot 0.5)\sqrt{\Delta_n} \mid j = 0, \dots, 196\}$ in the case $b_t = \sigma_t \equiv 1$, resulting in $\varepsilon = 2\sqrt{\Delta_n}$. In the latter case, we chose ε depending on $\sqrt{\Delta_n}$ since jumps of smaller size may be dominated by the Brownian component resulting in a loss of efficiency of the CP-Test (see also the results in Figure 3 below). The results in the two tables reveal a rather precise approximation of the nominal level of the tests ($\alpha = 5\%$) in all scenarios. In general, KSCP-Test 1 turns out to be slightly more conservative than KSCP-Test 2.

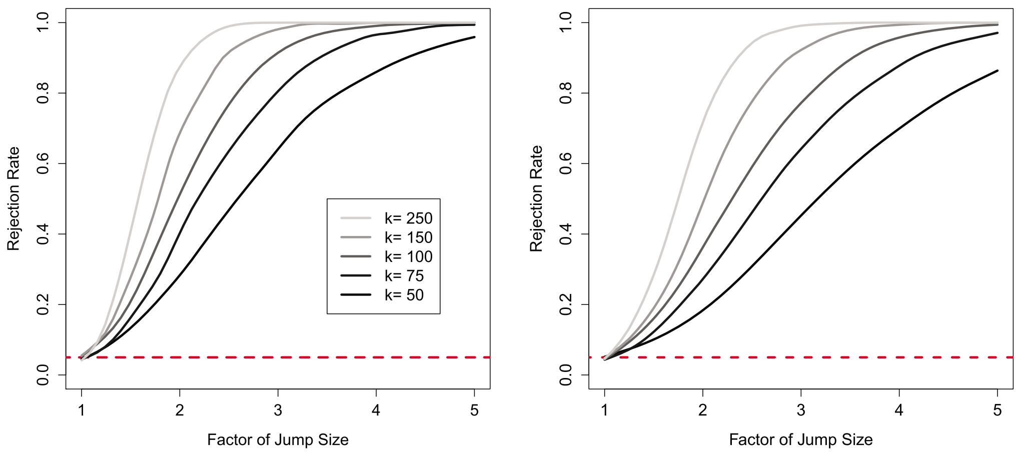

Figure 1. Rejection rate of the CP-Test for pure jump subordinator data (on the left-hand side) and a subordinator plus a drift and a Brownian motion (on the right-hand side). β changes from 1 to the factor of jump size.

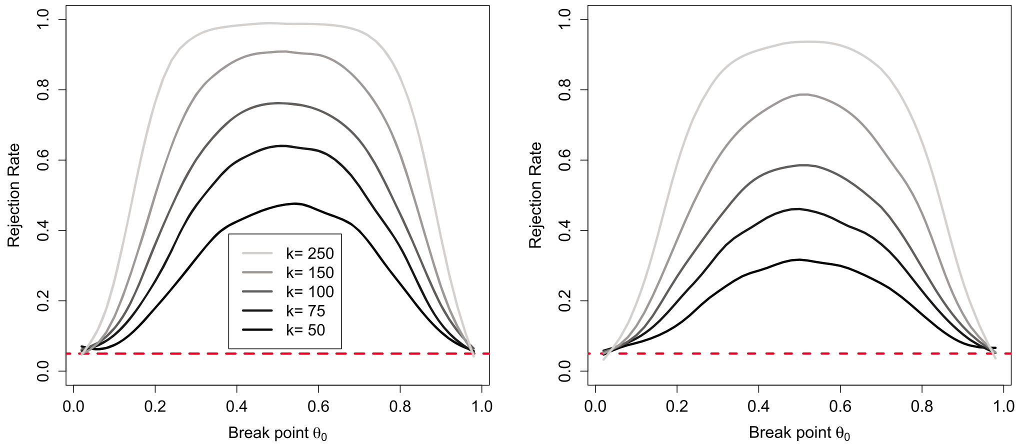

Figure 2. Rejection rate of the CP-Test for pure jump subordinator data (left panel) and a subordinator with a drift plus a Brownian motion (right panel) for different change point locations.

The results presented in Figure 1 consider the CP-Test for alternatives involving one fixed break point at $\theta_0 = 0.5$ and a varying height of the jump size, as measured through the value of β in (5.1). In contrast to the results in Tables 1 and 2, due to computational reasons, we subsequently used smaller grids $M = \{j \cdot 0.2 \mid j = 1, \dots, 20\}$ for the case $b_t = \sigma_t \equiv 0$, resulting in $\varepsilon = 0.2$, and $M = \{j \cdot 2.5 \cdot \sqrt{\Delta_n} \mid j = 1, \dots, 20\}$ for the case $b_t = \sigma_t \equiv 1$, resulting in $\varepsilon = 2.5\sqrt{\Delta_n}$. The left plot is based on the pure jump process ($b_t = \sigma_t \equiv 0$), whereas the right one is based on $b_t = \sigma_t \equiv 1$. The dashed red line indicates the nominal level of $\alpha = 5\%$. We observe

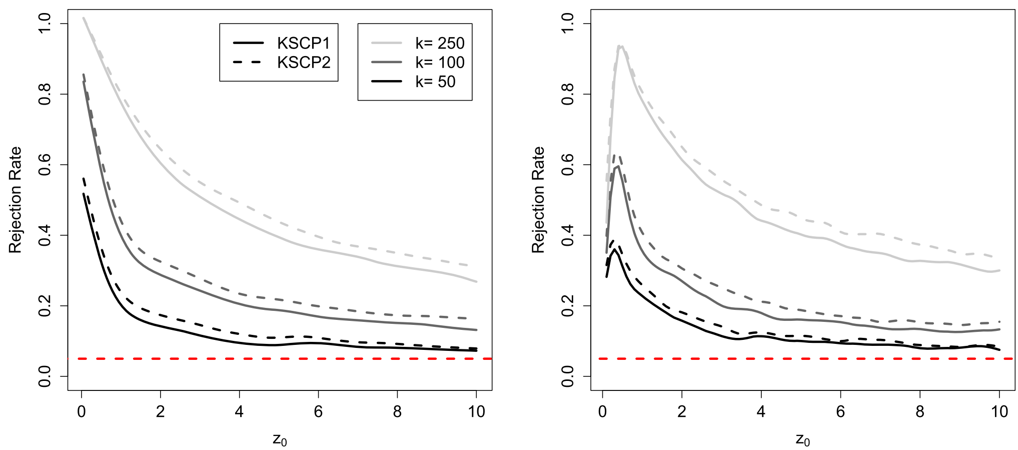

Figure 3. Rejection rates of the KSCP-Test 1 and KSCP-Test 2 for different z_0. Left panel: pure jump subordinator, right panel: subordinator with a drift plus Brownian motion.

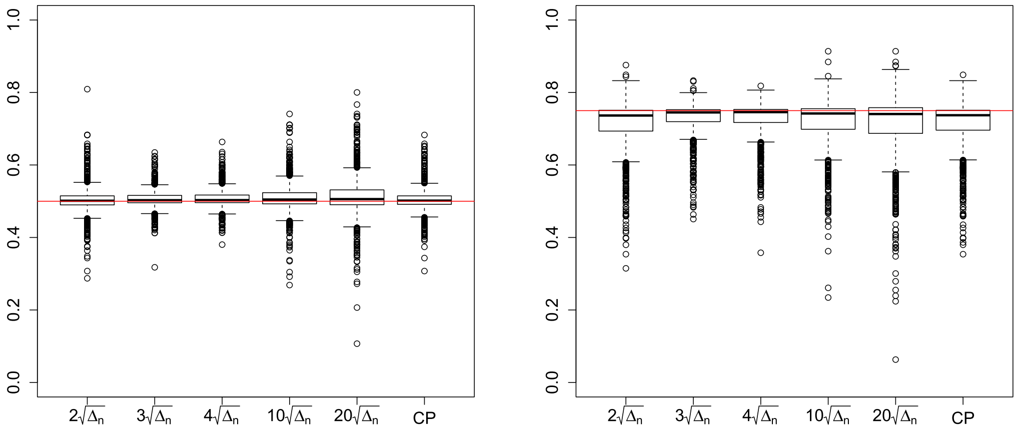

Figure 4. Box plots for the estimators $\tilde{\theta}_n^{(z_0)}$ and $\hat{\theta}_n^{(\varepsilon)}$ based on a subordinator with a drift plus Brownian motion and a change from $\beta = 1$ to $\beta = 4$ at $\theta_0 = 0.5$ (left panel) and $\theta_0 = 0.75$ (right panel). The first five box plots in each panel correspond to five different choices of z_0.

that the rejection rate of the test is increasing in β (as to be expected) and in k_n. The latter can be explained by the fact that k_n represents the effective sample size (interpretable as the number of trading days). Finally, the rejection rates turn out to be higher when no continuous component is involved in the underlying semimartingale.

The graphics in Figure 2 show the rate of rejection of the CP-Test under alternatives involving one break point from $\beta = 1$ to $\beta = 2.5$ within the model in (5.1) for varying locations of the change point $\theta_0 \in (0, 1)$. Again, the left and right plots correspond to $b_t = \sigma_t \equiv 0$ and $\equiv 1$, respectively. Additionally to the general conclusions drawn from the results in Figure 1, we observe that break points can be detected best if $\theta_0 = 1/2$, and that the rejection rates are symmetric around that point.

Figure 3 shows the rejection rates of the KSCP-Test 1 and KSCP-Test 2, evaluated at different points z_0, for one fixed alternative model involving a single change from $\beta = 1$ to $\beta = 2.5$ at the point $\theta_0 = 1/2$. The curves in the left plot are based on a pure jump process. We can see that the rejection rates are decreasing in z_0, explainable by the fact that there are only very few large jumps both for $\beta = 1$ and for $\beta = 2.5$. In the right plot, involving drift and volatility ($b_t = \sigma_t \equiv 1$), we observe a maximal value of the rejection rates that is increasing in the number of trading days, k_n. For values of z_0 smaller than this maximum, the contribution of the Brownian component (an independent normally distributed term with variance Δ_n within each increment $\Delta_j^n X$) predominates the jumps of that size and results in a decrease of the rejection rate.

Finally, in Figure 4, we depict box plots for the estimators $\tilde{\theta}_n^{(z_0)}$ and $\hat{\theta}_n^{(\varepsilon)}$ of the change point for certain values of z_0 and for M as specified in the case of Tables 1 and 2. The results are based on two models, involving a change in β from 1 to 4 at time point $\theta_0 = 0.5$ (left panel) and $\theta_0 = 0.75$ (right panel) for $k_n = 250$ and $\Delta_n^{-1} = 90$, and with $b_t = \sigma_t \equiv 1$. We observe a reasonable approximation of the true value (indicated by the red line) with more accurate approximations for $\theta_0 = 0.5$. For $\theta_0 = 0.75$, the distribution of the estimator is skewed, giving more weight to the left tail directing to $\theta_0 = 0.5$. This might be explained by the fact that the distribution of the argmax absolute value of a tight-down stochastic process indexed by $\theta \in [0, 1]$ gives very small weight to the boundaries of the unit interval. Moreover, as for the results presented in the right plot of Figure 3, the plots in Figure 4 reveal that the estimator $\tilde{\theta}_n^{(z_0)}$ behaves best for an intermediate choice of z_0. Results for $b_t = \sigma_t \equiv 0$ are not depicted for the sake of brevity, since they do not transfer any additional insight.

Appendix

A.1. Proof of Theorem 2.3

Write $X_t = Y_t + Z_t$, where

$$Y_t = X_0 + \int_0^t b_s\,ds + \int_0^t \sigma_s\,dW_s \tag{A.1}$$

is the sum of the first three summands in representation (1.1) while Z_t is a pure jump Lévy process with characteristics $(0, 0, \nu)$. Let $\mathbb{G}_n^\circ \in \ell^\infty(A_\varepsilon)$ denote the process defined in (2.4), but based on the increments $\Delta_j^n Z = Z_{j\Delta_n} - Z_{(j-1)\Delta_n}$ instead of $\Delta_j^n X$.

First, we will show the claim of Theorem 2.3 for the processes $\mathbb{G}_n^\circ$ and afterwards we will prove that $\|\mathbb{G}_n - \mathbb{G}_n^\circ\|_{A_\varepsilon} = o_{\mathbb{P}}(1)$ as $n \to \infty$. This yields the assertion by Lemma 1.10.2 in [24].

Weak convergence of the process $\mathbb{G}_n^\circ$ can be deduced from Theorem 11.16 of [18]. Note that $\mathbb{G}_n^\circ$ can be written as

$$\mathbb{G}_n^\circ(\theta, z) = \frac{1}{\sqrt{k_n}} \sum_{j=1}^{\lfloor n\theta \rfloor} \bigl\{ 1_{\{\Delta_j^n Z \in \mathcal{I}(z)\}} - \mathbb{P}\bigl(\Delta_j^n Z \in \mathcal{I}(z)\bigr) \bigr\} = \sum_{j=1}^{n} \bigl\{ f_{nj}(\omega; \theta, z) - \mathbb{E} f_{nj}(\cdot; \theta, z) \bigr\}$$

with the triangular array $\{f_{nj}(\omega; \theta, z) \mid n \geq 1; j = 1, \dots, n; (\theta, z) \in A_\varepsilon\}$ consisting of the processes

$$f_{nj}(\theta, z) := f_{nj}(\omega; \theta, z) := \frac{1}{\sqrt{k_n}} 1_{\{j \leq \lfloor n\theta \rfloor\}} 1_{\{\Delta_j^n Z(\omega) \in \mathcal{I}(z)\}},$$

which are independent within rows. By Theorem 11.16 in [18], the assertion for $\mathbb{G}_n^\circ$ holds if the following six conditions for $\{f_{nj}\}$ can be established:

(1) $\{f_{nj}\}$ is almost measurable Suslin (AMS);
(2) the $\{f_{nj}\}$ are manageable with envelopes $\{F_{nj} \mid n \in \mathbb{N}, j = 1, \dots, n\}$ given through $F_{nj} := k_n^{-1/2} 1_{\{|\Delta_j^n Z| \geq \varepsilon\}}$, which are also independent within rows;
(3) $H(\theta_1, z_1; \theta_2, z_2) = \lim_{n\to\infty} \mathbb{E}\{\mathbb{G}_n^\circ(\theta_1, z_1)\mathbb{G}_n^\circ(\theta_2, z_2)\}$ for all $(\theta_1, z_1), (\theta_2, z_2) \in A_\varepsilon$;
(4) $\limsup_{n\to\infty} \sum_{j=1}^n \mathbb{E} F_{nj}^2 < \infty$;
(5) $\lim_{n\to\infty} \sum_{j=1}^n \mathbb{E} F_{nj}^2 1_{\{F_{nj} > \eta\}} = 0$ for all $\eta > 0$;

(6) $\rho(\theta_1, z_1; \theta_2, z_2) = \lim_{n\to\infty} \rho_n(\theta_1, z_1; \theta_2, z_2)$ for every $(\theta_1, z_1), (\theta_2, z_2) \in A_\varepsilon$, where

$$\rho_n(\theta_1, z_1; \theta_2, z_2) := \left\{ \sum_{j=1}^{n} \mathbb{E}\big| f_{nj}(\cdot; \theta_1, z_1) - f_{nj}(\cdot; \theta_2, z_2)\big|^2 \right\}^{1/2}.$$

Moreover, $\rho_n(\theta_1^{(n)}, z_1^{(n)}; \theta_2^{(n)}, z_2^{(n)}) \to 0$ for all sequences $(\theta_1^{(n)}, z_1^{(n)})_{n\in\mathbb{N}}$ and $(\theta_2^{(n)}, z_2^{(n)})_{n\in\mathbb{N}} \subset A_\varepsilon$ such that $\rho(\theta_1^{(n)}, z_1^{(n)}; \theta_2^{(n)}, z_2^{(n)}) \to 0$.

Proof of (1). By Lemma 11.15 in [18], the triangular array $\{f_{nj}\}$ is AMS provided it is separable, i.e., provided for every $n \in \mathbb{N}$, there exists a countable subset $S_n \subset A_\varepsilon$, such that

$$\mathbb{P}^*\left(\sup_{(\theta_1, z_1)\in A_\varepsilon} \inf_{(\theta_2, z_2)\in S_n} \sum_{j=1}^{n} \big\{ f_{nj}(\omega; \theta_2, z_2) - f_{nj}(\omega; \theta_1, z_1)\big\}^2 > 0 \right) = 0.$$

Define $S_n := (\mathbb{Q}^2 \cap A_\varepsilon) \cup ((\mathbb{Q} \cap [0, 1]) \times \{-\varepsilon\}) \cup ((\mathbb{Q} \cap [0, 1]) \times \{\varepsilon\})$ for all $n \in \mathbb{N}$. Then, for every element ω of the underlying probability space and for every $(\theta_1, z_1) \in A_\varepsilon$, there exists a $(\theta_2, z_2) \in S_n$ such that

$$\sum_{j=1}^{n} \big\{ f_{nj}(\omega; \theta_2, z_2) - f_{nj}(\omega; \theta_1, z_1)\big\}^2 = 0.$$

Proof of (2). The $\{F_{nj}\}$ are obviously independent within rows since Z is a Lévy process. Therefore, according to Theorem 11.17 in [18], it suffices to prove that the triangular arrays

$$\big\{ \tilde{f}_{nj}(\omega; z) := k_n^{-1/2} 1_{\{\Delta_j^n Z \in \mathcal{I}(z)\}} \mid n \in \mathbb{N}; j = 1, \dots, n; |z| \geq \varepsilon \big\},$$

and

$$\big\{ \tilde{g}_{nj}(\omega; \theta) := 1_{\{j \leq \lfloor n\theta \rfloor\}} \mid n \in \mathbb{N}; j = 1, \dots, n; \theta \in [0, 1] \big\}$$

are manageable with envelopes $\{\tilde{F}_{nj}(\omega) := k_n^{-1/2} 1_{\{|\Delta_j^n Z| \geq \varepsilon\}} \mid n \in \mathbb{N}; j = 1, \dots, n\}$ and $\{\tilde{G}_{nj}(\omega) :\equiv 1 \mid n \in \mathbb{N}; j = 1, \dots, n\}$, respectively.

Concerning the first triangular array $\{\tilde{f}_{nj}\}$ define, for $\omega \in \Omega$ and $n \in \mathbb{N}$,

$$\mathcal{F}_{n\omega} := \big\{ \big(k_n^{-1/2} 1_{\{\Delta_1^n Z(\omega) \in \mathcal{I}(z)\}}, \dots, k_n^{-1/2} 1_{\{\Delta_n^n Z(\omega) \in \mathcal{I}(z)\}}\big) \mid |z| \geq \varepsilon \big\} \subset \mathbb{R}^n.$$

For any $j_1, j_2 \in \{1, \dots, n\}$, the projection $p_{j_1,j_2}(\mathcal{F}_{n\omega})$ of $\mathcal{F}_{n\omega}$ onto the j_1th and the j_2th coordinate is an element of the set

$$\begin{aligned}
&\big\{ \big\{(0,0)\big\}, \big\{(0,0), \big(k_n^{-1/2}, 0\big)\big\}, \big\{(0,0), \big(0, k_n^{-1/2}\big)\big\},\\
&\quad \big\{(0,0), \big(k_n^{-1/2}, k_n^{-1/2}\big)\big\}, \big\{(0,0), \big(k_n^{-1/2}, 0\big), \big(0, k_n^{-1/2}\big)\big\},\\
&\quad \big\{(0,0), \big(k_n^{-1/2}, 0\big), \big(k_n^{-1/2} k_n^{-1/2}\big)\big\}, \big\{(0,0), \big(0, k_n^{-1/2}\big), \big(k_n^{-1/2}, k_n^{-1/2}\big)\big\}\big\}.
\end{aligned}$$

Hence, for every $t \in \mathbb{R}^2$, no proper coordinate projection of $\mathcal{F}_{n\omega}$ can surround t in the sense of Definition 4.2 of [22]. Thus, $\mathcal{F}_{n\omega}$ is a subset of $\mathbb{R}^n$ of pseudodimension at most 1 (Definition 4.3 in [22]). Additionally, $\mathcal{F}_{n\omega}$ is a bounded set, whence Corollary 4.10 in [22] yields the existence of constants A and W, depending only on the pseudodimension, such that

$$D_2\big(x\big\|\alpha \odot \tilde{F}_n(\omega)\big\|_2, \alpha \odot \mathcal{F}_{n\omega}\big) \le Ax^{-W} =: \lambda(x),$$

for all $0 < x \le 1$, for every rescaling vector $\alpha \in \mathbb{R}^n$ with non-negative entries and for all $\omega \in \Omega$ and $n \in \mathbb{N}$. Therein, $\|\cdot\|_2$ denotes the Euclidean distance, D_2 denotes the packing number with respect to the Euclidean distance, $\odot$ denotes coordinatewise multiplication and $\tilde{F}_n(\omega) := (\tilde{F}_{n1}(\omega), \dots, \tilde{F}_{nn}(\omega)) \in \mathbb{R}^n$ is the vector of envelopes. Since $\int_0^1 \sqrt{\log \lambda(x)}\,dx < \infty$, the triangular array $\{\tilde{f}_{nj}\}$ is indeed manageable with envelopes $\{\tilde{F}_{nj}\}$.

Concerning the triangular array $\{\tilde{g}_{ni}\}$, we proceed similar and consider the set

$$\begin{aligned}\mathcal{G}_{n\omega} &:= \big\{\big(\tilde{g}_{n1}(\omega;\theta), \dots, \tilde{g}_{nn}(\omega;\theta)\big) \mid \theta \in [0,1]\big\} \\ &= \big\{(0,\dots,0), (1,0,\dots,0), (1,1,0,\dots,0), \dots, (1,\dots,1)\big\}.\end{aligned} \tag{A.2}$$

Then, for any $j_1, j_2 \in \{1,\dots,n\}$, the projection $p_{j_1,j_2}(\mathcal{G}_{n\omega})$ of $\mathcal{G}_{n\omega}$ onto the j_1th and the j_2th coordinate is either $\{(0,0),(1,0),(1,1)\}$ or $\{(0,0),(0,1),(1,1)\}$. Therefore, the same reasoning as above shows that $\mathcal{G}_{n\omega}$ is a set of pseudodimension at most one, whence the triangular array $\{\tilde{g}_{nj}\}$ is manageable with envelopes $\{\tilde{G}_{nj}\}$.

Proof of (3). For any $(\theta_1, z_1), (\theta_2, z_2) \in A_\varepsilon$, by independence of $\{f_{nj}\}$ within rows, we can write

$$\begin{aligned}&\mathbb{E}\big\{\mathbb{G}_n^\circ(\theta_1,z_1)\mathbb{G}_n^\circ(\theta_2,z_2)\big\} \\ &\quad= \sum_{j=1}^n \mathbb{E}\big[\big\{f_{nj}(\omega;\theta_1,z_1) - \mathbb{E}f_{nj}(\cdot;\theta_1,z_1)\big\}\big\{f_{nj}(\omega;\theta_2,z_2) - \mathbb{E}f_{nj}(\cdot;\theta_2,z_2)\big\}\big] \\ &\quad= \frac{1}{k_n}\sum_{j=1}^{\lfloor n(\theta_1\wedge\theta_2)\rfloor}\big\{\mathbb{P}\big(\Delta_j^n Z \in \mathcal{I}(z_1)\cap\mathcal{I}(z_2)\big) - \mathbb{P}\big(\Delta_j^n Z \in \mathcal{I}(z_1)\big)\mathbb{P}\big(\Delta_j^n Z \in \mathcal{I}(z_2)\big)\big\}.\end{aligned} \tag{A.3}$$

By Lemma 2.4, we have

$$\mathbb{P}\big(\Delta_j^n Z \in \mathcal{I}(z)\big) = \Delta_n \nu\big(\mathcal{I}(z)\big) + O\big(\Delta_n^2\big), \qquad n \to \infty \tag{A.4}$$

for all $|z| \ge \varepsilon$ and all $j = 1,\dots,n$, whence the right-hand side of equation (A.3) can be written as

$$\frac{\lfloor n(\theta_1\wedge\theta_2)\rfloor}{n}\big\{\nu\big(\mathcal{I}(z_1)\cap\mathcal{I}(z_2)\big) + O(\Delta_n)\big\} = H(\theta_1,z_1;\theta_2,z_2) + o(1), \qquad n \to \infty.$$

Proof of (4). Recall that $M_\varepsilon = (-\infty, -\varepsilon] \cup [\varepsilon, \infty)$. Again from (A.4), we have, as $n \to \infty$,

$$\sum_{j=1}^{n} \mathbb{E} F_{nj}^2 = \frac{1}{n\Delta_n} \sum_{j=1}^{n} \mathbb{P}\big(\big|\Delta_j^n Z\big| \geq \varepsilon\big) = \nu(M_\varepsilon) + O(\Delta_n) \to \nu(M_\varepsilon) < \infty.$$

Proof of (5). For $\eta > 0$ define $N := \min\{n \in \mathbb{N} \mid k_m^{-1/2} \leq \eta \text{ for all } m \geq n\}$. Choose $K = K(\varepsilon)$ as in Lemma 2.4. Then we have

$$\begin{aligned} \sum_{j=1}^{n} \mathbb{E} F_{nj}^2 \mathbb{1}_{\{F_{nj} > \eta\}} &\leq \sum_{j=1}^{N} \mathbb{E} F_{nj}^2 = \frac{1}{n\Delta_n} \sum_{j=1}^{N} \mathbb{P}\big(\Delta_j^n Z \in M_\varepsilon\big) \\ &\leq \frac{N}{n} \big\{\nu(M_\varepsilon) + K\Delta_n\big\} \to 0, \qquad n \to \infty. \end{aligned}$$

Proof of (6). For $(\theta_1, z_1), (\theta_2, z_2) \in A_\varepsilon$, we can write

$$\begin{aligned} &\rho_n^2(\theta_1, z_1; \theta_2, z_2) \\ &\quad= \sum_{j=1}^{n} \mathbb{E}\big| f_{nj}(\cdot; \theta_1, z_1) - f_{nj}(\cdot; \theta_2, z_2)\big|^2 \\ &\quad= \frac{1}{n\Delta_n} \left\{ \sum_{j=1}^{\lfloor n(\theta_1 \wedge \theta_2) \rfloor} \mathbb{P}\big(\Delta_j^n Z \in \mathcal{I}(z_1) \Delta \mathcal{I}(z_2)\big) + \sum_{j=\lfloor n(\theta_1 \wedge \theta_2) \rfloor + 1}^{\lfloor n(\theta_1 \vee \theta_2) \rfloor} \mathbb{P}\big(\Delta_j^n Z \in \mathcal{I}(z_{L(\theta_1, \theta_2)})\big) \right\} \\ &\quad= \big\{(\theta_1 \wedge \theta_2) + O\big(n^{-1}\big)\big\} \times \big\{\nu\big(\mathcal{I}(z_1) \Delta \mathcal{I}(z_2)\big) + O(\Delta_n)\big\} \\ &\qquad+ \big\{|\theta_1 - \theta_2| + O\big(n^{-1}\big)\big\} \times \big\{\nu\big(\mathcal{I}(z_{L(\theta_1, \theta_2)})\big) + O(\Delta_n)\big\} \end{aligned}$$

as $n \to \infty$, where the O-terms are uniform in $(\theta_1, z_1), (\theta_2, z_2) \in A_\varepsilon$ for the same reason as in equation (A.4). Thus, ρ_n^2 converges uniformly on each $A_\varepsilon \times A_\varepsilon$ to ρ^2. Consequently, for any two sequences $(\theta_1^{(n)}, z_1^{(n)})_{n \in \mathbb{N}}, (\theta_2^{(n)}, z_2^{(n)})_{n \in \mathbb{N}} \subset A_\varepsilon$ such that $\rho(\theta_1^{(n)}, z_1^{(n)}; \theta_2^{(n)}, z_2^{(n)}) \to 0$, it follows that $\rho_n(\theta_1^{(n)}, z_1^{(n)}; \theta_2^{(n)}, z_2^{(n)}) \to 0$.

Finally, ρ is a semimetric: applying first the triangle inequality in $\mathbb{R}^n$ and then the Minkowski inequality, one sees that each ρ_n satisfies the triangle inequality. Thus the triangle inequality also holds for ρ.

It remains to be shown that $\|\mathbb{G}_n - \mathbb{G}_n^\circ\|_{A_\varepsilon} = o_{\mathbb{P}}(1)$. Let $U_n^\circ(\theta, z)$ denote the quantity defined in (2.3) based on the increments $\Delta_j^n Z$. Then

$$\big\|\mathbb{G}_n - \mathbb{G}_n^\circ\big\|_{A_\varepsilon} \leq \sqrt{k_n}\big\|U_n - U_n^\circ\big\|_{A_\varepsilon} + \sqrt{k_n}\big\|\mathbb{E}\big|U_n - U_n^\circ\big|\big\|_{A_\varepsilon}, \tag{A.5}$$

and it suffices to treat both terms separately.

Let $p > 2$ and $0 < \tau < 1$ be the constants of Condition 2.1 and let $v_n := \Delta_n^{\tau/2}$. Distinguishing the cases $|\Delta_j^n Y| \geq v_n$ and $|\Delta_j^n Y| < v_n$, we get that

$$\begin{aligned}\left|U_n(\theta,z) - U_n^\circ(\theta,z)\right| &\leq k_n^{-1}\sum_{j=1}^n 1_{\{|\Delta_j^n Y|\geq v_n\}} + k_n^{-1}\sum_{j=1}^n 1_{\{\Delta_j^n Z\in(z-v_n,z+v_n)\}} \\ &\leq k_n^{-1}\sum_{j=1}^n 1_{\{|\Delta_j^n Y|\geq v_n\}} + k_n^{-1}\sum_{j=1}^n 1_{\{\Delta_j^n Z\in[z-v_n,z+v_n]\}} \\ &=: S_{n1} + S_{n2}(z)\end{aligned} \tag{A.6}$$

for any $(\theta, z) \in A_\varepsilon$. In the following, let $K > 0$ denote a generic constant whose value may change from line to line. By Hölder's inequality and the Burkholder–Davis–Gundy inequalities (see, for instance, page 39 in [14]) we have, for each $1 \leq j \leq n$,

$$\mathbb{E}\left|\int_{(j-1)\Delta_n}^{j\Delta_n} b_s\,ds\right|^p \leq \Delta_n^p \mathbb{E}\left(\frac{1}{\Delta_n}\int_{(j-1)\Delta_n}^{j\Delta_n} |b_s|^p\,ds\right) \leq \Delta_n^p \mathbb{E}S^p \leq K\Delta_n^p$$

and

$$\mathbb{E}\left|\int_{(j-1)\Delta_n}^{j\Delta_n} \sigma_s\,dW_s\right|^p \leq K\Delta_n^{p/2}\mathbb{E}\left(\frac{1}{\Delta_n}\int_{(j-1)\Delta_n}^{j\Delta_n} |\sigma_s|^2\,ds\right)^{p/2} \leq K\Delta_n^{p/2}\mathbb{E}S^p \leq K\Delta_n^{p/2},$$

where S is the bound on the coefficients in Condition 2.1(a). By Markov's inequality and the choice of τ in Condition 2.1(d), we get that

$$\sqrt{k_n}\mathbb{E}S_{n1} \leq K k_n^{-1/2} n \Delta_n^{p/2 - p\tau/2} = O\left(\left(n\Delta_n^{1+\tau}\right)^{1/2}\right) = o(1). \tag{A.7}$$

Similarly, by Lemma 2.4 and Condition 2.1(c),

$$\sup_{|z|\geq\varepsilon} \sqrt{k_n}\mathbb{E}S_{n2}(z) \leq K\left(k_n^{-1/2} n\Delta_n v_n + k_n^{-1/2} n\Delta_n^2\right) = O\left(\left(n\Delta_n^{1+\tau}\right)^{1/2}\right) = o(1). \tag{A.8}$$

Hence, by (A.6), the second summand on the right of (A.5) is $o(1)$.

Consider the first summand on the right of (A.5). From the right-hand side of (A.6), we get that

$$\left\|U_n - U_n^\circ\right\|_{A_\varepsilon} \leq S_{n1} + \sup_{|z|\geq\varepsilon} k_n^{-1/2}\left|\mathbb{G}_n^\circ(1, z - v_n) - \mathbb{G}_n^\circ(1, z+v_n)\right| + \sup_{|z|\geq\varepsilon}\mathbb{E}S_{n2}(z). \tag{A.9}$$

This expression is $o_{\mathbb{P}}(k_n^{-1/2})$ by the previous two displays and by Theorem 1.5.7 and its addendum in [24], applied to the process $\mathbb{G}_n^\circ(1,\cdot)$. The latter converges weakly by the first part of this proof. $\square$

A.2. Proof of Corollary 2.5

By Lemma 1.10.2(i) in [24] and the first part of the proof of Theorem 2.3, it suffices to show that $\|\mathbb{G}_n^\circ - \tilde{\mathbb{G}}_n\|_{A_\varepsilon} = o_{\mathbb{P}}(1)$. Clearly, for any $(\theta, z) \in A_\varepsilon$,

$$\big|\mathbb{G}_n^\circ(\theta, z) - \tilde{\mathbb{G}}_n(\theta, z)\big| \le \sqrt{k_n}\big|U_n(\theta, z) - U_n^\circ(\theta, z)\big| + \sqrt{k_n}\big|\mathbb{E}U_n^\circ(\theta, z) - \theta\nu\big(\mathcal{I}(z)\big)\big|.$$

By (A.9) the first term on the right-hand side of the last equation is a uniform $o_{\mathbb{P}}(1)$. For the second term in the last display, choose $K = K(\varepsilon)$ as in Lemma 2.4. Then

$$\begin{aligned}
&\sqrt{k_n}\big|\mathbb{E}U_n^\circ(\theta, z) - \theta\nu\big(\mathcal{I}(z)\big)\big| \\
&\quad\le \sqrt{k_n}\left|\frac{1}{n}\sum_{j=1}^{\lfloor n\theta\rfloor}\big\{\Delta_n^{-1}\mathbb{P}\big(\Delta_j^n Z \in \mathcal{I}(z)\big) - \nu\big(\mathcal{I}(z)\big)\big\}\right| + \sqrt{k_n}\nu\big(\mathcal{I}(z)\big)\left|\frac{\lfloor n\theta\rfloor}{n} - \theta\right| \qquad \text{(A.10)}\\
&\quad\le K\sqrt{k_n}\Delta_n + \nu(M_\varepsilon)\sqrt{\Delta_n/n} \to 0.
\end{aligned}$$

The convergence is uniform in $(\theta, z) \in A_\varepsilon$ and this yields the assertion. □

A.3. Proof of Theorem 2.6

We use the extended continuous mapping theorem (Theorem 1.11.1 in [24]). For $n \in \mathbb{N}_0$, define $g_n : \ell^\infty(A_\varepsilon) \to \ell^\infty(A_\varepsilon)$ through

$$g_n(f)(\theta, z) = f(\theta, z) - \frac{\lfloor n\theta\rfloor}{n} f(1, z) \qquad (n \in \mathbb{N}), \qquad g_0(f)(\theta, z) = f(\theta, z) - \theta f(1, z).$$

Note that g_n is Lipschitz continuous for any $n \in \mathbb{N}_0$. Obviously, $\mathbb{T}_n = g_n(\mathbb{G}_n) + \mathbb{E}\mathbb{T}_n$ for each $n \in \mathbb{N}$ and $\mathbb{T} = g_0(\mathbb{G})$. We have

$$\mathbb{E}\mathbb{T}_n(\theta, z) = \sqrt{k_n}\lambda_n(\theta)\left\{\frac{n}{\lfloor n\theta\rfloor}\mathbb{E}U_n(\theta, z) - \frac{n}{n - \lfloor n\theta\rfloor}\big[\mathbb{E}U_n(1, z) - \mathbb{E}U_n(\theta, z)\big]\right\}.$$

Observing that (A.10) and (A.9) together with (A.7) as well as (A.8) imply that

$$\sqrt{k_n}\big|\mathbb{E}U_n(\theta, z) - \theta\nu\big(\mathcal{I}(z)\big)\big| \to 0$$

in $\ell^\infty(A_\varepsilon)$, we can conclude that also $\mathbb{E}\mathbb{T}_n$ converges to 0 in $\ell^\infty(A_\varepsilon)$. Thus, by Slutsky's theorem ([24], Example 1.4.7), it suffices to verify $g_n(\mathbb{G}_n) \rightsquigarrow g_0(\mathbb{G})$.

Due to Theorem 1.11.1 in [24] (note that $\mathbb{G}$ is separable as it is tight; see Lemma 1.3.2 in the last-named reference) this weak convergence is valid, if we can show that, for any sequence $(f_n)_{n\in\mathbb{N}} \subset \ell^\infty(A_\varepsilon)$ with $f_n \to f_0$ for some $f_0 \in \ell^\infty(A_\varepsilon)$, we have $g_n(f_n) \to g_0(f_0)$. This can be established by the following calculation:

$$\begin{aligned}
\big\|g_n(f_n) - g_0(f_0)\big\|_{A_\varepsilon} &= \big\|f_n(\theta, z) - \big(\lfloor n\theta\rfloor/n\big)f_n(1, z) - f_0(\theta, z) + \theta f_0(1, z)\big\|_{A_\varepsilon} \\
&\le n^{-1}\|f_0\|_{A_\varepsilon} + 2\|f_n - f_0\|_{A_\varepsilon}.
\end{aligned}$$

Obviously, $\mathbb{T}$ is a tight, mean-zero Gaussian process. Moreover, from Theorem 2.3,

$$\begin{aligned}\operatorname{Cov}\{\mathbb{T}(\theta_1, z_1), \mathbb{T}(\theta_2, z_2)\} &= H(\theta_1, z_1; \theta_2, z_2) - \theta_1 H(1, z_1; \theta_2, z_2) \\ &\quad - \theta_2 H(\theta_1, z_1; 1, z_2) + \theta_1\theta_2 H(1, z_1; 1, z_2) \\ &= \{(\theta_1 \wedge \theta_2) - \theta_1\theta_2\}\nu(\mathcal{I}(z_1) \cap \mathcal{I}(z_2))\end{aligned}$$

for any $(\theta_1, z_1), (\theta_2, z_2) \in A_\varepsilon$. □

A.4. Proof of Proposition 2.8

Because of Corollary 2.5 (and the continuous mapping theorem) $U_{1:n}(z_0) = U_n(1, z_0)$ converges to $\nu(\mathcal{I}(z_0)) > 0$ in probability. Therefore, it follows easily that the random variable $\{U_n(1, z_0)\}^{-1/2}\mathbb{1}_{\{U_n(1,z_0)>0\}}$ converges to $\{\nu(\mathcal{I}(z_0))\}^{-1/2}$ in probability. Hence, by Slutsky's theorem ([24], Example 1.4.7) we obtain

$$\mathbb{V}_n^{(z_0)}(\theta) \rightsquigarrow \frac{1}{\sqrt{\nu(\mathcal{I}(z_0))}}\mathbb{T}(\theta, z_0)$$

in $\ell^\infty([0, 1])$. By Theorem 2.6, the process on the right-hand side of this display is a tight mean zero Gaussian with covariance function $k(\theta_1, \theta_2) = \theta_1 \wedge \theta_2 - \theta_1\theta_2$. Thus, the law of that process is the law of a standard Brownian bridge on $\ell^\infty([0, 1])$. □

A.5. Proof of Theorem 3.3

Recall the decomposition $X_t = Y_t + Z_t$ prior to (A.1) and the triangular array $\{f_{nj}(\omega; \theta, z) \mid n \geq 1; j = 1, \dots, n; (\theta, z) \in A_\varepsilon\}$ consisting of the processes

$$f_{nj}(\omega; \theta, z) := k_n^{-1/2}\mathbb{1}_{\{j \leq \lfloor n\theta\rfloor\}}\mathbb{1}_{\{\Delta_j^n Z \in \mathcal{I}(z)\}}.$$

Set $\mu_{nj}(\theta, z) := \mathbb{E} f_{nj}(\cdot; \theta, z) = k_n^{-1/2}\mathbb{1}_{\{j \leq \lfloor n\theta\rfloor\}}\mathbb{P}(\Delta_j^n Z \in \mathcal{I}(z))$ and let

$$\hat{\mu}_{nj}(\theta, z) := \hat{\mu}_{nj}(\omega; \theta, z) := k_n^{-1/2}\mathbb{1}_{\{j \leq \lfloor n\theta\rfloor\}}\eta_n^\circ(z)$$

be an estimator for $\mu_{nj}(\theta, z)$, where $\eta_n^\circ(z) := n^{-1}\sum_{i=1}^n \mathbb{1}_{\{\Delta_i^n Z \in \mathcal{I}(z)\}}$.

First, we want to show the assertion of Theorem 3.3 for $\hat{\mathbb{G}}_n^\circ$, the process being defined exactly as $\hat{\mathbb{G}}_n$ but based on the increments $\Delta_j^n Z$. This process can be written as

$$\hat{\mathbb{G}}_n^\circ(\theta, z) = \hat{\mathbb{G}}_n^\circ(\omega; \theta, z) = \sum_{j=1}^n \xi_j\{f_{nj}(\omega; \theta, z) - \hat{\mu}_{nj}(\omega; \theta, z)\}.$$

Due to Theorem 3 in [17] the proof for $\hat{\mathbb{G}}_n^\circ$ is complete, if we show the following properties for the triangular array $\{\hat{\mu}_{nj}(\omega; \theta, z) \mid n \geq 1; j = 1, \dots, n; (\theta, z) \in A_\varepsilon\}$:

(i) $\{\hat{\mu}_{nj}\}$ is almost measurable Suslin.

(ii) $\sup_{(\theta,z)\in A_\varepsilon} \sum_{j=1}^n \{\hat{\mu}_{nj}(\omega;\theta,z) - \mu_{nj}(\theta,z)\}^2 \xrightarrow{\mathbb{P}^*} 0$.

(iii) The triangular array $\{\hat{\mu}_{nj}\}$ is manageable with envelopes $\{\hat{F}_{nj}\}$ given through $\hat{F}_{nj}(\omega) := k_n^{-1/2} n^{-1} \sum_{i=1}^n \mathbb{1}_{\{|\Delta_i^n Z| \geq \varepsilon\}}$.

(iv) There exists a constant $M < \infty$ such that $M \vee \sum_{j=1}^n \hat{F}_{nj}^2 \xrightarrow{\mathbb{P}^*} M$.

Proof of (i). As in the proof of (1) in Theorem 2.3, it suffices to verify that the triangular array $\{\hat{\mu}_{nj}\}$ is separable. This can be seen by taking $S_n := (\mathbb{Q}^2 \cap A_\varepsilon) \cup ((\mathbb{Q} \cap [0,1]) \times \{-\varepsilon\}) \cup ((\mathbb{Q} \cap [0,1]) \times \{\varepsilon\})$.

Proof of (ii). We have

$$\begin{aligned}
&\sup_{(\theta,z)\in A_\varepsilon} \sum_{j=1}^n \bigl\{\hat{\mu}_{nj}(\omega;\theta,z) - \mu_{nj}(\theta,z)\bigr\}^2 \\
&\quad = \sup_{|z|\geq\varepsilon} n^{-3}\Delta_n^{-1} \sum_{j=1}^n \Biggl[\sum_{i=1}^n \bigl\{\mathbb{1}_{\{\Delta_i^n Z \in \mathcal{I}(z)\}} - \mathbb{P}\bigl(\Delta_j^n Z \in \mathcal{I}(z)\bigr)\bigr\}\Biggr]^2 \\
&\quad = n^{-1} \sup_{|z|\geq\varepsilon} \bigl\{\mathbb{G}_n^\circ(1,z)\bigr\}^2.
\end{aligned}$$

The last quantity in the above display converges to 0 in probability because in the proof of Theorem 2.3 we have seen $\mathbb{G}_n^\circ \rightsquigarrow \mathbb{G}$.

Proof of (iii). Following (A.2), we have already shown that the triangular array

$$\bigl\{\tilde{g}_{nj}(\theta) := \tilde{g}_{nj}(\omega;\theta) := \mathbb{1}_{\{j \leq \lfloor n\theta \rfloor\}} \mid n \in \mathbb{N};\ j = 1,\dots,n;\ \theta \in [0,1]\bigr\}$$

is manageable with envelopes $\{\tilde{G}_{nj}(\omega) \stackrel{\text{def}}{\equiv} 1 \mid n \in \mathbb{N};\ j = 1,\dots,n\}$. Therefore, due to Theorem 11.17 in [18], it suffices to prove that the triangular array

$$\left\{\tilde{h}_{nj}(\omega;z) := \frac{1}{n\sqrt{k_n}} \sum_{i=1}^n \mathbb{1}_{\{\Delta_i^n Z \in \mathcal{I}(z)\}} \,\middle|\, n \in \mathbb{N};\ j = 1,\dots,n;\ |z| \geq \varepsilon\right\}$$

is manageable with envelopes $\{\hat{F}_{nj}(\omega) \mid n \in \mathbb{N};\ j = 1,\dots,n\}$. But $\tilde{h}_{nj}(\omega;z)$ does not depend on j at all, such that every projection of $\mathcal{H}_{n\omega} := \{(\tilde{h}_{n1}(\omega;z),\dots,\tilde{h}_{nn}(\omega;z)) \mid |z| \geq \varepsilon\}$ onto two coordinates lies in the straight line $\{(x,y) \in \mathbb{R}^2 \mid x = y\}$. Consequently, the set $\mathcal{H}_{n\omega}$ has a pseudodimension of at most 1 and is bounded. Hence, the same arguments as in the proof of (2) in Theorem 2.3 show the desired manageability.

Proof of (iv). A straight forward calculation using (A.4) yields

$$\mathbb{E}\left\{\sum_{j=1}^{n} \hat{F}_{nj}^2\right\} = n^{-2}\Delta_n^{-1} \sum_{i_1=1}^{n}\sum_{i_2=1}^{n} \mathbb{E}\{1_{\{|\Delta_{i_1}^n Z|\geq\varepsilon\}} 1_{\{|\Delta_{i_2}^n Z|\geq\varepsilon\}}\} = O(\Delta_n).$$

Thus $\sum_{j=1}^n \hat{F}_{nj}^2$ is $o_{\mathbb{P}}(1)$.

So far, we have established $\hat{\mathbb{G}}_n^\circ \rightsquigarrow_\xi \mathbb{G}$, and due to Lemma A.1 it suffices to show $\|\hat{\mathbb{G}}_n - \hat{\mathbb{G}}_n^\circ\|_{A_\varepsilon} = o_{\mathbb{P}}(1)$ in order to finish the proof. This can be done following the lines of the proof of Theorem 2.3. The only difference regards showing that

$$\sup_{|z|\geq\varepsilon} \frac{1}{\sqrt{k_n}} \sum_{j=1}^{n} |\xi_j| 1_{\{\Delta_j^n Z\in(z-v_n,z+v_n)\}} \tag{A.11}$$

is $o_{\mathbb{P}}(1)$. Arguing as in (A.9),

$$\begin{aligned}
&\sup_{|z|\geq\varepsilon} \frac{1}{\sqrt{k_n}} \sum_{i=1}^{n} 1_{\{\Delta_i^n Z\in(z-v_n,z+v_n)\}} \\
&\quad \leq \sup_{|z|\geq\varepsilon} \left|\mathbb{G}_n^\circ(1, z-v_n) - \mathbb{G}_n^\circ(1, z+v_n)\right| + o_{\mathbb{P}}(1) = o_{\mathbb{P}}(1).
\end{aligned}$$

Therefore with the strong law of large numbers the expression inside of the supremum in (A.11) can be written as

$$\begin{aligned}
&\frac{1}{\sqrt{k_n}} \sum_{j=1}^{n} \big\{|\xi_j| - \mathbb{E}|\xi_j|\big\} 1_{\{\Delta_j^n Z\in(z-v_n,z+v_n)\}} + o_{\mathbb{P}}(1) \\
&\quad = \frac{1}{\sqrt{k_n}} \sum_{j=1}^{n} \big\{|\xi_j| - \mathbb{E}|\xi_j|\big\} \left\{1_{\{\Delta_j^n Z\in(z-v_n,z+v_n)\}} - \frac{1}{n}\sum_{i=1}^{n} 1_{\{\Delta_i^n Z\in(z-v_n,z+v_n)\}}\right\} + o_{\mathbb{P}}(1),
\end{aligned}$$

where the $o_{\mathbb{P}}(1)$-terms are uniform in $|z| \geq \varepsilon$. After a standardization of the new multipliers $\xi_j^* := |\xi_j| - \mathbb{E}|\xi_j|$, the main term in the above display can be seen to be $o_{\mathbb{P}}(1)$ by asymptotic uniform ρ-equicontinuity in probability as a consequence of $\hat{\mathbb{G}}_n^\circ \rightsquigarrow_\xi \mathbb{G}$ and Remark 3.2(ii), see Theorem 1.5.7 and its addendum in [24] again. □

A.6. Proof of Theorem 3.4

This result follows by the same lines as in the proof of Theorem 2.6, with an application of Proposition 10.7(i) in [18]. □

A.7. Proof of Proposition 4.1

The assertion involving $V_n^{(z_0)}$ is trivial. Regarding $W_n^{(z_0)}$, note that Proposition A.2 and the continuous mapping theorem imply that, for any fixed $B \in \mathbb{N}$,

$$\big(W_n^{(z_0)}, \hat{W}_{n,\xi^{(1)}}^{(z_0)}, \dots, \hat{W}_{n,\xi^{(B)}}^{(z_0)}\big) \rightsquigarrow \big(W^{(z_0)}, W^{(z_0),(1)}, \dots, W^{(z_0),(B)}\big)$$

in $\mathbb{R}^{B+1}$, where $W^{(z_0)} := \sup_{\theta\in[0,1]} |\mathbb{T}(\theta, z_0)|$ with the limit process $\mathbb{T}$ of Theorem 2.6 and where $W^{(z_0),(1)}, \dots, W^{(z_0),(B)}$ are independent copies of $W^{(z_0)}$. According to the corollary to Proposition 3 in [20], $W^{(z_0)}$ has a continuous c.d.f. since $\nu(\mathcal{I}(z_0)) > 0$. Thus, Proposition F.1 in the supplement to [6] implies that

$$\lim_{B\to\infty} \lim_{n\to\infty} \mathbb{P}\big\{W_n^{(z_0)} \geq \hat{q}_{1-\alpha}^{(B)}\big(W_n^{(z_0)}\big)\big\} = \alpha$$

for all $\alpha \in (0, 1)$, as asserted. A similar reasoning gives the claim for $T^{(\varepsilon)}$. □

A.8. Proof of Proposition 4.2

This proof is a simple consequence of the auxiliary Propositions A.3 and A.4. □

A.9. Proof of Proposition 4.3

Let $X^{(1)}(n)$ and $X^{(2)}(n)$ denote two independent Itô semimartingales with characteristics $(b_t^{(n)}, \sigma_t^{(n)}, \nu_1)$ and $(b_t^{(n)}, \sigma_t^{(n)}, \nu_2)$, respectively. For $n \in \mathbb{N}$ and $j = 0, \dots, n$, set $Y_j(n) = X_{j\Delta_n}^{(1)}(n)$ and $Z_j(n) = X_{j\Delta_n}^{(2)}(n)$. Let $U_n^{(1)}$ and $U_n^{(2)}$ denote the quantity defined in (2.3), based on the observations $Y_j(n)$ and $Z_j(n)$, respectively, instead on $X_{j\Delta_n}$. Moreover, define a random element S_n with values in $\ell^\infty(A_\varepsilon)$ through

$$\begin{aligned} S_n(\theta, z) := {} & \frac{n - \lfloor n\theta\rfloor}{n} U_n^{(1)}(\theta, z) - \frac{\lfloor n\theta\rfloor}{n}\big\{U_n^{(1)}(\theta_0, z) - U_n^{(1)}(\theta, z)\big\} \\ & - \frac{\lfloor n\theta\rfloor}{n}\big\{U_n^{(2)}(1, z) - U_n^{(2)}(\theta_0, z)\big\}, \end{aligned}$$

for $(\theta, z) \in A_\varepsilon$ with $\theta \leq \theta_0$, whereas for $(\theta, z) \in A_\varepsilon$ with $\theta \geq \theta_0$,

$$\begin{aligned} S_n(\theta, z) := {} & \frac{n - \lfloor n\theta\rfloor}{n} U_n^{(1)}(\theta_0, z) + \frac{n - \lfloor n\theta\rfloor}{n}\big\{U_n^{(2)}(\theta, z) - U_n^{(2)}(\theta_0, z)\big\} \\ & - \frac{\lfloor n\theta\rfloor}{n}\big\{U_n^{(2)}(1, z) - U_n^{(2)}(\theta, z)\big\}. \end{aligned}$$

Obviously we have the distributional equality

$$\begin{aligned} & \big(\Delta_1^n X(n), \dots, \Delta_{\lfloor n\theta_0\rfloor}^n X(n), \Delta_{\lfloor n\theta_0\rfloor+1}^n X(n), \dots, \Delta_n^n X(n)\big) \\ & \stackrel{\mathcal{D}}{=} \big(\Delta_1^n X^{(1)}(n), \dots, \Delta_{\lfloor n\theta_0\rfloor}^n X^{(1)}(n), \Delta_{\lfloor n\theta_0\rfloor+1}^n X^{(2)}(n), \dots, \Delta_n^n X^{(2)}(n)\big). \end{aligned}$$

Hence, for any $(\theta_1, z_1), \dots, (\theta_g, z_g) \in A_\varepsilon$ and $g \in \mathbb{N}$, we also have that

$$\big(k_n^{-1/2}\mathbb{T}_n(\theta_1, z_1), \dots, k_n^{-1/2}\mathbb{T}_n(\theta_g, z_g)\big) \stackrel{\mathcal{D}}{=} \big(S_n(\theta_1, z_1), \dots, S_n(\theta_g, z_g)\big).$$

Now, from the previous display, and from the fact that the function T is continuous in (θ, z) and that the functions $\mathbb{T}_n(\theta, z)$ depend only through $\lfloor n\theta \rfloor$ on θ and are either left-continuous or right-continuous in z, we immediately get that

$$\begin{aligned} \sup_{(\theta,z)\in A_\varepsilon} \big|k_n^{-1/2}\mathbb{T}_n(\theta, z) - T(\theta, z)\big| &= \sup_{(\theta,z)\in A_\varepsilon\cap\mathbb{Q}^2} \big|k_n^{-1/2}\mathbb{T}_n(\theta, z) - T(\theta, z)\big| \\ &\stackrel{\mathcal{D}}{=} \sup_{(\theta,z)\in A_\varepsilon\cap\mathbb{Q}^2} \big|S_n(\theta, z) - T(\theta, z)\big|. \end{aligned}$$

This expression is in fact $o_{\mathbb{P}}(1)$ as a consequence of Corollary 2.5 and the continuous mapping theorem. Note that Corollary 2.5 is in fact applicable in this setup, because under $\mathbf{H}_1$ the characteristics $b_t^{(n)}$ and $\sigma_t^{(n)}$ have a uniform bound S in $n \in \mathbb{N}$ and $t \ge 0$. □

A.10. Proof of Proposition 4.4

Under $\mathbf{H}_1$, choose $\varepsilon > 0$ such that there exists a $|z_0| \ge \varepsilon$ with $\nu_1(z_0) \ne \nu_2(z_0)$. Then, according to Proposition 4.3 and the continuous mapping theorem, the random functions $\theta \mapsto \sup_{|z|\ge\varepsilon} |k_n^{-1/2}\mathbb{T}_n(\theta, z)|$ converge weakly in $\ell^\infty([0, 1])$ to the continuous function $\theta \mapsto \sup_{|z|\ge\varepsilon} |T(\theta, z)|$, which has a unique maximum at θ_0. Thus, the asserted convergences follow from the argmax-continuous mapping theorem (Theorem 2.7 in [16]). The claim regarding $\mathbf{H}_1^{(z_0)}$ can be shown similarly. □

A.11. Additional auxiliary results

The first auxiliary result is needed for validating the bootstrap procedures defined in Section 3. It is proved in [4], Lemma A.1.

Lemma A.1. *Consider two bootstrapped statistics $\hat{G}_n = \hat{G}_n(X_1, \dots, X_n, \xi_1, \dots, \xi_n)$ and $\hat{H}_n = \hat{H}_n(X_1, \dots, X_n, \xi_1, \dots, \xi_n)$ in a metric space $(\mathbb{D}, d)$ with $d(\hat{G}_n, \hat{H}_n) \xrightarrow{\mathbb{P}^*} 0$. Then, for a tight Borel measurable process G in $\mathbb{D}$, we have $\hat{G}_n \rightsquigarrow_\xi G$ if and only if $\hat{H}_n \rightsquigarrow_\xi G$.*

The proof of Proposition 4.1 is based on the following auxiliary result, establishing unconditional weak convergence of the vector of processes $(\mathbb{T}_n, \hat{\mathbb{T}}_{n,\xi^{(1)}}, \dots, \hat{\mathbb{T}}_{n,\xi^{(B)}})$.

Proposition A.2. *Suppose the conditions from Theorem 3.3 are met. Then, under $\mathbf{H}_0$, for all $B \in \mathbb{N}$, we have*

$$(\mathbb{T}_n, \hat{\mathbb{T}}_{n,\xi^{(1)}}, \dots, \hat{\mathbb{T}}_{n,\xi^{(B)}}) \rightsquigarrow \big(\mathbb{T}, \mathbb{T}^{(1)}, \dots, \mathbb{T}^{(B)}\big)$$

in $(\ell^\infty(A_\varepsilon), \|\cdot\|_{A_\varepsilon})^{B+1}$, *where* $\rightsquigarrow$ *denotes* (*unconditional*) *weak convergence* (*with respect to the probability measure* $\mathbb{P}$), *and where* $\mathbb{T}^{(1)}, \dots, \mathbb{T}^{(B)}$ *are independent copies of* $\mathbb{T}$.

Proof. We are going to apply Corollary 1.4.5 in [24]. Therefore, let $f^{(0)}, f^{(1)}, \dots, f^{(B)} \in \mathrm{BL}_1(\ell^\infty(A_\varepsilon))$. Since $\mathbb{T}_n, \hat{\mathbb{T}}_{n,\xi^{(1)}}, \dots, \hat{\mathbb{T}}_{n,\xi^{(B)}}$ are independent conditional on the data, we have

$$\begin{aligned}&\mathbb{E}_\xi\big\{f^{(0)}(T_n)\cdot f^{(1)}(\hat{T}_{n,\xi^{(1)}})\cdot \ldots \cdot f^{(B)}(\hat{T}_{n,\xi^{(B)}})\big\}\\ &\quad = f^{(0)}(T_n)\cdot \mathbb{E}_\xi f^{(1)}(\hat{T}_{n,\xi^{(1)}})\cdot \ldots \cdot \mathbb{E}_\xi f^{(B)}(\hat{T}_{n,\xi^{(B)}}) =: S_n.\end{aligned}$$

By Definition 3.1 and Theorem 3.4, $\mathbb{E}_\xi f^{(b)}(\hat{T}_{n,\xi^{(b)}})$ converges in outer probability to $\mathbb{E}(f^{(b)}(\mathbb{T}^{(b)})) =: c_b$ for each $b \in \{1, \dots, B\}$. Therefore,

$$S_n \rightsquigarrow c_1 \cdot \ldots \cdot c_B \cdot f^{(0)}(\mathbb{T}) =: \mathcal{S}$$

by using the continuous mapping theorem, Slutsky's lemma and Lemma 1.10.2 in [24] several times.

Choose an $M > 0$ with $|S_n| \vee |\mathcal{S}| \le M$ for all $\omega \in \Omega$, $n \in \mathbb{N}$ and let $g: \mathbb{R} \longrightarrow \mathbb{R}$ be a bounded and continuous function with $g(x) = x$ on $[-M, M]$. Then

$$\begin{aligned}&\mathbb{E}_X^*\big[\mathbb{E}_\xi^*\big\{f^{(0)}(\mathbb{T}_n)\cdot f^{(1)}(\hat{\mathbb{T}}_{n,\xi^{(1)}})\cdot \ldots \cdot f^{(B)}(\hat{\mathbb{T}}_{n,\xi^{(B)}})\big\}\big]\\ &\quad = \mathbb{E}_X^* S_n = \mathbb{E}_X^* g(S_n) \xrightarrow{(1)} \mathbb{E}\big(g(\mathcal{S})\big) = \mathbb{E}\mathcal{S} \qquad \text{(A.12)}\\ &\quad \overset{(2)}{=} \mathbb{E}\big\{f^{(0)}(\mathbb{T})\cdot f^{(1)}\big(\mathbb{T}^{(1)}\big)\cdot \ldots \cdot f^{(B)}\big(\mathbb{T}^{(B)}\big)\big\}.\end{aligned}$$

Note that (1) uses the fact that a coordinate projection on a product probability space is perfect (Lemma 1.2.5 in [24]). Moreover, (2) holds because the limit processes are independent.

By Theorem 2.6, Remark 3.2(ii), Theorem 3.4 and Lemma 1.3.8 and Lemma 1.4.4 in [24] the vector of processes $(\mathbb{T}_n, \hat{\mathbb{T}}_{n,\xi^{(1)}}, \dots, \hat{\mathbb{T}}_{n,\xi^{(B)}})$ is (jointly) asymptotically measurable. Consequently, equation (A.12), Fubini's theorem (Lemma 1.2.6 in [24]) and Corollary 1.4.5 in this reference yield the desired weak convergence. Note that the limit process $(\mathbb{T}, \mathbb{T}^{(1)}, \dots, \mathbb{T}^{(B)})$ is separable because it is tight (Lemma 1.3.2 in the previously mentioned reference). □

Proposition A.3. *Under* $\mathbf{H}_1$, *there exists an* $\varepsilon > 0$ *such that, for all* $K > 0$,

$$\lim_{n\to\infty} \mathbb{P}\big(T_n^{(\varepsilon)} \ge K\big) = 1.$$

If $\mathbf{H}_1^{(z_0)}$ *is true, the same assertion holds for* $V_n^{(z_0)}$ *and* $W_n^{(z_0)}$.

Proof. Choose $\varepsilon > 0$ such that there exists a $|\hat{z}| \ge \varepsilon$ with $\nu_1(\hat{z}) \ne \nu_2(\hat{z})$. Then $c := \sup_{\theta\in[0,1]} \sup_{|z|\ge\varepsilon} |T(\theta, z)| \in (0, \infty)$, with the function T defined in Proposition 4.3. But Proposition 4.3 and the continuous mapping theorem show that $k_n^{-1/2} T_n^{(\varepsilon)} = c + o_{\mathbb{P}}(1)$ and this yields the assertion for $T_n^{(\varepsilon)}$. The same argument implies the claim for $W_n^{(z_0)}$.

Finally, let us prove the claim for $V_n^{(z_0)}$. As in the proof of Proposition 4.3, let $X^{(1)}(n)$ and $X^{(2)}(n)$ be independent Itô semimartingales with characteristics $(b_t^{(n)}, \sigma_t^{(n)}, \nu_1)$ and $(b_t^{(n)}, \sigma_t^{(n)}, \nu_2)$, respectively. For $n \in \mathbb{N}$ and $j = 0, \dots, n$, set $Y_j(n) = X^{(1)}_{j\Delta_n}(n)$ and $Z_j(n) = X^{(2)}_{j\Delta_n}(n)$. Let $U_n^{(1)}$ and $U_n^{(2)}$ denote the quantity defined in (2.3), based on the observations $Y_j(n)$ and $Z_j(n)$, respectively, instead on $X_{j\Delta_n}$.

Then the quantities $V_n^{(z_0)}$ and $W_n^{(z_0)}$ differ only by a factor $A_n^{-1/2} 1_{\{A_n > 0\}}$, with A_n being equal in distribution to

$$U_n^{(1)}(\theta_0, z_0) + U_n^{(2)}(1, z_0) - U_n^{(2)}(\theta_0, z_0).$$

This expression converges to $\theta_0 \nu_1(z_0) + (1 - \theta_0)\nu_2(z_0) > 0$, in probability, which in turn implies the assertion regarding $V_n^{(z_0)}$. $\square$

Proposition A.4. *Under* $\mathbf{H}_1$, *for all* $\varepsilon > 0$ *and all* $b \in \{1, \dots, B\}$,

$$\hat{T}^{(\varepsilon)}_{n,\xi^{(b)}} = \mathrm{O}_{\mathbb{P}}(1), \qquad \text{that is } \lim_{K \to \infty} \limsup_{n \to \infty} \mathbb{P}\big(\hat{T}^{(\varepsilon)}_{n,\xi^{(b)}} > K\big) = 0.$$

Moreover, under $\mathbf{H}_1^{(z_0)}$, *for all* $b \in \{1, \dots, B\}$,

$$\hat{W}^{(z_0)}_{n,\xi^{(b)}} = \mathrm{O}_{\mathbb{P}}(1), \qquad \text{that is } \lim_{K \to \infty} \limsup_{n \to \infty} \mathbb{P}\big(\hat{W}^{(z_0)}_{n,\xi^{(b)}} > K\big) = 0.$$

Proof. Since the results are independent of b, we omit this index throughout the proof. Also note that, for both assertions, it suffices to show that $\sup_{\theta \in [0,1]} \sup_{|z| \geq \varepsilon} |\hat{\mathbb{G}}_n(\theta, z)| = O_{\mathbb{P}}(1)$ under $\mathbf{H}_1$.

For $n \in \mathbb{N}$ and $j = 0, \dots, n$, let $Y_j(n) = X^{(1)}_{j\Delta_n}(n)$ and $Z_j(n) = X^{(2)}_{j\Delta_n}(n)$ be defined as in the proof of Proposition 4.3. Let $U_n^{(1)}$, $\eta_n^{(1)}$ and $U_n^{(2)}$, $\eta_n^{(2)}$ denote the corresponding quantities, based on the observations $Y_j(n)$ and $Z_j(n)$, respectively.

Then, for $\theta \leq \theta_0$, we can write $\hat{\mathbb{G}}_n(\theta, z)$ as

$$\frac{1}{\sqrt{k_n}} \sum_{j=1}^{\lfloor n\theta \rfloor} \xi_j \big\{ 1_{\{\Delta_j^n Y \in \mathcal{I}(z)\}} - \eta_n^{(1)}(z) \big\} + \left\{ \frac{1}{\sqrt{n}} \sum_{j=1}^{\lfloor n\theta \rfloor} \xi_j \right\} \times \big\{ \Delta_n^{-1/2} \big(\eta_n^{(1)}(z) - \eta_n(z) \big) \big\}.$$

The first term of this display is $O_{\mathbb{P}}(1)$, uniformly in $\theta \leq \theta_0$ and $|z| \geq \varepsilon$, by Theorem 3.3 and Remark 3.2(ii). By the classical Donsker theorem, the term in curly brackets on the right-hand side is also $O_{\mathbb{P}}(1)$ uniformly in $\theta \leq \theta_0$. The quantity $\Delta_n^{-1/2} \eta_n^{(1)}(z) = \Delta_n^{1/2} U_n^{(1)}(1, z)$ is $o_{\mathbb{P}}(1)$ uniformly in $|z| \geq \varepsilon$ by Corollary 2.5. Finally, the same argument as in the proof of Proposition 4.3 yields

$$\Delta_n^{-1/2} \sup_{|z| \geq \varepsilon} \big| \eta_n(z) \big| = \sqrt{\Delta_n} \sup_{|z| \geq \varepsilon} \big| U_n^{(1)}(\theta_0, z) + U_n^{(2)}(1, z) - U_n^{(2)}(\theta_0, z) \big| = o_{\mathbb{P}}(1).$$

To conclude,

$$\sup_{\theta \leq \theta_0} \sup_{|z| \geq \varepsilon} \big| \hat{\mathbb{G}}(\theta, z) \big| = O_{\mathbb{P}}(1).$$

The supremum over $\theta > \theta_0$ and $|z| \geq \varepsilon$ can be treated similarly. □

Acknowledgements

The authors would like to thank an unknown referee and an Associate Editor for helpful comments on an earlier version of this manuscript, which led to a substantial improvement of the paper. This work has been supported by the Collaborative Research Center "Statistical modeling of nonlinear dynamic processes" (SFB 823, Teilprojekt A1, A7, C1) of the German Research Foundation (DFG) which is gratefully acknowledged.

References

[1] Aït-Sahalia, Y. and Jacod, J. (2009). Estimating the degree of activity of jumps in high frequency data. *Ann. Statist.* **37** 2202–2244. MR2543690

[2] Aït-Sahalia, Y. and Jacod, J. (2009). Testing for jumps in a discretely observed process. *Ann. Statist.* **37** 184–222. MR2488349

[3] Aït-Sahalia, Y. and Jacod, J. (2014). *High-Frecuency Financial Econometrics*. Princeton: Princeton Univ. Press.

[4] Bücher, A. (2011). Statistical inference for copulas and extremes. Ph.D. thesis, Ruhr-Universität Bochum.

[5] Bücher, A., Hoffmann, M., Vetter, M. and Dette, H. (2014). Nonparametric tests for detecting breaks in the jump behaviour of a time-continuous process. Preprint. Available at arXiv:1412.5376v1.

[6] Bücher, A. and Kojadinovic, I. (2014). A dependent multiplier bootstrap for the sequential empirical copula process under strong mixing. *Bernoulli* **22** 927–968.

[7] Bücher, A. and Vetter, M. (2013). Nonparametric inference on Lévy measures and copulas. *Ann. Statist.* **41** 1485–1515. MR3113819

[8] Cont, R. and Tankov, P. (2004). *Financial Modelling with Jump Processes*. *Chapman & Hall/CRC Financial Mathematics Series*. Boca Raton, FL: Chapman & Hall/CRC. MR2042661

[9] Delbaen, F. and Schachermayer, W. (1994). A general version of the fundamental theorem of asset pricing. *Math. Ann.* **300** 463–520. MR1304434

[10] Figueroa-López, J.E. (2008). Small-time moment asymptotics for Lévy processes. *Statist. Probab. Lett.* **78** 3355–3365. MR2479503

[11] Figueroa-López, J.E. and Houdré, C. (2009). Small-time expansions for the transition distributions of Lévy processes. *Stochastic Process. Appl.* **119** 3862–3889. MR2552308

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

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

[14] Jacod, J. and Protter, P. (2012). *Discretization of Processes*. *Stochastic Modelling and Applied Probability* **67**. Heidelberg: Springer. MR2859096

[15] Jacod, J. and Shiryaev, A.N. (2003). *Limit Theorems for Stochastic Processes*, 2nd ed. Berlin: Springer. MR1943877

[16] Kim, J. and Pollard, D. (1990). Cube root asymptotics. *Ann. Statist.* **18** 191–219. MR1041391

[17] Kosorok, M.R. (2003). Bootstraps of sums of independent but not identically distributed stochastic processes. *J. Multivariate Anal.* **84** 299–318. MR1965224

[18] Kosorok, M.R. (2008). *Introduction to Empirical Processes and Semiparametric Inference*. *Springer Series in Statistics*. New York: Springer. MR2724368

[19] Lee, S., Nishiyama, Y. and Yoshida, N. (2006). Test for parameter change in diffusion processes by cusum statistics based on one-step estimators. *Ann*. *Inst*. *Statist*. *Math*. **58** 211–222. MR2246154

[20] Lifshits, M.A. (1982). Absolute continuity of functionals of "supremum" type for Gaussian processes. *Zap*. *Nauchn*. *Sem*. *S*.*-Peterburg*. *Otdel*. *Mat*. *Inst*. *Steklov*. (*POMI*) **119** 154–166. MR0666093

[21] Mykland, P.A. and Zhang, L. (2012). The econometrics of high-frequency data. In *Statistical Methods for Stochastic Differential Equations*. *Monogr*. *Statist*. *Appl*. *Probab*. **124** (M. Kessler, A. Lindner and M. Sørensen, eds.) 109–190. Boca Raton, FL: CRC Press. MR2976983

[22] Pollard, D. (1990). *Empirical Processes*: *Theory and Applications*. Hayward, CA: IMS. MR1089429

[23] Rüschendorf, L. and Woerner, J.H.C. (2002). Expansion of transition distributions of Lévy processes in small time. *Bernoulli* **8** 81–96. MR1884159

[24] van der Vaart, A.W. and Wellner, J.A. (1996). *Weak Convergence and Empirical Processes*. *Springer Series in Statistics*. New York: Springer. MR1385671

[25] Vetter, M. (2014). Inference on the Lévy measure in case of noisy observations. *Statist*. *Probab*. *Lett*. **87** 125–133. MR3168946

[26] Vostrikova, L. (1981). Detecting disorder in multidimensional random processes. *Sov*. *Math*., *Dokl*. **24** 55–59.

Received December 2014 and revised September 2015