SFB 823

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

Michael Hoffmann

Nr. 4/2018

Discussion Paper

SFB 823

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

Michael Hoffmann*

Ruhr-Universität Bochum

February 23, 2018

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.

Keywords and Phrases: Lévy measure; jump compensator; transition kernel; empirical processes; weak convergence; multiplier bootstrap; change points; gradual changes

AMS Subject Classification: 60F17, 60G51, 62G10, 62M99.

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 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

[1]Ruhr-Universität Bochum, Fakultät für Mathematik, 44780 Bochum, Germany. E-mail: michael.hoffmann@rub.de

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

$$\mathbf{H}_0 : \nu_s(dz) = \nu_0(dz) \tag{1.1}$$

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, B and C.

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

$$N_\rho^{(n)}(\theta, t) = \frac{1}{n\Delta_n} \sum_{i=1}^{\lfloor n\theta \rfloor} \rho(\Delta_i^n X^{(n)}) \mathbb{1}_{(-\infty,t]}(\Delta_i^n X^{(n)}) \mathbb{1}_{\{|\Delta_i^n X^{(n)}| > v_n\}}, \quad (\theta, t) \in [0,1] \times \mathbb{R}, \tag{2.3}$$

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

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

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 6.1, 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 6.8, in which we prove conditional weak convergence in a suitable sense of the bootstrapped version

$$\hat{G}_\rho^{(n)}(\theta,t) = \frac{1}{\sqrt{n\Delta_n}} \sum_{i=1}^{\lfloor n\theta \rfloor} \xi_i \rho\big(\Delta_i^n X^{(n)}\big) \mathbb{1}_{(-\infty,t]}\big(\Delta_i^n X^{(n)}\big) \mathbb{1}_{\{|\Delta_i^n X^{(n)}| > v_n\}}, \quad (\theta,t) \in [0,1] \times \mathbb{R}$$

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) $h_y(z) \leq K|z|^{-(1+\beta)}$ 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 $C_n := \{z \in \mathbb{R} \mid \frac{1}{n} \leq |z| \leq n\}$. Then for each $n \in \mathbb{N}$ there exists a $K_n > 0$ with $h_y(z) \leq K_n$ for each $z \in C_n$ and all $y \in [0,1] \setminus L$.

(3) $h_y(z) \leq K|z|^{-(2p\vee 2)-\epsilon}$ whenever $|z| \geq M$ 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 6.12 and Proposition 6.13 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. Below we give the main assumptions which are sufficient for the convergence results in this paper.

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:

$$t_1 := (1+\tau)^{-1} \quad \text{and} \quad t_2 := ((7\tau+1)/2)^{-1} \wedge 1.$$

Then we have $0 < t_1 < t_2 \leq 1$ 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

$$m_b = \frac{6+10\tau}{3-5\tau} \leq 4 \quad \text{and} \quad m_\sigma = \frac{6+10\tau}{1-5\tau}$$

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

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

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

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

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

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

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* $\mathbb{V}_{\rho,t_0}^{(n)} \rightsquigarrow \mathbb{K} + \bar{\mathbb{V}}_{\rho,t_0}^{(g_1)}$ *in* $\ell^\infty([0,1])$, *where* $\mathbb{K}$ *denotes a standard Brownian bridge and with the deterministic function*

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

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

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

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, $\hat{Y}_n = \hat{Y}_n(X_1, \ldots, X_n, \xi_1, \ldots, \xi_n)$, which depend on random variables $X_1, \ldots, X_n$ defined on a probability space $(\Omega_X, \mathcal{F}_X, \mathbb{P}_X)$ and on random weights $\xi_1, \ldots, \xi_n$ 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 $\hat{Y}_n = \hat{Y}_n(X_1, \ldots, X_n; \xi_1, \ldots, \xi_n) \colon (\Omega, \mathcal{A}, \mathbb{P}) \to \mathbb{D}$ be a random element taking values in some metric space $\mathbb{D}$ depending on some random variables $X_1, \ldots, X_n$ and some random weights $\xi_1, \ldots, \xi_n$. 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 $X_1, X_2, \ldots$ in probability, 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 $X_1, \ldots, X_n$, whereas $\mathrm{BL}_1(\mathbb{D})$ is the space of all real-valued Lipschitz continuous functions f on $\mathbb{D}$ with sup-norm $\|f\|_\mathbb{D} \le 1$ 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) $\Rightarrow$ (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 $\{X_{i\Delta_n}^{(n)} \mid i = 0, 1, \dots, n\}$ 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 $\hat{\mathbb{T}}_\rho^{(n)}$ of the processes $\mathbb{T}_\rho^{(n)}$ are given by

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

In the following theorem we establish conditional weak convergence of $\hat{\mathbb{T}}_\rho^{(n)}$ 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*

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

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

$$\begin{aligned}
\mathbb{E}\{\mathbb{T}_\rho(\theta_1, t_1)\mathbb{T}_\rho(\theta_2, t_2)\} = &\int_0^{\theta_1\wedge\theta_2}\int_{-\infty}^{t_1\wedge t_2}\rho^2(z)g_0(y,dz)dy - \theta_1\int_0^{\theta_2}\int_{-\infty}^{t_1\wedge t_2}\rho^2(z)g_0(y,dz)dy \\
&- \theta_2\int_0^{\theta_1}\int_{-\infty}^{t_1\wedge t_2}\rho^2(z)g_0(y,dz)dy + \theta_1\theta_2\int_0^{1}\int_{-\infty}^{t_1\wedge t_2}\rho^2(z)g_0(y,dz)dy.
\end{aligned} \tag{3.7}$$

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 (3.7) 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 $\hat{\mathbb{T}}^{(n)}_{\rho,\xi^{(b)}}$ and $\hat{T}^{(n)}_{\rho,\xi^{(b)}}$ 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

$$T_\rho^{(n)} \geq \hat{q}_{1-\alpha}^{(B)}\Big(T_\rho^{(n)}\Big), \tag{3.8}$$

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

$$W_\rho^{(n,t_0)} := \sup_{\theta\in[0,1]} |\mathbb{T}_\rho^{(n)}(\theta,t_0)| \geq \hat{q}_{1-\alpha}^{(B)}\Big(W_\rho^{(n,t_0)}\Big), \tag{3.9}$$

where $\hat{q}_{1-\alpha}^{(B)}(W_\rho^{(n,t_0)})$ denotes the $(1-\alpha)$-sample quantile of $\hat{W}^{(n,t_0)}_{\rho,\xi^{(1)}}, \dots, \hat{W}^{(n,t_0)}_{\rho,\xi^{(B)}}$, and where $\hat{W}^{(n,t_0)}_{\rho,\xi^{(b)}} := \sup_{\theta\in[0,1]} |\hat{\mathbb{T}}^{(n)}_{\rho,\xi^{(b)}}(\theta,t_0)|$ 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

$$V_{\rho,t_0}^{(n)} \geq q_{1-\alpha}^K, \tag{3.10}$$

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 $\nu_0 \neq 0$*

$$\begin{aligned}\mathbb{P}\big(L_\rho(T_{\rho,g_1}) > 1-\alpha\big) &\leq \liminf_{B\to\infty} \lim_{n\to\infty} \mathbb{P}\big(T_\rho^{(n)} \geq \hat{q}_{1-\alpha}^{(B)}(T_\rho^{(n)})\big) \\ &\leq \limsup_{B\to\infty} \lim_{n\to\infty} \mathbb{P}\big(T_\rho^{(n)} \geq \hat{q}_{1-\alpha}^{(B)}(T_\rho^{(n)})\big) \leq \mathbb{P}\big(L_\rho(T_{\rho,g_1}) \geq 1-\alpha\big) \end{aligned} \tag{3.11}$$

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

$$\mathbb{P}\big(\bar{V}_{\rho,t_0}^{(g_1)} > q_{1-\alpha}^K\big) \leq \liminf_{n\to\infty} \mathbb{P}\big(V_{\rho,t_0}^{(n)} \geq q_{1-\alpha}^K\big) \leq \limsup_{n\to\infty} \mathbb{P}\big(V_{\rho,t_0}^{(n)} \geq q_{1-\alpha}^K\big) \leq \mathbb{P}\big(\bar{V}_{\rho,t_0}^{(g_1)} \geq q_{1-\alpha}^K\big), \tag{3.12}$$

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

$$\begin{aligned}\mathbb{P}\big(L_\rho^{(t_0)}(T_{\rho,g_1}^{(t_0)}) > 1-\alpha\big) &\leq \liminf_{B\to\infty} \lim_{n\to\infty} \mathbb{P}\big(W_\rho^{(n,t_0)} \geq \hat{q}_{1-\alpha}^{(B)}(W_\rho^{(n,t_0)})\big) \\ &\leq \limsup_{B\to\infty} \lim_{n\to\infty} \mathbb{P}\big(W_\rho^{(n,t_0)} \geq \hat{q}_{1-\alpha}^{(B)}(W_\rho^{(n,t_0)})\big) \leq \mathbb{P}\big(L_\rho^{(t_0)}(T_{\rho,g_1}^{(t_0)}) \geq 1-\alpha\big). \end{aligned} \tag{3.13}$$

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, (3.11) basically states that under the local alternative for large $B, n \in \mathbb{N}$ the probability that the test (3.8) 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 (3.12) and (3.13) can be interpreted in the same way.

Corollary 3.12. *Under $\mathbf{H}_0$ the tests (3.8), (3.9) and (3.10) have asymptotic level α, that is if $\nu_0 \neq 0$ we have for each $\alpha \in (0,1)$*

$$\lim_{B\to\infty} \lim_{n\to\infty} \mathbb{P}\big(T_\rho^{(n)} \geq \hat{q}_{1-\alpha}^{(B)}(T_\rho^{(n)})\big) = \alpha \tag{3.14}$$

and furthermore

$$\lim_{n\to\infty} \mathbb{P}\big(V_{\rho,t_0}^{(n)} \geq q_{1-\alpha}^K\big) = \alpha, \quad \lim_{B\to\infty} \lim_{n\to\infty} \mathbb{P}\big(W_\rho^{(n,t_0)} \geq \hat{q}_{1-\alpha}^{(B)}(W_\rho^{(n,t_0)})\big) = \alpha, \tag{3.15}$$

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

Proposition 3.13. *The tests (3.8), (3.9) and (3.10) 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 $\mathbf{H}_1^{(\rho,t_0)}$, 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 The 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

$$\tilde{\theta}_\rho^{(n)} := \arg\max_{\theta\in[0,1]} \sup_{t\in\mathbb{R}} \big|\mathbb{T}_\rho^{(n)}(\theta,t)\big| \tag{3.16}$$

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).

4 Statistical inference for gradual changes

As a generalization of Proposition 3.14 one can show that $k_n^{-1/2}\mathbb{T}_\rho^{(n)}(\theta,t)$ 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 minor regularity conditions, $\text{argmax}_{\theta\in[0,1]}|\mathbb{T}_\rho^{(n)}(\theta,t)|$ is a consistent estimator of $\text{argmax}_{\theta\in[0,1]}|\mathbb{T}_{\rho,g_0}(\theta,t)|$. 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

$$\tilde{\mathcal{D}}_\rho^{(g_0)}(\theta):=\sup_{t\in\mathbb{R}}\sup_{0\le\zeta\le\theta}\left|D_\rho^{(g_0)}(\zeta,\theta,t)\right|.$$

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

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

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 $\nu_1 \neq \nu_2$ 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

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

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

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

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*

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

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 $\mathbb{H}_\rho^{(n)}$

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

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 $|\aleph(\theta)| \leq K\big(\theta - \theta_0\big)^{\varpi+\eta}$ 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

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

should satisfy

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

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

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

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

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

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

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

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 the constants* K_1 *and* K_2 *can be chosen as*

$$K_1 = \left(\frac{1-\varphi}{c}\right)^{2/\varpi} \quad \text{and} \quad K_2 = \left(\frac{1+\varphi}{c}\right)^{2/\varpi}$$

for arbitrary $\varphi \in (0,1)$.

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 $0 < r \le 1$ we define the following empirical distribution function

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

and we denote by

$$F_{n,B}^{(\rho,r)-}(y) := \inf\Big\{x \in \mathbb{R} \;\Big|\; F_{n,B}^{(\rho,r)}(x) \ge y\Big\}$$

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,*

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

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*

1. $\alpha_n^2 B_n \to \infty$,
2. $(n\Delta_n)^{\frac{1-r}{2r}} \alpha_n \to \infty$,
3. $\alpha_n^{-1} n \Delta_n^{1+\tau} \to 0$*, 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 6.7.

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 $\theta_0 < \theta \leq 1$. 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

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

for $y \in U$ and $t \in \mathbb{R}$. Then we have

$$k_0 := \min\Big\{k \in \mathbb{N} \;\Big|\; \exists t \in \mathbb{R}\colon N_k(t) \neq 0\Big\} < \infty,$$

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

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

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 $D^{(g_0)}_\rho$ in (4.1) and define for $t_0 \in \mathbb{R}$ and $\theta \in [0,1]$ the quantities

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

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 $\mathcal{D}^{(g_0)}_\rho(1) > 0$.

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 $\mathcal{D}^{(g_0)}_{\rho,t_0}(1) > 0$.

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

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

holds for each $t \in \mathbb{R}$ and all pairs $(\zeta_1,\theta_1),(\zeta_2,\theta_2) \in C = \{(\zeta,\theta) \in [0,1]^2 \mid \zeta \le \theta\}$ 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 $(\xi^{(b)})_{b=1,\dots,B}$ 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 $\{X_{i\Delta_n}^{(n)} \mid i = 0,\dots,n\}$ and the b-th bootstrap multipliers $\xi_1^{(b)},\dots,\xi_n^{(b)}$. For a given level $\alpha \in (0,1)$, we propose to reject $\mathbf{H}_0$ in favor of $\mathbf{H}_1^*$, if

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

where $\hat{q}_{1-\alpha}^{(B)}\big(\mathbb{H}_{\rho,*}^{(n)}(1)\big)$ denotes the $(1-\alpha)$-quantile of the sample $\hat{\mathbb{H}}_{\rho,*}^{(n,1)}(1),\dots,\hat{\mathbb{H}}_{\rho,*}^{(n,B)}(1)$ 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

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

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* $\alpha \in (0,1)$

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

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*

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

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* $B \in \mathbb{N}$

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

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

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

5 Finite-sample properties

In this section we present the results of an extensive simulation study assessing the finite-sample properties of the new statistical procedures. The design of this study is as follows:

- We apply our estimators and test statistics to n data points $\{X_{\Delta_n}, \dots, X_{n\Delta_n}\}$ 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 $\Delta_n^{-1} = 450, 300, 225, 150, 90$.

- 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

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

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

- In order to simulate data points $\{X_{\Delta_n}, \dots, X_{n\Delta_n}\}$ 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.

- 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

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

k_n	Test (3.8)	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	(3.9)	0.062	0.036	0.024	0.036	0.026	0.036
		(3.10)	0.060	0.042	0.030	0.030	0.016	0.020
75	0.052	(3.9)	0.058	0.048	0.046	0.040	0.046	0.050
		(3.10)	0.040	0.046	0.032	0.036	0.028	0.030
100	0.050	(3.9)	0.046	0.054	0.042	0.046	0.038	0.042
		(3.10)	0.038	0.038	0.036	0.040	0.028	0.032
150	0.068	(3.9)	0.038	0.054	0.054	0.054	0.058	0.066
		(3.10)	0.036	0.036	0.050	0.042	0.052	0.044
250	0.060	(3.9)	0.068	0.056	0.056	0.058	0.064	0.060
		(3.10)	0.046	0.034	0.034	0.032	0.044	0.052

Table 1: **Test procedures (3.8), (3.9) and (3.10) under $\mathbf{H}_0$.** *Simulated rejection probabilities in the application of the test (3.8), the test (3.9) and the test (3.10) using 500 pure jump subordinator data vectors under the null hypothesis.*

- 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.
- 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.
- 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 $T_2 := \{0.1 + j \cdot 0.3 \mid j = 0, 1, \dots, 9\}$.
- For the function ρ we use $\rho_{L,p}$ from (2.8) in Example 2.5 with parameters $L = 1$ and $p = 2$.
- 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 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 $\Theta = \{\hat\theta_1, \dots, \hat\theta_{500}\}$ 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.

k_n	Test (3.8)	Pointwise Tests	$t_0 = 0.5$	$t_0 = 1$	$t_0 = 1.5$	$t_0 = 2$	$t_0 = 2.5$	$t_0 = 3$
50	0.040	(3.9)	0.038	0.042	0.036	0.054	0.034	0.036
		(3.10)	0.036	0.030	0.028	0.042	0.026	0.028
75	0.058	(3.9)	0.024	0.050	0.030	0.048	0.058	0.050
		(3.10)	0.030	0.032	0.020	0.042	0.046	0.036
100	0.050	(3.9)	0.044	0.050	0.040	0.046	0.048	0.052
		(3.10)	0.034	0.040	0.026	0.046	0.040	0.048
150	0.054	(3.9)	0.040	0.050	0.048	0.056	0.048	0.060
		(3.10)	0.040	0.032	0.038	0.038	0.030	0.038
250	0.060	(3.9)	0.046	0.058	0.036	0.056	0.062	0.058
		(3.10)	0.036	0.050	0.030	0.044	0.054	0.046

Table 2: **Test procedures (3.8), (3.9) and (3.10) under $\mathbf{H}_0$.** *Simulated rejection probabilities in the application of the test (3.8), the test (3.9) and the test (3.10) using 500 pure jump subordinator data vectors plus a drift and plus a Brownian motion under the null hypothesis.*

5.1 Finite-sample performance of the procedures in Section 3

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

Simulated rejection probabilities for the tests (3.8), (3.9) and (3.10)

Table 1 and Table 2 show a reasonable approximation of the nominal level $\alpha = 0.05$ of the tests (3.8), (3.9) and (3.10) under $\mathbf{H}_0$. Test (3.10) appears to be slightly more conservative than the test (3.9). Moreover, in Table 2 the underlying process includes a continuous component with $b = \sigma = 1$ and the results are similar to Table 1.

In Figure 1 we depict rejection probabilities for the test (3.8) 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 $\theta_0 = 0.5$. The results are as expected: Large differences of the factor of jump size before and after the change yield higher rejection probabilities. Moreover, due to better approximations the relative frequencies of rejections increase with the effective sample size $k_n = n\Delta_n$. Notice also that the results for pure jump Itō semimartingales and for data including a continuous component are almost the same. This fact indicates a reasonable performance of the proposed truncation technique for an ordinary sample size $n = 22500$.

Figure 2 shows 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 the middle of the data set, i.e. for $\theta_0 \approx 0.5$. Furthermore, in this case the power of the test is increasing in the effective sample size $k_n = n\Delta_n$ as well and the performance for data including a continuous component is nearly the same.

In Figure 3 we display relative frequencies of rejection for different values of the parameter $p \in [2, 20]$ of the function $\rho_{1,p}$ defined in (2.8). This function is used to calculate the process

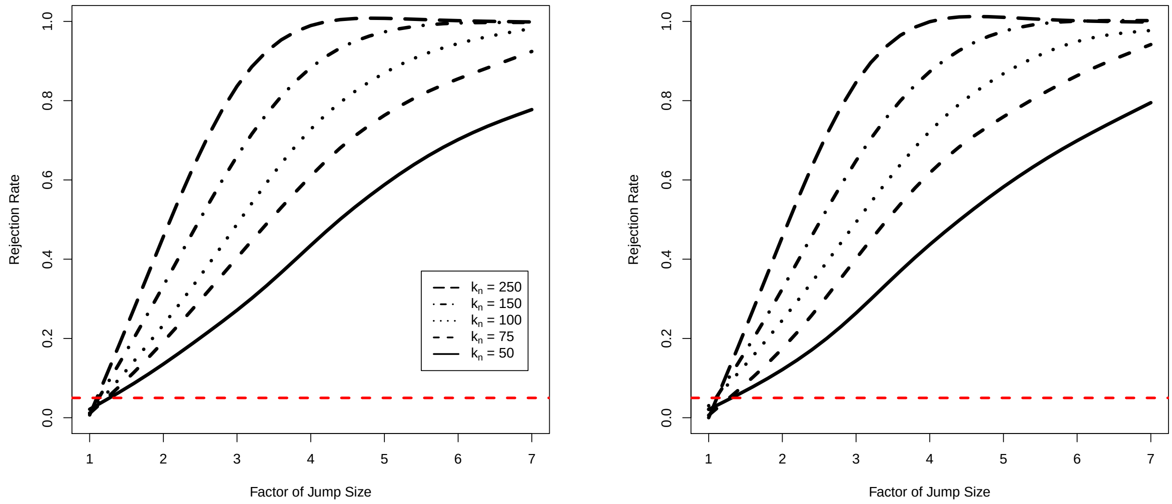

Figure 1: *Simulated rejection probabilities of the test* (3.8) *for different factors of jump size* ψ *in* (5.2)*. Pure jump data on the left-hand side and data including a Brownian motion with drift on the right-hand side. The dashed red line indicates* $\alpha = 5\%$*.*

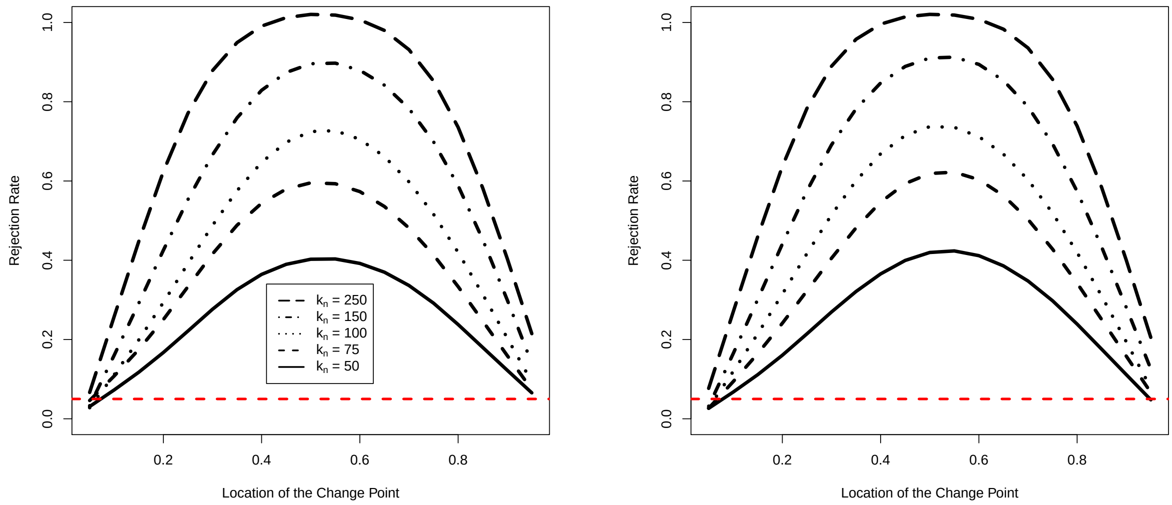

Figure 2: *Simulated relative frequencies of rejection of the test* (3.8) *for different locations of the change point* θ_0 *in* (5.2) *for pure jump data (left-hand side) and with an additional Brownian motion with drift (right-hand side). The dashed red line indicates the nominal level* $\alpha = 5\%$*.*

$\mathbb{T}^{(n)}_{\rho_{1,p}}(\theta, t)$ for certain values of $\theta \in [0,1]$ and $t \in T_1$. In each run the change point is located at $\theta_0 = 0.5$ and we have $\psi = 3$ 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 4 we depict rejection probabilities of the tests (3.9) and (3.10) 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

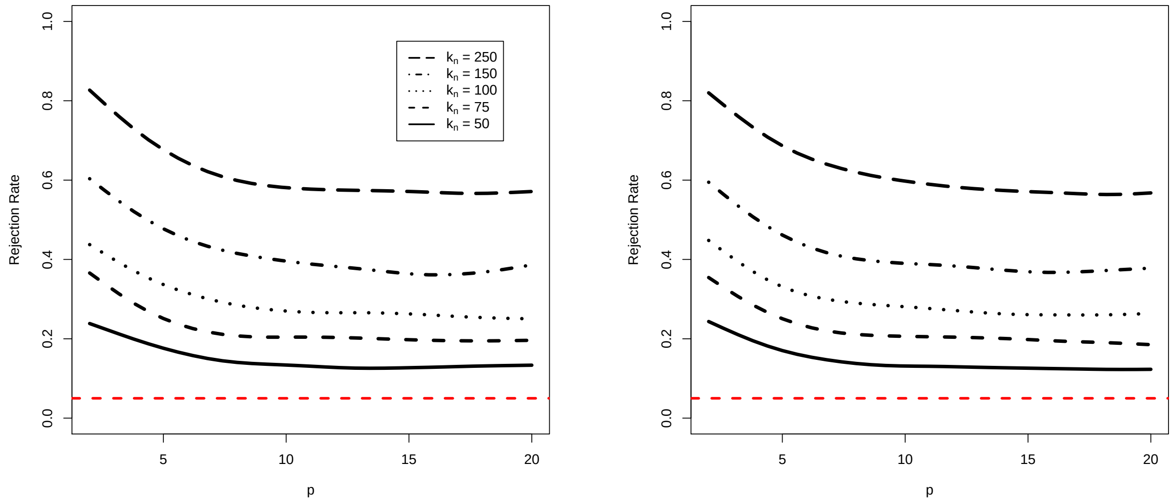

Figure 3: *Simulated rejection probabilities of the test* (3.8) *for different values of the parameter* $p \geq 2$ *in the function* $\rho_{1,p}$ *for pure jump data (left panel) and plus an additional Brownian motion with drift (right panel). The dashed red line indicates the confidence level* $\alpha = 5\%$.

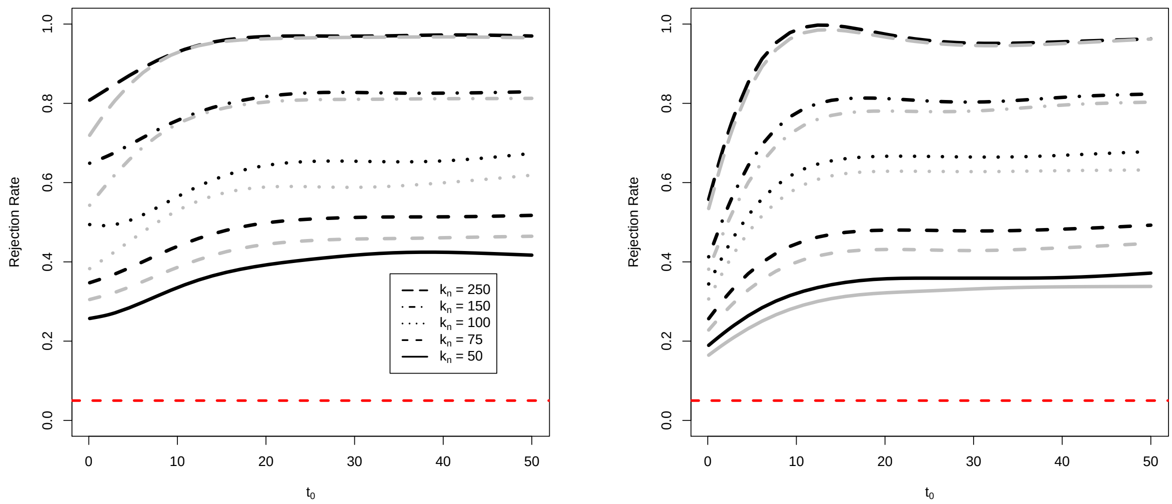

Figure 4: *Simulated rejection probabilities of the test* (3.9) *(black lines) and the test* (3.10) *(grey lines) for different values* t_0 *for pure jump data (left-hand side) and with an additional Brownian motion with drift (right-hand side). The dashed red line indicates the nominal level* $\alpha = 5\%$.

$\psi = 3$. We observe that test (3.9) has slightly more power than test (3.10) 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

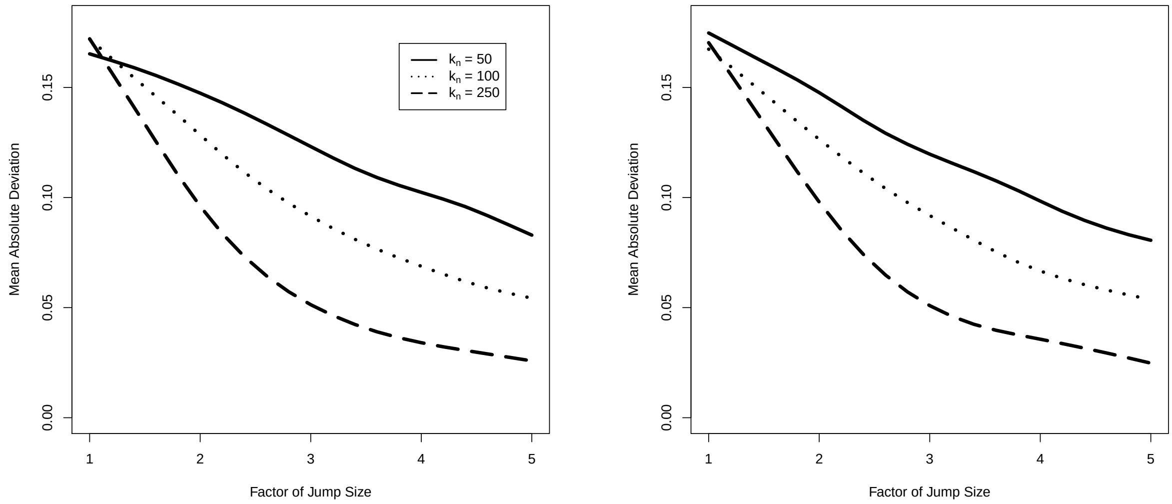

Figure 5: *Mean absolute deviation of estimator* (3.16) *for different values* $\psi \geq 1$ *in* (5.2) *for pure jump data (left panel) and plus an additional Brownian motion with drift (right panel).*

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 semimartingales 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.

Finite-sample performance of the argmax-estimator (3.16)

In Figure 5 we display mean absolute deviations of estimator (3.16) for different values $\psi \in [1, 5]$ in (5.2), where the true change point is located at $\theta_0 = 0.5$. The results correspond to Figure 1 in the sense that large values of ψ yield a better performance of the statistical procedure. Because of better approximations the mean absolute deviation is also decreasing in the effective sample size $k_n = n\Delta_n$. Additionally, we also observe in the estimation results that due to the truncation approach the mean absolute deviation is nearly unaffected by the presence of a Brownian component.

Figure 6 shows mean absolute deviations of estimator (3.16) for different locations of the change point $\theta_0 \in (0, 1)$, where we choose $\psi = 3$ in (5.2). Similar to Figure 2 the results suggest that a change point can be detected best if it is located around the middle of the data set, i.e. if $\theta_0 \approx 0.5$. Furthermore, as in Figure 5 the estimation error is decreasing in the effective sample size k_n.

5.2 Finite-sample performance of the procedures in Section 4

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

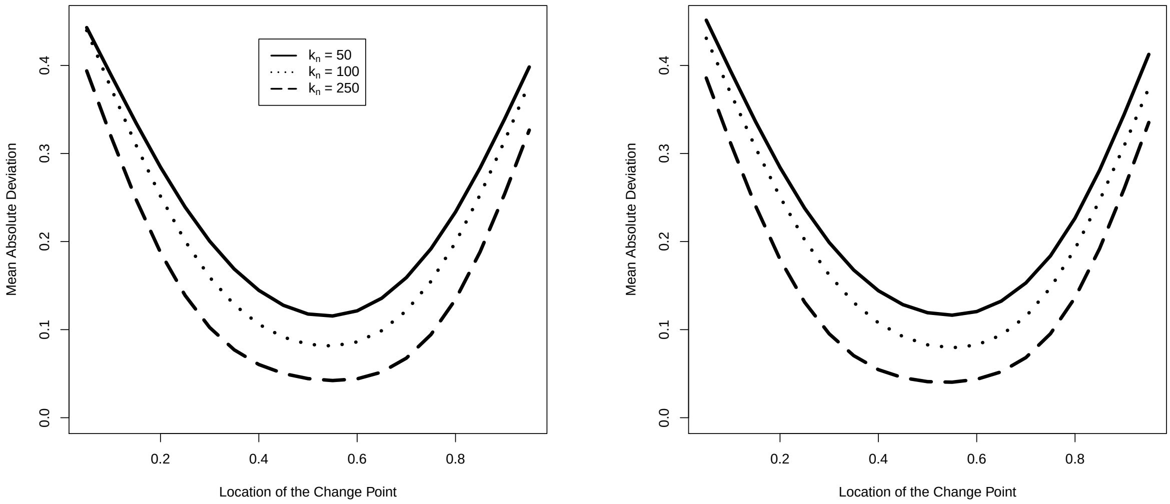

Figure 6: *Mean absolute deviation of estimator* (3.16) *for different values* $\theta_0 \in (0,1)$ *in* (5.2) *for pure jump Itō semimartingales (left-hand side) and plus an additional Brownian component (right-hand side).*

	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

Table 3: **Test procedures** (4.19) **and** (4.20) under $\mathbf{H}_0$. *Simulated rejection probabilities of the test* (4.19) *and the test* (4.20) *using* 500 *pure jump Itō semimartingale data vectors under the null hypothesis.*

Finite-sample performance of the tests (4.19) **and** (4.20)

Table 3 and Table 4 show simulated rejection probabilities of test (4.19) and test (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 $T_1 = \{0.1 \cdot j \mid j = 1, \dots, 30\}$. In both cases for pure jump Itō semimartingales ($b = \sigma = 0$; Table 3) and for Itō semimartingales including a Brownian component ($b = \sigma = 1$; Table 4) we observe a reasonable approximation of the nominal level $\alpha = 5\%$.

For computational reasons the results on the tests (4.19) and (4.20) depicted below are obtained for a sample size $n = 10000$ and effective sample size $k_n \in \{50, 100, 200\}$. In each run we choose the confidence level $\alpha = 5\%$ and the supremum in $t \in \mathbb{R}$ is approximated by the maximum over the finite grid $T_1 = \{0.1 \cdot j \mid j = 1, \dots, 30\}$.

	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	0.044	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

Table 4: **Test procedures** (4.19) **and** (4.20) under $\mathbf{H}_0$. *Simulated rejection probabilities of the test* (4.19) *and the test* (4.20) *using* 500 *pure jump Itō semimartingale data vectors plus a drift and plus a Brownian motion under the null hypothesis.*

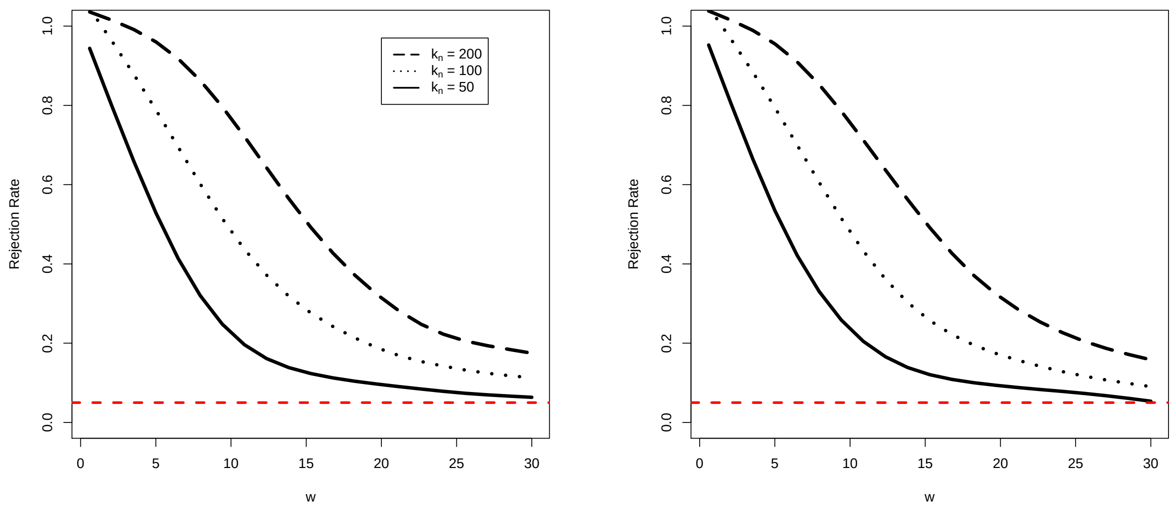

Figure 7: *Simulated rejection probabilities of the test* (4.19) *for different values* $w \in [0.6, 30]$ *in* (5.3) *for pure jump Itō semimartingales (left panel) and plus a Brownian motion and a drift (right panel). The dashed red line indicates the nominal level* $\alpha = 5\%$.

Figure 7 shows 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 the power of the test is increasing in $k_n = n\Delta_n$ as well. Furthermore, all simulation results on the tests (4.19) and (4.20) are similar for pure jump processes and processes including a Brownian component. This indicates that our truncation approach also works in this setup.

In Figure 8 we depict rejection rates of the test (4.19) for different locations of the change 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 effec-

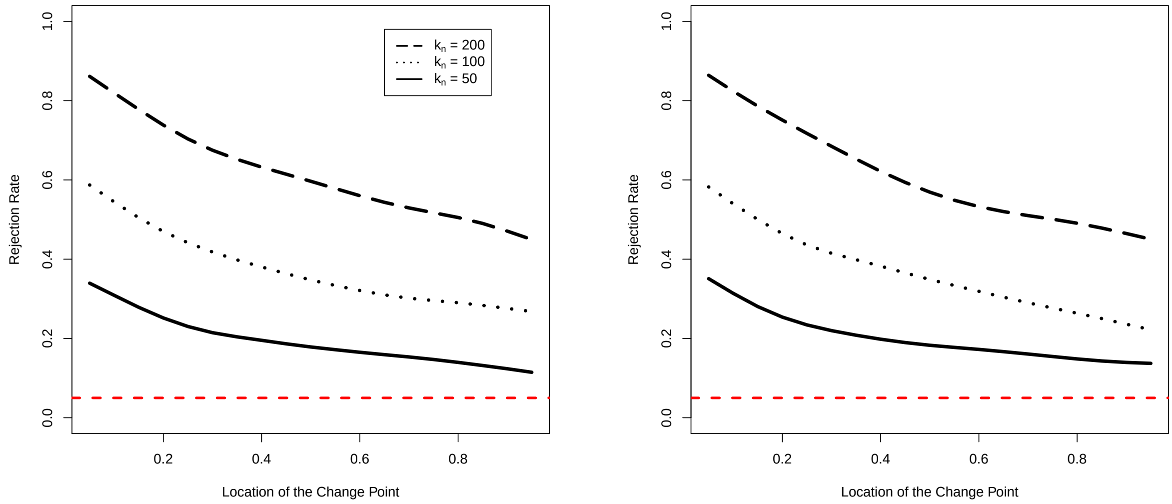

Figure 8: *Simulated relative frequencies of rejections of the test (4.19) for different locations of the change point $\theta_0 \in (0,1)$ in (5.3) for pure jump processes (left-hand side) and plus an additional Brownian motion with drift (right-hand side). The dashed red line indicates the confidence level $\alpha = 5\%$.*

tive sample size $k_n = n\Delta_n$ and moreover it is decreasing in θ_0. The latter can be explained by the shape of the model (5.3): For large θ_0 the jump characteristic is "close" to the null hypothesis.

Finite-sample performance of the estimator $\hat{\theta}_\rho^{(n)}$ defined 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.

From the theoretical standpoint of 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

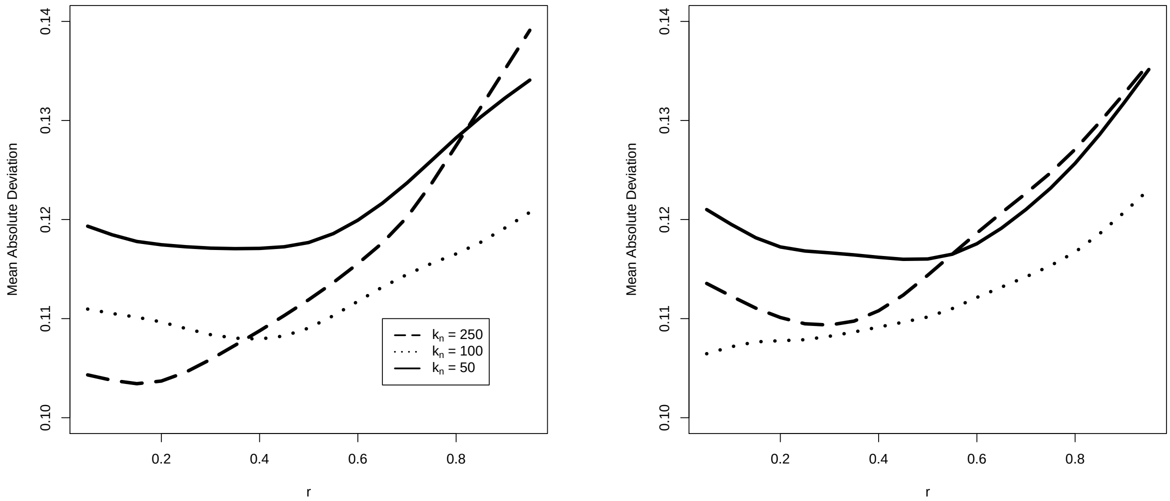

Figure 9: *Mean absolute deviations of estimator* (4.14) *for different choices of* $r \in (0, 1]$ *in Step 1 for pure jump Itō semimartingales (left panel) and with an additional Brownian component (right panel).*

declared 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 the sample size $n = 22500$ and we vary the effective sample size in $k_n = n\Delta_n \in \{50, 100, 250\}$. 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 declared 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.

Figure 9 shows mean absolute deviations for different choices of $r \in (0, 1]$ in Step 1. The graphics suggest 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 Figure 10 we depict mean absolute deviations of the estimator (4.14) for different choices of the preliminary estimate $\hat{\theta}^{(pr)} \in (0, 1)$ in Step 1. The estimation error is smallest, if the preliminary estimate is chosen close to 1. This finding corresponds to another simulation study which is not included below and which has shown that the estimation 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.

Figure 11 shows simulated mean absolute deviations of estimator (4.14) for different degrees of smoothness of the change w in (5.3). The results correspond to Figure 7 and confirm the intuitive idea that a smooth change is more difficult to detect. Moreover, better approximations

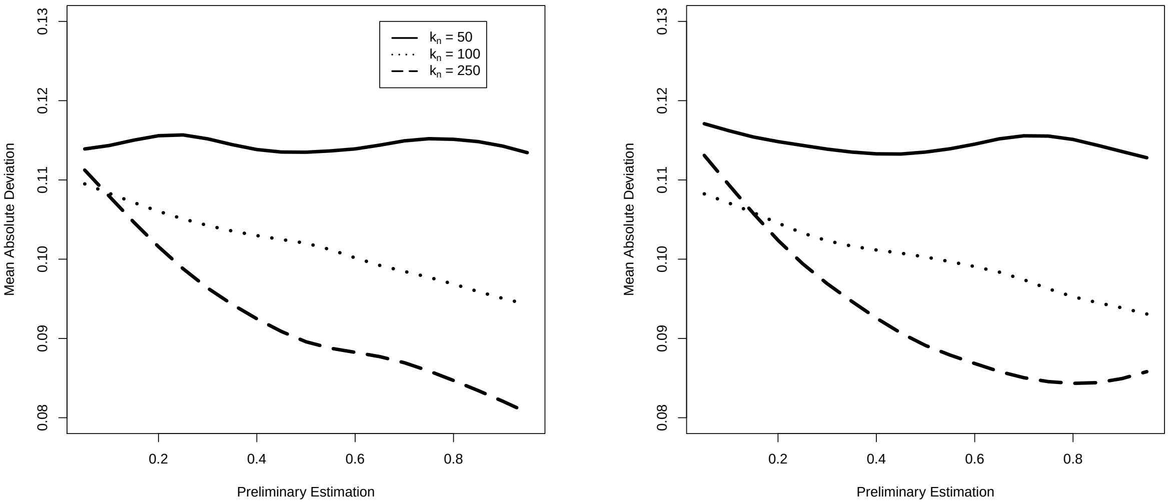

Figure 10: *Mean absolute deviations of estimator* (4.14) *for different choices of the preliminary estimate* $\hat{\theta}^{(pr)} \in (0,1)$ *in Step 1 for pure jump Itō semimartingales (left-hand side) and plus an additional Brownian motion with drift (right-hand side).*

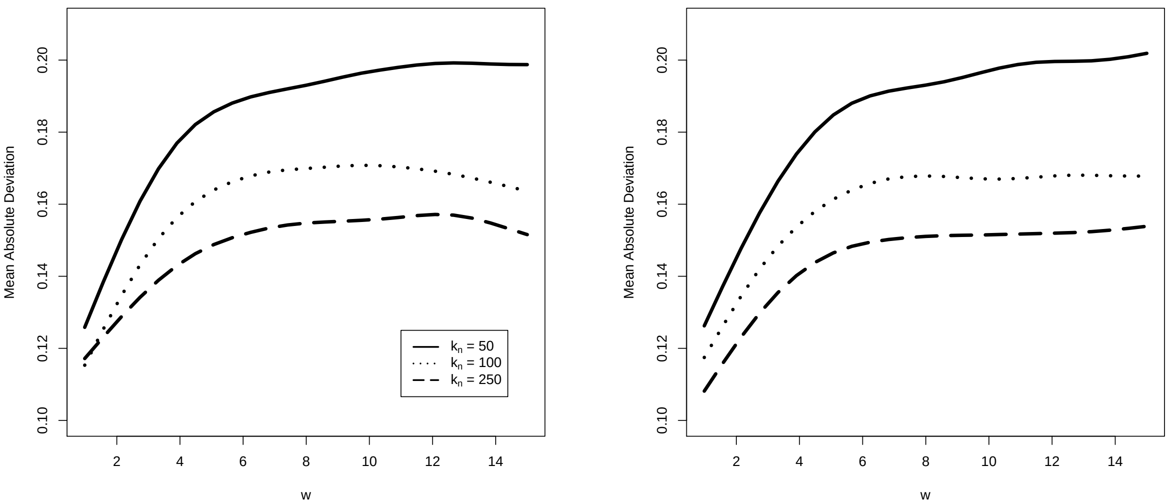

Figure 11: *Mean absolute deviations of estimator* (4.14) *for different degrees of smoothness of the change* w *in* (5.3) *for pure jump processes (left panel) and with an additional continuous component (right panel).*

for large effective sample sizes $k_n = n\Delta_n$ reduce the estimation error.

In Figure 12 we display simulated mean absolute deviations of estimator $\hat{\theta}_\rho^{(n)}$ for different locations of the change point $\theta_0 \in (0,1)$ in (5.3). The results correspond to Figure 8 and show that for small values of θ_0 the change point can be detected best. This is a consequence of model (5.3): For large $\theta_0 \in (0,1)$ the jump behaviour is nearly constant.

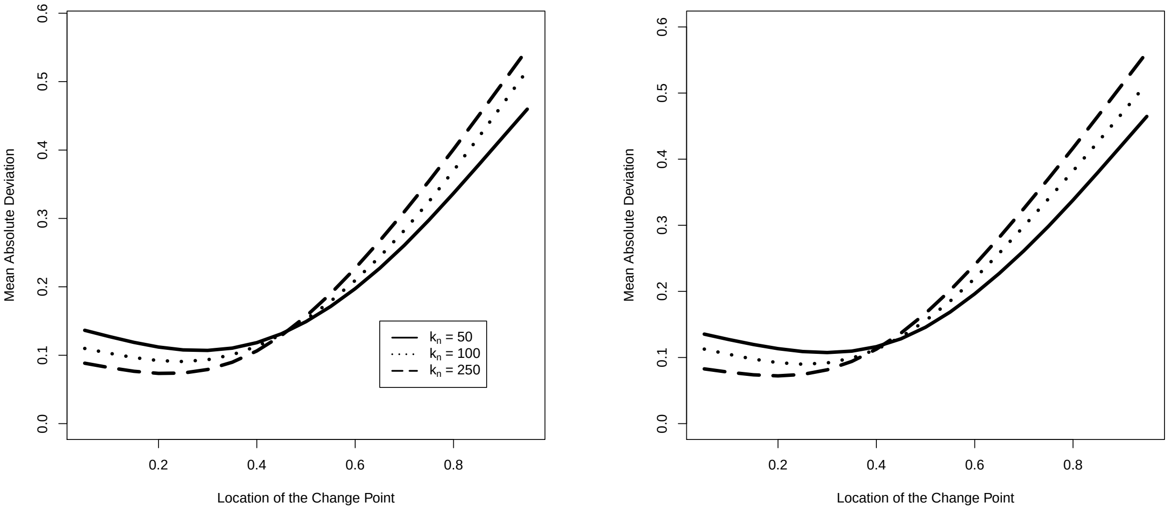

Figure 12: *Simulated mean absolute deviation of estimator* (4.14) *for different locations of the change point for pure jump processes (left-hand side) and plus an additional Brownian motion with drift (right-hand side).*

Acknowledgements 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).

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.

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

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

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.

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.

Billingsley, P. (1999). *Convergence of Probability Measures.* J. Wiley and Sons. ISBN: 0-471-19745-9.

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

Bücher, A. (2011). *Statistical Inference for Copulas and Extremes.* PhD thesis, Ruhr-Universität Bochum.

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.

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

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.

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

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

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.

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.

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.

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.

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

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

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

Jacod, J. (1979). *Calcul Stochastique et Problèmes de Martingales*. Springer. ISBN: 3-540-09253-6.

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

Jacod, J. and Shiryaev, A. (2002). *Limit Theorems for Stochastic Processes*. Springer. ISBN: 3-540-17882-1.

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.

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

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

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.

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

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

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.

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

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

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.

Pollard, D. (1990). *Empirical Processes: Theory and Applications*, volume 2. Institute of Mathematical Statistics and the American Statistical Association.

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.

Sato, K.-I. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press.

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

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.

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.

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

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

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

6 Proofs and technical details

Before we prove the results of Section 3 and Section 4 we begin with a generalization of the results in Hoffmann and Vetter (2017).

6.1 Weak convergence of the empirical truncated Lévy distribution function

Theorem 6.1 below is a central limit theorem for $N_\rho^{(n)}$ centered at $N_\rho(g^{(n)};\cdot,\cdot)$. Precisely, we consider 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.

Theorem 6.1. *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

$$H_\rho((\theta_1,t_1);(\theta_2,t_2)) := \int_0^{\theta_1\wedge\theta_2}\int_{-\infty}^{t_1\wedge t_2} \rho^2(z)g_0(y,dz)dy. \tag{6.1}$$

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

$$d_\rho((\theta_1,t_1);(\theta_2,t_2)) = \Big\{\int_0^{\theta_1}\int_{t_1\wedge t_2}^{t_1\vee t_2} \rho^2(z)g_0(y,dz)dy + \int_{\theta_1}^{\theta_2}\int_{-\infty}^{t_2} \rho^2(z)g_0(y,dz)dy\Big\}^{1/2} \tag{6.2}$$

for $\theta_1 \le \theta_2$.

Theorem 6.1 is a generalization of Hoffmann and Vetter (2017) in the sense that their Theorem 3.1 can be obtained if we set $g_0(y,dz) = \nu(dz)$ for a Lévy measure ν and $g_1 = g_2 = 0$. In order to prove Theorem 6.1 we divide the process $G_\rho^{(n)}$ into two parts which correspond to large and small jumps of the underlying process $X^{(n)}$, respectively. To this end we choose an auxiliary function $\Psi\colon \mathbb{R}_+ \to \mathbb{R}$ which is $\mathcal{C}^\infty$ and satisfies $\mathbb{1}_{[1,\infty)}(z) \le \Psi(z) \le \mathbb{1}_{[1/2,\infty)}(z)$ for all $z \in \mathbb{R}_+$. For $\alpha > 0$ define $\Psi_\alpha\colon \mathbb{R}\to\mathbb{R}$ via $\Psi_\alpha(z) = \Psi(|z|/\alpha)$ and let $\Psi_\alpha^\circ\colon \mathbb{R}\to\mathbb{R}$ be the function $\Psi_\alpha^\circ(z) = 1-\Psi_\alpha(z)$.

For the function ρ we define $\rho_\alpha(z) = \rho(z)\Psi_\alpha(z)$ and $\rho_\alpha^\circ(z) = \rho(z)\Psi_\alpha^\circ(z)$. Furthermore, let

$$\chi_t^{(\alpha)}(z) = \rho(z)\Psi_\alpha(z)\mathbb{1}_{(-\infty,t]}(z) \quad \text{and} \quad \chi_t^{\circ(\alpha)}(z) = \rho(z)\Psi_\alpha^\circ(z)\mathbb{1}_{(-\infty,t]}(z), \tag{6.3}$$

for $t, z \in \mathbb{R}$ and define the following empirical processes:

$$G_{\rho,n}^{(\alpha)}(\theta,t) = \sqrt{n\Delta_n}\Big\{\frac{1}{n\Delta_n}\sum_{i=1}^{\lfloor n\theta\rfloor} \chi_t^{(\alpha)}(\Delta_i^n X^{(n)})\mathbb{1}_{\{|\Delta_i^n X^{(n)}|>v_n\}} - N_{\rho_\alpha}(g^{(n)};\theta,t)\Big\},$$

$$G_{\rho,n}^{\circ(\alpha)}(\theta,t) = \sqrt{n\Delta_n}\Big\{\frac{1}{n\Delta_n}\sum_{i=1}^{\lfloor n\theta\rfloor} \chi_t^{\circ(\alpha)}(\Delta_i^n X^{(n)})\mathbb{1}_{\{|\Delta_i^n X^{(n)}|>v_n\}} - N_{\rho_\alpha^\circ}(g^{(n)};\theta,t)\Big\}.$$

Then, of course, we have $G_\rho^{(n)}(\theta,t) = G_{\rho,n}^{(\alpha)}(\theta,t) + G_{\rho,n}^{\circ(\alpha)}(\theta,t)$. In Section 6.4 we show that it suffices to prove three auxiliary lemmas in order to establish Theorem 6.1. The first one is concerned with the behaviour of the large jumps, i.e. it holds for $G_{\rho,n}^{(\alpha)}$ and a fixed $\alpha > 0$.

Lemma 6.2. *If Assumption 2.3 is satisfied, we have weak convergence*

$$G_{\rho,n}^{(\alpha)} \rightsquigarrow \mathbb{G}_{\rho_\alpha}$$

in $\ell^\infty([0,1]\times\mathbb{R})$ for each fixed $\alpha > 0$, where $\mathbb{G}_{\rho_\alpha}$ denotes a tight centered Gaussian process with covariance function

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

The sample paths of $\mathbb{G}_{\rho_\alpha}$ are almost surely uniformly continuous with respect to the semimetric

$$d_{\rho_\alpha}((\theta_1,t_1);(\theta_2,t_2)) = \Big\{ \int_0^{\theta_1} \int_{t_1\wedge t_2}^{t_1\vee t_2} \rho_\alpha^2(z) g_0(y,dz)dy + \int_{\theta_1}^{\theta_2} \int_{-\infty}^{t_2} \rho_\alpha^2(z) g_0(y,dz)dy \Big\}^{1/2}$$

for $\theta_1 \leq \theta_2$.

The general idea behind the proof of Lemma 6.2 is to replace the increments of the underlying process $X^{(n)}$ by increments of pure jump Itō semimartingales. Precisely, let $\mu^{(n)}$ be the Poisson random measure associated with the jumps of $X^{(n)}$. Then we consider

$$L^{(n)} = \big(z\mathbb{1}_{\{|z|>v_n\}}\big) \star \mu^{(n)} \tag{6.4}$$

with the truncation $v_n = \gamma\Delta_n^{\overline{w}}$ as above. The main advantage of the processes $L^{(n)}$ is that they have deterministic characteristics. Therefore, their increments are independent (see Theorem II.4.15 in Jacod and Shiryaev (2002)) and we can use a central limit theorem tailored for triangular arrays of independent stochastic processes from Kosorok (2008) to prove weak convergence of

$$Y_f^{(n)}(\theta,t) = \sqrt{n\Delta_n}\Big\{\frac{1}{n\Delta_n}\sum_{i=1}^{\lfloor n\theta\rfloor}[f(\Delta_i^n L^{(n)})\mathbb{1}_{(-\infty,t]}(\Delta_i^n L^{(n)}) - \mathbb{E}(f(\Delta_i^n L^{(n)})\mathbb{1}_{(-\infty,t]}(\Delta_i^n L^{(n)}))]\Big\} \tag{6.5}$$

to $\mathbb{G}_{\rho_\alpha}$, where $(\theta,t)\in[0,1]\times\mathbb{R}$ and where $f\colon\mathbb{R}\to\mathbb{R}$ is a bounded continuous function, for which we plug in ρ_α and ρ_α° later. In order to prove Lemma 6.2 we need to ensure that the distance between $Y_{\rho_\alpha}^{(n)}$ and $G_{\rho,n}^{(\alpha)}$ is small. To this end, our next claim shows that the bias due to estimating $(n\Delta_n)^{-1}\sum_{i=1}^{\lfloor n\theta\rfloor}\mathbb{E}(\rho_\alpha(\Delta_i^n L^{(n)})\mathbb{1}_{(-\infty,t]}(\ \Delta_i^n L^{(n)}))$ instead of $N_{\rho_\alpha}(g^{(n)};\theta,t)$ is small compared to the rate of convergence. In order to state the result recall that for a real-valued non-negative function $f:\Xi\to\mathbb{R}_+$ on a measure space $(\Xi,\mathcal{B},\vartheta)$ the essential supremum with respect to ϑ is given by

$$\vartheta - \text{ess sup}_{x\in\Xi}(f) = \inf_{B\in\mathcal{B},\vartheta(B)=0}\ \sup_{x\in\Xi\setminus B} f(x).$$

Moreover, recall that λ_1 denotes the restriction of the one-dimensional Lebesgue measure to $[0,1]$.

Proposition 6.3. *Let $(\mu^{(n)})_{n\in\mathbb{N}}$ be a sequence of Poisson random measures with predictable compensators $\bar{\mu}^{(n)}(ds,dz) = \nu_s^{(n)}(dz)ds$ such that (2.6) is satisfied for each $n\in\mathbb{N}$ with a null sequence $\Delta_n>0$ and a sequence of transition kernels $g^{(n)}$ from $([0,1],\mathbb{B}([0,1]))$ into $(\mathbb{R},\mathbb{B})$ with*

$$\lambda_1 - \text{ess sup}_{y\in[0,1]}\Big(\int(1\wedge|z|^\beta)g^{(n)}(y,dz)\Big) \leq K$$

for each $n\in\mathbb{N}$ and some $\beta\in[0,2]$, $K>0$. Furthermore, let $f\colon\mathbb{R}\to\mathbb{R}$ be a bounded Borel measurable function satisfying $|f(z)|\leq K|z|^p$ on a neighbourhood of 0 for some $K>0$, $p\geq\beta$.

Then if $v_n > 0$ is a null sequence and $L^{(n)}$ is defined as in (6.4) we have

$$\sup_{i=1,\dots,n} \sup_{t\in\overline{\mathbb{R}}} \Big| \mathbb{E}\left\{ f\big(\Delta_i^n L^{(n)}\big) \mathbb{1}_{(-\infty,t]}\big(\Delta_i^n L^{(n)}\big) \right\} - n\Delta_n \big(N_f(g^{(n)}; i/n, t) - N_f(g^{(n)}; (i-1)/n, t)\big) \Big| \\ = O(\Delta_n^2 v_n^{-2\beta} + \Delta_n v_n^{p-\beta}), \quad (6.6)$$

with $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$.

The following proposition establishes the desired weak convergence of $Y_f^{(n)}$.

Proposition 6.4. *Suppose Assumption 2.3 is satisfied and let $f\colon \mathbb{R} \to \mathbb{R}$ be a continuous function with $|f(z)| \leq K(1 \wedge |z|^p)$ for all $z \in \mathbb{R}$ and some $K > 0$. Then the processes $Y_f^{(n)}$ from (6.5) converge weakly in $\ell^\infty([0,1] \times \mathbb{R})$ to the tight mean zero Gaussian process $\mathbb{G}_f$ from Lemma 6.2, that is*

$$Y_f^{(n)} \rightsquigarrow \mathbb{G}_f.$$

In order to obtain the result from Theorem 6.1 the following lemma ensures that the limiting process $\mathbb{G}_{\rho_\alpha}$ converges in a suitable sense as $\alpha \to 0$.

Lemma 6.5. *Under Assumption 2.3 the weak convergence*

$$\mathbb{G}_{\rho_\alpha} \rightsquigarrow \mathbb{G}_\rho$$

holds in $\ell^\infty([0,1] \times \mathbb{R})$ as $\alpha \to 0$.

Its proof is a direct consequence of the following result.

Proposition 6.6. *Suppose Assumption 2.3 is satisfied and let $f_n\colon \mathbb{R} \to \mathbb{R}$ ($n \in \mathbb{N}_0$) be Borel measurable functions with $|f_n(z)| \leq K(1 \wedge |z|^p)$ for a constant $K > 0$ and all $n \in \mathbb{N}_0$, $z \in \mathbb{R}$. Assume further that $f_n(z) \to f_0(z)$ converges for all z outside a set $B \in \mathbb{B}$ such that $[0,1] \times B$ is a $g_0(y, dz)dy$-null set. Then we have weak convergence*

$$\mathbb{G}_{f_n} \rightsquigarrow \mathbb{G}_{f_0}$$

in $\ell^\infty([0,1] \times \mathbb{R})$ for $n \to \infty$.

Our final lemma shows that the contribution due to small jumps are uniformly small as α tends to zero.

Lemma 6.7. *Suppose Assumption 2.3 is satisfied. Then for each $\eta > 0$ we have:*

$$\lim_{\alpha\to 0} \limsup_{n\to\infty} \mathbb{P}\Big(\sup_{(\theta,t)\in[0,1]\times\mathbb{R}} \big| G_{\rho,n}^{\circ(\alpha)}(\theta, t) \big| > \eta \Big) = 0.$$

6.2 The multiplier bootstrap approach

In this section we investigate a bootstrapped version of $G_\rho^{(n)}$ given by

$$\hat{G}_\rho^{(n)}(\theta, t) = \frac{1}{\sqrt{n\Delta_n}} \sum_{i=1}^{\lfloor n\theta \rfloor} \xi_i \rho\big(\Delta_i^n X^{(n)}\big) \mathbb{1}_{(-\infty,t]}\big(\Delta_i^n X^{(n)}\big) \mathbb{1}_{\{|\Delta_i^n X^{(n)}| > v_n\}}, \tag{6.7}$$

for $(\theta, t) \in [0,1] \times \mathbb{R}$ and where the sequence of multipliers $(\xi_i)_{i\in\mathbb{N}}$ satisfies Assumption 3.6. The bootstrap results in Section 3 and Section 4 are a simple consequence of Proposition 10.7 in Kosorok (2008) and the following theorem, which establishes weak convergence conditional on the data in probability of $\hat{G}_\rho^{(n)}$.

Theorem 6.8. *Let Assumption 2.3 be valid and let the multipliers $(\xi_i)_{i\in\mathbb{N}}$ satisfy Assumption 3.6. Then we have*

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

in $\ell^\infty([0,1] \times \mathbb{R})$, where $\mathbb{G}_\rho$ is the tight mean zero Gaussian process of Theorem 6.1.

Similar to the proof of Theorem 6.1 we show Theorem 6.8 by treating small and large increments of $X^{(n)}$ separately. Therefore, with the quantities defined prior to (6.3) we consider the processes

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

Section 6.5 reveals that the claim of Theorem 6.8 can be obtained using the following two lemmas which state weak convergence conditional on the data in probability of the previously established processes for fixed $\alpha > 0$.

Lemma 6.9. *If Assumption 2.3 and Assumption 3.6 are satisfied, we have*

$$\hat{G}_{\rho,n}^{(\alpha)} \rightsquigarrow_\xi \mathbb{G}_{\rho_\alpha}$$

in $\ell^\infty([0,1] \times \mathbb{R})$ for each fixed $\alpha > 0$.

Lemma 6.10. *Suppose Assumption 2.3 and Assumption 3.6 are valid. Then for each $\alpha > 0$ in a neighbourhood of 0*

$$\hat{G}_{\rho,n}^{\circ(\alpha)} \rightsquigarrow_\xi \mathbb{G}_{\rho_\alpha^\circ}$$

holds in $\ell^\infty([0,1] \times \mathbb{R})$.

The lemmas above will be verified by approximating the truncated increments of the underlying processes by the increments of the pure jump Itō semimartingales from (6.4)

$$L^{(n)} = (z\mathbb{1}_{\{|z|>v_n\}}) \star \mu^{(n)},$$

with the usual truncation $v_n = \gamma\Delta_n^{\overline{w}}$. The main advantage of the processes $L^{(n)}$ is the fact, that they have deterministic characteristics and therefore independent increments. As a consequence, we can use a result from Kosorok (2008) for triangular arrays of processes which are independent within rows to prove weak convergence conditional on the data in probability of the bootstrapped analogs of $Y_f^{(n)}$ from (6.5) which are given by

$$\hat{Y}_f^{(n)}(\theta, t) = \frac{1}{\sqrt{n\Delta_n}} \sum_{i=1}^{\lfloor n\theta \rfloor} \xi_i f(\Delta_i^n L^{(n)}) \mathbb{1}_{(-\infty,t]}(\Delta_i^n L^{(n)}),$$

for $(\theta, t) \in [0,1] \times \mathbb{R}$ and where $f\colon \mathbb{R} \to \mathbb{R}$ is a bounded continuous function. Precisely, the following proposition is the main tool in order to obtain Lemma 6.9 and Lemma 6.10.

Proposition 6.11. *Suppose Assumption 2.3 and Assumption 3.6 are satisfied. Then for a continuous function $f\colon \mathbb{R} \to \mathbb{R}$ satisfying $|f(z)| \leq K(1 \wedge |z|^p)$ for all $z \in \mathbb{R}$ and some $K > 0$ we have*

$$\hat{Y}_f^{(n)} \rightsquigarrow_\xi \mathbb{G}_f$$

in $\ell^\infty([0,1] \times \mathbb{R})$, where $\mathbb{G}_f$ is the tight mean zero Gaussian process defined in Theorem 6.1.

6.3 Alternative Assumptions

All results in this paper also hold under the weaker assumptions given below. Here and throughout the following proofs, K or $K(\delta)$ denote generic constants which sometimes depend on an auxiliary quantity δ and may change from place to place.

Assumption 6.12. 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 $\{i\Delta_n \mid i = 0, 1, \ldots, n\}$. 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

$$\nu_s^{(n)}(dz) = g^{(n)}\Big(\frac{s}{n\Delta_n}, dz\Big), \tag{6.8}$$

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 $y \in [0,1]$

$$g^{(n)}(y, dz) = g_0(y, dz) + \frac{1}{\sqrt{n\Delta_n}} g_1(y, dz) + \mathcal{R}_n(y, dz) \tag{6.9}$$

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 $m_1, m_2 \in \{g_0, g_1, g_2\}$

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

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

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

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
$$m_b = \frac{1+2\overline{w}}{1-\overline{w}} \leq 4 \quad \text{and} \quad m_\sigma = \frac{1+2\overline{w}}{1/2-\overline{w}},$$
we have
$$\sup_{n\in\mathbb{N}} \sup_{s\in\mathbb{R}_+} \left\{\mathbb{E}\big|b_s^{(n)}\big|^{m_b} \vee \mathbb{E}\big|\sigma_s^{(n)}\big|^{m_\sigma}\right\} < \infty.$$

In the sequel, we will work with Assumption 6.12 without further mention. This is due to the following result which proves that Assumption 2.3 implies the set of conditions above.

Proposition 6.13. *Assumption 2.3 is sufficient for Assumption 6.12.*

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 6.12 define the following quantities:
$$\overline{r} := 3\tau, \quad \overline{v} := \frac{\tau}{1+3\beta}, \quad q := \overline{r} - (1+3\beta)\overline{v} = 2\tau, \tag{6.10}$$
and recall that $\overline{w} = (1+5\tau)/4$.

ρ is suitable for Assumption 6.12(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 6.12(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
$$\begin{aligned}(1+2\overline{r}-2\overline{w}) \vee 1 = t_2^{-1} < 1+\tau = t_1^{-1} = (1+\frac{q}{2})\\ = 2(1-\beta\overline{w}(1+\epsilon)) < (1+2\overline{v}(p-\beta)) \wedge (1+2\overline{w})\end{aligned} \tag{6.11}$$
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 $\lambda_1 - \operatorname{ess\,sup}\big(\int (1 \wedge |z|^{(\beta+\delta)\wedge 2}) g_i(y,dz)\big) \leq K(\delta) < \infty$ for every $\delta > 0$ and all $i \in \{0,1,2\}$. Moreover, due to Definition 2.1(3) we have
$$h_y^{(i)}(z)|z|^{\overline{p}} \leq K|z|^{-1-\epsilon}, \tag{6.12}$$

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 6.12(a3).

Furthermore we have

$$\frac{1+2\overline{w}}{1-\overline{w}} = \frac{6+10\tau}{3-5\tau} \quad \text{and} \quad \frac{1+2\overline{w}}{1/2-\overline{w}} = \frac{6+10\tau}{1-5\tau},$$

so Assumption 2.3(c) yields Assumption 6.12(c).

We are thus left with proving Assumption 6.12(a(4)I) and (a(4)II). Obviously, $0 < \overline{v} < \overline{r}$ holds with the choice in (6.10). First, we verify Assumption 6.12(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

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

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

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

for some $\xi(z) \in [z - \Delta_n^{\overline{r}}, z + \Delta_n^{\overline{r}}]$. Finally, we distinguish the cases $\beta < 1$ and $\beta \geq 1$ 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,

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

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 6.12(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:

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

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,

$$\begin{aligned} &\int\int \mathbb{1}_{\{|x-z|\le\Delta_n^{\overline{r}}\}} \mathbb{1}_{\{|x|>\alpha\}} \mathbb{1}_{\{|z|>\alpha\}} m_1(y_1, dx) m_2(y_2, dz) \\ &\quad \le O(\Delta_n^{\overline{r}}) + K\Delta_n^{\overline{r}} \int_{M'}^{\infty} \frac{1}{|z^2 - \Delta_n^{2\overline{r}}|} z^{-2} \mathbb{1}_{\{z>\alpha\}} dz = o(\Delta_n^q) \end{aligned} \tag{6.13}$$

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 (6.13) holds since the last integral is finite. □

6.4 Proof of Theorem 6.1

Proof of Proposition 6.6. In order to show weak convergence we want to use Theorem 1.5.4 and Theorem 1.5.7 in Van der Vaart and Wellner (1996). To this end, we prove asymptotical uniform d-equicontinuity in probability of $\mathbb{G}_{f_n}$ for some suitable semimetric d on $[0,1] \times \mathbb{R}$ with Theorem 2.2.4 in this reference.

First, recall that $\mathbb{G}_{f_n}$ are tight centered Gaussian processes in $\ell^\infty([0,1] \times \mathbb{R})$ with covariance function

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

and their sample paths are almost surely uniformly continuous with respect to the semimetric

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

for $\theta_1 \le \theta_2$. Due to Lemma B.1 in Appendix B we obtain for the L^8-norm

$$\|\mathbb{G}_{f_n}(\theta_1, t_1) - \mathbb{G}_{f_n}(\theta_2, t_2)\|_8 = 105^{\frac{1}{8}} d_{f_n}((\theta_1, t_1); (\theta_2, t_2)), \tag{6.14}$$

for each $n \in \mathbb{N}$ and $(\theta_1, t_1), (\theta_2, t_2) \in [0,1] \times \mathbb{R}$. Additionally, the convex, non-decreasing, non-zero function $\varphi(x) = x^8$ clearly satisfies $\varphi(0) = 0$ and $\limsup_{x,y\to\infty} \varphi(x)\varphi(y)/\varphi(cxy) < \infty$ for some constant $c > 0$. Furthermore, by Lemma B.2 in Appendix B the process $\mathbb{G}_{f_n}$ is separable for each $n \in \mathbb{N}_0$ in the sense of Theorem 2.2.4 in Van der Vaart and Wellner (1996). Thus, this theorem can be applied and due to (6.14) it yields a constant $K > 0$, which does not depend on $n \in \mathbb{N}_0$, such that for all $\zeta, \delta > 0$

$$\Big\| \sup_{d_{f_n}((\theta_1,t_1);(\theta_2,t_2))\le\delta} \big|\mathbb{G}_{f_n}(\theta_1,t_1) - \mathbb{G}_{f_n}(\theta_2,t_2)\big|\Big\|_8 \le K\Big\{ \int_0^\zeta \big(D(\varepsilon, d_{f_n})\big)^{\frac{1}{8}} d\varepsilon + \delta\big(D(\zeta, d_{f_n})\big)^{\frac{1}{4}}\Big\}, \tag{6.15}$$

where $D(\varepsilon, d_{f_n})$ denotes the packing number of $[0,1]\times\mathbb{R}$ with respect to the semimetric d_{f_n} at distance ε. According to Lemma B.3 we have $D(\varepsilon, d_{f_n}) \le K/\varepsilon^4$ for every $n \in \mathbb{N}_0$, where $K > 0$ does not depend on $n \in \mathbb{N}_0$. Therefore, with (6.15) we conclude that there exists a $K > 0$ which is independent of n such that

$$\Big\| \sup_{d_{f_n}((\theta_1,t_1);(\theta_2,t_2))\le\delta} \big|\mathbb{G}_{f_n}(\theta_1,t_1) - \mathbb{G}_{f_n}(\theta_2,t_2)\big|\Big\|_8 \le K\Big\{ \int_0^\zeta \varepsilon^{-1/2} d\varepsilon + \delta/\zeta\Big\} \le K\big(\zeta^{1/2} + \delta/\zeta\big), \tag{6.16}$$

for each $\zeta, \delta > 0$ and $n \in \mathbb{N}_0$. Now, for arbitrary $\varepsilon, \eta > 0$ and $K > 0$ from (6.16) choose a $\zeta > 0$ with $2^8 K^8 \zeta^4/\varepsilon^8 < \eta/2$ and for this ζ choose a $\delta > 0$ with $(2^8K^8\delta^8)/(\zeta^8\varepsilon^8) < \eta/2$. Then due to (6.16) we obtain for each $n \in \mathbb{N}$ with the Markov inequality

$$\begin{aligned} \mathbb{P}\Big(\sup_{d_{f_n}((\theta_1,t_1);(\theta_2,t_2))<\delta} \big|\mathbb{G}_{f_n}(\theta_1,t_1) - \mathbb{G}_{f_n}(\theta_2,t_2)\big| > \varepsilon\Big) &\le \frac{K^8(\zeta^{1/2}+\delta/\zeta)^8}{\varepsilon^8} \\ &\le \frac{K^8(2\max\{\zeta^{1/2},\delta/\zeta\})^8}{\varepsilon^8} = \frac{2^8K^8\max\{\zeta^4,\delta^8/\zeta^8\}}{\varepsilon^8} \\ &\le \frac{2^8K^8\zeta^4}{\varepsilon^8} + \frac{2^8K^8\delta^8}{\zeta^8\varepsilon^8} < \eta. \end{aligned} \tag{6.17}$$

Furthermore, d_{f_n} converges uniformly to d_{f_0} by Lebesgue's dominated convergence theorem. Thus, $\mathbb{G}_{f_n}$ is asymptotically uniformly d_{f_0}-equicontinuous in probability, because for each $\varepsilon, \eta > 0$ we have

$$\limsup_{n\to\infty} \mathbb{P}\Big(\sup_{d_{f_0}((\theta_1,t_1);(\theta_2,t_2))<\delta/2} \big|\mathbb{G}_{f_n}(\theta_1,t_1) - \mathbb{G}_{f_n}(\theta_2,t_2)\big| > \varepsilon\Big) < \eta,$$

with $\delta > 0$ from (6.17). Moreover, Lemma B.3 also shows that $([0,1]\times\mathbb{R}, d_{f_0})$ is totally bounded. Trivially, the marginals of $\mathbb{G}_{f_n}$ converge to the corresponding marginals of $\mathbb{G}_{f_0}$, because these are centered multivariate normal distributions and their covariance functions converge again by Lebesgue's dominated convergence theorem. Therefore, the desired result holds due to Theorem 1.5.4 and Theorem 1.5.7 in Van der Vaart and Wellner (1996). □

Proof of Proposition 6.3. Let $\overline{F}_n = \{z\colon |z| > v_n\}$, $\tilde{N}^{(n)} = \mathbb{1}_{\overline{F}_n}(z) \star \mu^{(n)}$ and let $i \in \{1,\dots,n\}$ be fixed in the entire proof. According to Proposition II.1.14 in Jacod and Shiryaev (2002) for each $n \in \mathbb{N}$ there exist a thin random set D_n with an exhausting sequence of stopping times $(T_m^{(n)})_{m\in\mathbb{N}}$ and an $\mathbb{R}$-valued optional process $\xi^{(n)}$ such that

$$\mu^{(n)}(\omega; ds, dz) = \sum_{m\in\mathbb{N}} \epsilon_{\big(T_m^{(n)}(\omega), \xi^{(n)}_{T_m^{(n)}(\omega)}(\omega)\big)}(ds, dz), \tag{6.18}$$

where $\epsilon_{(s,x)}$ denotes the Dirac measure with mass in $(s,x) \in \mathbb{R}_+ \times \mathbb{R}$. Furthermore, due to Lemma A.13 $\tilde{N}^{(n)}_{t_2} - \tilde{N}^{(n)}_{t_1}$ follows a Poisson distribution for $0 \leq t_1 \leq t_2$ and the sets $\tilde{A}^{(i)}_n := \big\{\tilde{N}^{(n)}_{i\Delta_n} - \tilde{N}^{(n)}_{(i-1)\Delta_n} \leq 1\big\}$ satisfy

$$\mathbb{P}\big((\tilde{A}^{(i)}_n)^C\big) = O(\Delta_n^2 v_n^{-2\beta}), \tag{6.19}$$

where M^C denotes the complement of a set M. Thus, we calculate for $n \in \mathbb{N}$ large enough

$$\begin{aligned}
\gamma_n^{(i,t)} :=& \Big|\mathbb{E}\left\{f(\Delta_i^n L^{(n)})\mathbb{1}_{(-\infty,t]}(\Delta_i^n L^{(n)})\right\} - n\Delta_n\big(N_f(g^{(n)}; i/n, t) - N_f(g^{(n)}; (i-1)/n, t)\big)\Big| \\
=& \Big| \int\limits_{\tilde{A}^{(i)}_n} f(\Delta_i^n L^{(n)})\mathbb{1}_{(-\infty,t]}(\Delta_i^n L^{(n)})d\mathbb{P} + \int\limits_{(\tilde{A}^{(i)}_n)^C} f(\Delta_i^n L^{(n)})\mathbb{1}_{(-\infty,t]}(\Delta_i^n L^{(n)})d\mathbb{P} - \\
&- n\Delta_n \int\limits_{(i-1)/n}^{i/n} \int\limits_{-\infty}^{t} f(z)g^{(n)}(y,dz)dy\Big| \\
\leq& K\Delta_n^2 v_n^{-2\beta} + \Big| \int\limits_{\{\tilde{N}^{(n)}_{i\Delta_n} - \tilde{N}^{(n)}_{(i-1)\Delta_n} = 1\}} f(\Delta_i^n L^{(n)})\mathbb{1}_{(-\infty,t]}(\Delta_i^n L^{(n)})d\mathbb{P} - \\
&- n\Delta_n \int\limits_{(i-1)/n}^{i/n} \int\limits_{-\infty}^{t} f(z)g^{(n)}(y,dz)dy\Big|,
\end{aligned}$$

where the inequality above follows because of two reasons, first (6.19) as well as the fact that f is bounded lead to the term $K\Delta_n^2 v_n^{-2\beta}$ and secondly for each $\omega \in \big\{\tilde{N}^{(n)}_{i\Delta_n} - \tilde{N}^{(n)}_{(i-1)\Delta_n} = 0\big\}$ and $m \in \mathbb{N}$ we have $\big(T^{(n)}_m(\omega), \xi^{(n)}_{T^{(n)}_m(\omega)}(\omega)\big) \notin ((i-1)\Delta_n, i\Delta_n] \times \overline{F}_n$ such that $\Delta_i^n L^{(n)}(\omega) = 0$ and thus $f(\Delta_i^n L^{(n)}(\omega)) = 0$ by the assumptions on f. However, $\big(T^{(n)}_m(\omega), \xi^{(n)}_{T^{(n)}_m(\omega)}(\omega)\big) \in ((i-1)\Delta_n, i\Delta_n] \times \overline{F}_n$ holds for exactly one $m \in \mathbb{N}$ if $\omega \in \big\{\tilde{N}^{(n)}_{i\Delta_n} - \tilde{N}^{(n)}_{(i-1)\Delta_n} = 1\big\}$. This observation yields the following bound

$$\begin{aligned}
\gamma_n^{(i,t)} \leq& \Big| \int\limits_{\{\tilde{N}^{(n)}_{i\Delta_n} - \tilde{N}^{(n)}_{(i-1)\Delta_n} = 1\}} \int\limits_{(i-1)\Delta_n}^{i\Delta_n} \int\limits_{-\infty}^{t} f(z)\mathbb{1}_{\{|z|>v_n\}}\mu^{(n)}(\omega; ds, dz)\mathbb{P}(d\omega) - \\
&- n\Delta_n \int\limits_{(i-1)/n}^{i/n} \int\limits_{-\infty}^{t} f(z)g^{(n)}(y,dz)dy\Big| + K\Delta_n^2 v_n^{-2\beta} \\
\leq& \Big| \int\limits_{\Omega} \int\limits_{(i-1)\Delta_n}^{i\Delta_n} \int\limits_{-\infty}^{t} f(z)\mathbb{1}_{\{|z|>v_n\}}\mu^{(n)}(\omega; ds, dz)\mathbb{P}(d\omega) - \\
&- n\Delta_n \int\limits_{(i-1)/n}^{i/n} \int\limits_{-\infty}^{t} f(z)g^{(n)}(y,dz)dy\Big| + K\Delta_n^2 v_n^{-2\beta} + \delta_n^{(i,t)},
\end{aligned} \tag{6.20}$$

with

$$\delta_n^{(i,t)} = \Big| \int\limits_{\{\tilde N^{(n)}_{i\Delta_n} - \tilde N^{(n)}_{(i-1)\Delta_n} \geq 2\}} \int\limits_{(i-1)\Delta_n}^{i\Delta_n} \int\limits_{-\infty}^{t} f(z) \mathbb{1}_{\{|z|>v_n\}} \mu^{(n)}(\omega; ds, dz) \mathbb{P}(d\omega)\Big|. \tag{6.21}$$

We apply the defining relation of the predictable compensator of an optional $\mathcal{P}'$-σ-finite random measure. But notice that it cannot be guaranteed that the integrand in the stochastic integral with respect to $\mu^{(n)}$ in the first line of (6.20) is $\mathcal{P}'$-measurable. Therefore, we treat the leading term after the last inequality sign in (6.20) and $\delta_n^{(i,t)}$ separately. However, the integrand $f(z)\mathbb{1}_{(-\infty,t]}(z)\mathbb{1}_{\{|z|>v_n\}}\mathbb{1}_{((i-1)\Delta_n, i\Delta_n]}(s)$ on the right-hand side of (6.20) is $\mathcal{P}'$-measurable. Thus, Theorem II.1.8 in Jacod and Shiryaev (2002) yields

$$\begin{aligned}
\gamma_n^{(i,t)} \leq K\Delta_n^2 v_n^{-2\beta} + \delta_n^{(i,t)} + \Big| \int\limits_{(i-1)\Delta_n}^{i\Delta_n} \int\limits_{-\infty}^{t} f(z) \mathbb{1}_{\{|z|>v_n\}} \nu_s^{(n)}(dz) ds - \\
- n\Delta_n \int\limits_{(i-1)/n}^{i/n} \int\limits_{-\infty}^{t} f(z) g^{(n)}(y, dz) dy\Big| \\
= K\Delta_n^2 v_n^{-2\beta} + \delta_n^{(i,t)} + \Big| n\Delta_n \int\limits_{(i-1)/n}^{i/n} \int\limits_{-\infty}^{t} f(z) \mathbb{1}_{\{|z|\leq v_n\}} g^{(n)}(y, dz) dy\Big|.
\end{aligned}$$

Now, because of $|f(z)| \leq K|z|^p$ on a neighbourhood of 0, the above display yields for $n \in \mathbb{N}$ large enough

$$\begin{aligned}
\gamma_n^{(i,t)} &\leq K\Delta_n^2 v_n^{-2\beta} + \delta_n^{(i,t)} + K\Delta_n v_n^{p-\beta} n \int\limits_{(i-1)/n}^{i/n} \int (1 \wedge |z|^\beta) g^{(n)}(y, dz) dy \\
&\leq K\Delta_n^2 v_n^{-2\beta} + \delta_n^{(i,t)} + K\Delta_n v_n^{p-\beta}. \qquad (6.22)
\end{aligned}$$

Finally, (6.18) and the assumption that f is bounded by some constant $K > 0$ gives an estimate for $\delta_n^{(i,t)}$ from (6.21)

$$\begin{aligned}
\delta_n^{(i,t)} &\leq \sum_{\ell=2}^{\infty} \int\limits_{\{\tilde N^{(n)}_{i\Delta_n} - \tilde N^{(n)}_{(i-1)\Delta_n} = \ell\}} \sum_{m\in\mathbb{N}} \big|f\big(\xi^{(n)}_{T_m^{(n)}}\big)\big| \mathbb{1}_{(-\infty,t]}\big(\xi^{(n)}_{T_m^{(n)}}\big) \mathbb{1}_{((i-1)\Delta_n, i\Delta_n]}\big(T_m^{(n)}\big) \mathbb{1}_{\overline{F}_n}\big(\xi^{(n)}_{T_m^{(n)}}\big) d\mathbb{P} \\
&\leq \sum_{\ell=2}^{\infty} K\ell \times \mathbb{P}\Big(\tilde N^{(n)}_{i\Delta_n} - \tilde N^{(n)}_{(i-1)\Delta_n} = \ell\Big),
\end{aligned}$$

because for $\omega \in \big\{\tilde N^{(n)}_{i\Delta_n} - \tilde N^{(n)}_{(i-1)\Delta_n} = \ell\big\}$ we have $\#\big\{m \in \mathbb{N} \mid \big(T_m^{(n)}(\omega), \xi^{(n)}_{T_m^{(n)}(\omega)}(\omega)\big) \in ((i-1)\Delta_n, i\Delta_n] \times \overline{F}_n\big\} = \ell$, where $\#M$ denotes the cardinality of a set M. With Lemma A.13 and the previous inequality we obtain

$$\delta_n^{(i,t)} \leq \exp\big\{-\zeta_i^{(n)}\big\} \sum_{\ell=2}^{\infty} K\ell \times \frac{\big(\zeta_i^{(n)}\big)^\ell}{\ell!} \leq K\big(\zeta_i^{(n)}\big)^2,$$

with

$$\zeta_i^{(n)} = n\Delta_n \int\limits_{(i-1)/n}^{i/n} \int\limits_{\overline{F}_n} g^{(n)}(y,dz)dy \leq n\Delta_n v_n^{-\beta} \int\limits_{(i-1)/n}^{i/n} \int (1 \wedge |z|^\beta) g^{(n)}(y,dz)dy \leq K\Delta_n v_n^{-\beta},$$

for $n \in \mathbb{N}$ large enough. Thus, $\delta_n^{(i,t)} \leq K\Delta_n^2 v_n^{-2\beta}$ holds and (6.22) yields (6.6), because neither of the bounds for $\gamma_n^{(i,t)}$ or $\delta_n^{(i,t)}$ depends on i or t. □

Proof of Proposition 6.4. The processes $Y_f^{(n)}$ have the form

$$Y_f^{(n)}(\omega;(\theta,t)) = \sum_{i=1}^{m_n} \{g_{ni}(\omega;(\theta,t)) - \mathbb{E}(g_{ni}(\cdot;(\theta,t)))\},$$

with $m_n = n$ and the triangular array $\{g_{ni}(\omega;(\theta,t)) \mid n \in \mathbb{N}; i = 1,\dots,n; (\theta,t) \in [0,1] \times \mathbb{R}\}$ of processes

$$g_{ni}(\omega;(\theta,t)) = \frac{1}{\sqrt{n\Delta_n}} f(\Delta_i^n L^{(n)}(\omega)) \mathbb{1}_{(-\infty,t]}(\Delta_i^n L^{(n)}(\omega)) \mathbb{1}_{\{i \leq \lfloor n\theta \rfloor\}},$$

which is independent within rows, because $L^{(n)}$ has independent increments as it has deterministic characteristics (see Theorem II.4.15 in Jacod and Shiryaev (2002)). Thus, by Theorem 11.16 in Kosorok (2008), the proof is complete once we can show the following six conditions of the triangular array $\{g_{ni}\}$ (see for instance Kosorok (2008) for the notions of AMS and manageability):

(A) $\{g_{ni}\}$ is almost measurable Suslin (AMS);

(B) $\{g_{ni}\}$ is manageable with envelopes $\{G_{ni} \mid n \in \mathbb{N}; i = 1,\dots,n\}$ which are also independent within rows. Here, we set $G_{ni} = \frac{K}{\sqrt{n\Delta_n}}(1 \wedge |\Delta_i^n L^{(n)}|^p)$ with $K > 0$ such that $|f(x)| \leq K(1 \wedge |z|^p)$;

(C)
$$H_f((\theta_1,t_1);(\theta_2,t_2)) = \lim_{n\to\infty} \mathbb{E}\left\{Y_f^{(n)}(\theta_1,t_1) Y_f^{(n)}(\theta_2,t_2)\right\}$$
for all $(\theta_1,t_1),(\theta_2,t_2) \in [0,1] \times \mathbb{R}$, with H_f defined in (6.1);

(D) $\limsup\limits_{n\to\infty} \sum\limits_{i=1}^n \mathbb{E}G_{ni}^2 < \infty$;

(E)
$$\lim_{n\to\infty} \sum_{i=1}^n \mathbb{E}G_{ni}^2 \mathbb{1}_{\{G_{ni} > \epsilon\}} = 0$$
for each $\epsilon > 0$;

(F) For $(\theta_1,t_1),(\theta_2,t_2) \in [0,1] \times \mathbb{R}$ and d_f defined in (6.2) the limit

$$d_f((\theta_1,t_1);(\theta_2,t_2)) = \lim_{n\to\infty} d_f^{(n)}((\theta_1,t_1);(\theta_2,t_2))$$

with

$$d_f^{(n)}((\theta_1,t_1);(\theta_2,t_2)) = \left\{ \sum_{i=1}^n \mathbb{E}\left| g_{ni}(\cdot;(\theta_1,t_1)) - g_{ni}(\cdot;(\theta_2,t_2)) \right|^2 \right\}^{1/2}$$

exists, and for all deterministic sequences $((\theta_n^{(1)}, t_n^{(1)}))_{n\in\mathbb{N}}, ((\theta_n^{(2)}, t_n^{(2)}))_{n\in\mathbb{N}} \subset [0,1]\times\mathbb{R}$ with $d_f((\theta_n^{(1)}, t_n^{(1)});(\theta_n^{(2)}, t_n^{(2)})) \to 0$ we also have $d_f^{(n)}((\theta_n^{(1)}, t_n^{(1)});(\theta_n^{(2)}, t_n^{(2)})) \to 0$.

Proof of (A). Using Lemma 11.15 in Kosorok (2008) the triangular array $\{g_{ni}\}$ is AMS if it is separable, that is for each $n\in\mathbb{N}$ there exists a countable subset $S_n \subset [0,1]\times\mathbb{R}$ such that

$$\mathbb{P}^*\Bigg(\sup_{(\theta_1,t_1)\in[0,1]\times\mathbb{R}} \inf_{(\theta_2,t_2)\in S_n} \sum_{i=1}^n (g_{ni}(\omega;(\theta_1,t_1)) - g_{ni}(\omega;(\theta_2,t_2)))^2 > 0 \Bigg) = 0.$$

But if we choose $S_n = ([0,1]\times\mathbb{R})\cap\mathbb{Q}^2$ for all $n\in\mathbb{N}$, we obtain

$$\sup_{(\theta_1,t_1)\in\mathbb{R}} \inf_{(\theta_2,t_2)\in S_n} \sum_{i=1}^n (g_{ni}(\omega;(\theta_1,t_1)) - g_{ni}(\omega;(\theta_2,t_2)))^2 = 0$$

for each $\omega\in\Omega$ and $n\in\mathbb{N}$.

Proof of (B). G_{ni} are independent within rows since the $L^{(n)}$ have deterministic characteristics. Thus, according to Theorem 11.17 in Kosorok (2008), it suffices to show that the triangular arrays

$$\{\tilde{g}_{ni}(\omega;t) := \frac{1}{\sqrt{n\Delta_n}} f(\Delta_i^n L^{(n)}(\omega)) \mathbb{1}_{(-\infty,t]}(\Delta_i^n L^{(n)}(\omega)) \mid n\in\mathbb{N}; i=1,\ldots,n; t\in\mathbb{R}\},$$

and

$$\{\tilde{h}_{ni}(\omega;\theta) := \mathbb{1}_{\{i\le\lfloor n\theta\rfloor\}} \mid n\in\mathbb{N}; i=1,\ldots,n; \theta\in[0,1]\}$$

are manageable with envelopes $\{G_{ni} \mid n\in\mathbb{N}; i=1,\ldots,n\}$ and $\{\tilde{H}_{ni}(\omega) :\equiv 1 \mid n\in\mathbb{N}; i = 1,\ldots,n\}$, respectively. Concerning the first triangular array $\{\tilde{g}_{ni}\}$ define for $n\in\mathbb{N}$ and $\omega\in\Omega$ the set

$$\mathcal{G}_{n\omega} = \Bigg\{ \Bigg(\frac{1}{\sqrt{n\Delta_n}} f(\Delta_1^n L^{(n)}(\omega)) \mathbb{1}_{(-\infty,t]}(\Delta_1^n L^{(n)}(\omega)), \ldots \\ \ldots, \frac{1}{\sqrt{n\Delta_n}} f(\Delta_n^n L^{(n)}(\omega)) \mathbb{1}_{(-\infty,t]}(\Delta_n^n L^{(n)}(\omega)) \Bigg) \Bigg| t\in\mathbb{R} \Bigg\} \subset \mathbb{R}^n.$$

These sets are bounded with envelope vector

$$G_n(\omega) = (G_{n1}(\omega),\ldots,G_{nn}(\omega)) \in \mathbb{R}^n.$$

For $i_1,i_2\in\{1,\ldots,n\}$ the projection

$$p_{i_1,i_2}(\mathcal{G}_{n\omega}) = \Bigg\{ \Bigg(\frac{1}{\sqrt{n\Delta_n}} f(\Delta_{i_1}^n L^{(n)}(\omega)) \mathbb{1}_{(-\infty,t]}(\Delta_{i_1}^n L^{(n)}(\omega)), \\ \frac{1}{\sqrt{n\Delta_n}} f(\Delta_{i_2}^n L^{(n)}(\omega)) \mathbb{1}_{(-\infty,t]}(\Delta_{i_2}^n L^{(n)}(\omega)) \Bigg) \mid t\in\mathbb{R} \Bigg\} \subset \mathbb{R}^2$$

onto the i_1-th and the i_2-th coordinate is an element of the set

$$\Big\{\{(0,0)\},\{(0,0),(s_{i_1,n}(\omega),0)\},\{(0,0),(0,s_{i_2,n}(\omega))\},\{(0,0),(s_{i_1,n}(\omega),s_{i_2,n}(\omega))\},$$
$$\{(0,0),(s_{i_1,n}(\omega),0),(s_{i_1,n}(\omega),s_{i_2,n}(\omega))\},\{(0,0),(0,s_{i_2,n}(\omega)),(s_{i_1,n}(\omega),s_{i_2,n}(\omega))\}\Big\}.$$

with $s_{i_1,n}(\omega) = \frac{1}{\sqrt{n\Delta_n}} f(\Delta_{i_1}^n L^{(n)}(\omega))$ and $s_{i_2,n}(\omega) = \frac{1}{\sqrt{n\Delta_n}} f(\Delta_{i_2}^n L^{(n)}(\omega))$. Consequently, in the sense of Definition 4.2 in Pollard (1990), for every $s \in \mathbb{R}^2$ no proper coordinate projection of $\mathcal{G}_{n\omega}$ can surround s and therefore $\mathcal{G}_{n\omega}$ has a pseudo dimension of at most 1 (Definition 4.3 in Pollard (1990)). Thus, by Corollary 4.10 in the same reference, there exist constants A and W which depend only on the pseudodimension such that

$$D_2\big(x\|\alpha\odot G_n(\omega)\|_2,\alpha\odot\mathcal{G}_{n\omega}\big) \le Ax^{-W} =: \zeta(x),$$

for all $0 < x \le 1$, $n \in \mathbb{N}$, $\omega \in \Omega$ and each rescaling vector $\alpha \in \mathbb{R}^n$ with non-negative entries, where $\|\cdot\|_2$ denotes the Euclidean distance on $\mathbb{R}^n$, D_2 denotes the packing number with respect to the Euclidean distance and $\odot$ denotes coordinate-wise multiplication. Obviously, we have

$$\int_0^1 \sqrt{\log\zeta(x)}dx < \infty,$$

and therefore the triangular array $\{\tilde{g}_{ni}\}$ is indeed manageable with envelopes $\{G_{ni}\}$.

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

$$\begin{aligned}\mathcal{H}_{n\omega} &:= \{(\tilde{h}_{n1}(\omega;\theta),\dots,\tilde{h}_{nn}(\omega;\theta)) \mid \theta\in[0,1]\}\\ &= \{(0,\dots,0),(1,0,\dots,0),(1,1,0,\dots,0),\dots,(1,\dots,1)\}.\end{aligned}$$

Then, for any $i_1,i_2 \in \{1,\dots,n\}$, the projection $p_{i_1,i_2}(\mathcal{H}_{n\omega})$ of $\mathcal{H}_{n\omega}$ onto the i_1-th and the i_2-th coordinate is either $\{(0,0),(1,0),(1,1)\}$ or $\{(0,0),(0,1),(1,1)\}$. Therefore, the same reasoning as above shows that $\mathcal{H}_{n\omega}$ is a set of pseudodimension at most one, whence the triangular array $\{\tilde{h}_{ni}\}$ is manageable with envelopes $\{\tilde{H}_{ni}\}$.

Proof of (C). Using independence within rows of the triangular array $\{g_{ni}\}$ we calculate for $(\theta_1,t_1),(\theta_2,t_2) \in [0,1]\times\mathbb{R}$ as follows:

$$\mathbb{E}\Big\{Y_f^{(n)}(\theta_1,t_1)Y_f^{(n)}(\theta_2,t_2)\Big\} = \frac{1}{n\Delta_n}\sum_{i=1}^{\lfloor n(\theta_1\wedge\theta_2)\rfloor}\Big\{\mathbb{E}\big[f^2(\Delta_i^nL^{(n)})\mathbb{1}_{(-\infty,t_1\wedge t_2]}(\Delta_i^nL^{(n)})\big] - \Big(\mathbb{E}\big[f(\Delta_i^nL^{(n)})\mathbb{1}_{(-\infty,t_1]}(\Delta_i^nL^{(n)})\big]\mathbb{E}\big[f(\Delta_i^nL^{(n)})\mathbb{1}_{(-\infty,t_2]}(\Delta_i^nL^{(n)})\big]\Big)\Big\}.$$

Due to Lemma A.20 we have $\mathbb{E}\big[f(\Delta_i^nL^{(n)})\mathbb{1}_{(-\infty,t]}(\Delta_i^n\ L^{(n)})\big] = O(\Delta_n)$ for all $t\in\mathbb{R}$ and $i = 1,\dots,n$. Thus, an application of Proposition 6.3 yields for some small $\delta > 0$ and $n \in \mathbb{N}$ large enough

$$\mathbb{E}\big\{Y_f^{(n)}(\theta_1,t_1)Y_f^{(n)}(\theta_2,t_2)\big\} = \int_0^{\frac{\lfloor n(\theta_1\wedge\theta_2)\rfloor}{n}} \int_{-\infty}^{t_1\wedge t_2} f^2(z)g^{(n)}(y,dz)dy + \\ + O\big(\Delta_n v_n^{-2((\beta+\delta)\wedge 2)} + v_n^{2p-((\beta+\delta)\wedge 2)}\big) + O(\Delta_n)$$

$$= \int_0^{\frac{\lfloor n(\theta_1\wedge\theta_2)\rfloor}{n}} \int_{-\infty}^{t_1\wedge t_2} f^2(z)g_0(y,dz)dy + \frac{1}{\sqrt{n\Delta_n}} \int_0^{\frac{\lfloor n(\theta_1\wedge\theta_2)\rfloor}{n}} \int_{-\infty}^{t_1\wedge t_2} f^2(z)g_1(y,dz)dy \\ + \int_0^{\frac{\lfloor n(\theta_1\wedge\theta_2)\rfloor}{n}} \int_{-\infty}^{t_1\wedge t_2} f^2(z)\mathcal{R}_n(y,dz)dy + o(1),$$

where the final equality above follows using (6.9), as well as $p > \beta$ and $\overline{w} < 1/(2\beta)$. Furthermore, due to Assumption 6.12(a1) and $p > \beta$ the mapping $\big(y \mapsto \int_{-\infty}^{t_1\wedge t_2} f^2(z)g_i(y,dz)\big)$ is Lebesgue-almost surely bounded on $[0,1]$ for each $i \in \{0,1,2\}$. Thus, we have

$$\mathbb{E}\big\{Y_f^{(n)}(\theta_1,t_1)Y_f^{(n)}(\theta_2,t_2)\big\} = \int_0^{\frac{\lfloor n(\theta_1\wedge\theta_2)\rfloor}{n}} \int_{-\infty}^{t_1\wedge t_2} f^2(z)g_0(y,dz)dy + O\big((n\Delta_n)^{-1/2}\big) + o(1) \\ = H_f((\theta_1,t_1);(\theta_2,t_2)) + o(1).$$

Proof of (D). Using Proposition 6.3 we obtain for some small $\delta > 0$

$$\limsup_{n\to\infty} \sum_{i=1}^n \mathbb{E}G_{ni}^2 = \limsup_{n\to\infty} \frac{K}{n\Delta_n} \sum_{i=1}^n \mathbb{E}\Big\{1 \wedge \big|\Delta_i^n L^{(n)}\big|^{2p}\Big\} \\ \leq \limsup_{n\to\infty} \Big\{K\int_0^1 \int \big(1\wedge|z|^{2p}\big)g^{(n)}(y,dz)dy + K\big(\Delta_n v_n^{-2((\beta+\delta)\wedge 2)} + v_n^{2p-((\beta+\delta)\wedge 2)}\big)\Big\} \\ = K\int_0^1 \int \big(1\wedge|z|^{2p}\big)g_0(y,dz)dy < \infty,$$

where the final equality above is a consequence of Assumption 6.12(a1), $p > \beta$ and $\overline{w} < 1/(2\beta)$.

Proof of (E). We have $n\Delta_n \to \infty$. Thus, for $\epsilon > 0$, we can choose

$$N_\epsilon = \min\Big\{m \in \mathbb{N} \Big| \frac{K}{\sqrt{n\Delta_n}} \leq \epsilon \text{ for all } n \geq m\Big\} < \infty.$$

So for $n \geq N_\epsilon$ the integrand satisfies $G_{ni}^2 \mathbb{1}_{\{G_{ni}>\epsilon\}} = 0$ for all $1 \leq i \leq n$ and this yields the assertion.

Proof of (F). Due to symmetry of the semimetrics let $\theta_1 \leq \theta_2$ without loss of generality.

Then an application of Proposition 6.3 gives, for $(\theta_1, t_1), (\theta_2, t_2) \in [0,1] \times \mathbb{R}$,

$$\begin{aligned}
\left(d_f^{(n)}((\theta_1,t_1);(\theta_2,t_2))\right)^2 &= \sum_{i=1}^{n} \mathbb{E}\,|g_{ni}((\theta_1,t_1)) - g_{ni}((\theta_2,t_2))|^2 \\
&= \frac{1}{n\Delta_n} \sum_{i=1}^{\lfloor n\theta_1 \rfloor} \mathbb{E} f^2(\Delta_i^n L^{(n)}) \mathbb{1}_{(t_1\wedge t_2, t_1 \vee t_2]}(\Delta_i^n L^{(n)}) + \\
&\qquad + \frac{1}{n\Delta_n} \sum_{i=\lfloor n\theta_1 \rfloor + 1}^{\lfloor n\theta_2 \rfloor} \mathbb{E} f^2(\Delta_i^n L^{(n)}) \mathbb{1}_{(-\infty, t_2]}(\Delta_i^n L^{(n)}) \\
&= \int_0^{\frac{\lfloor n\theta_1 \rfloor}{n}} \int_{t_1 \wedge t_2}^{t_1 \vee t_2} f^2(z) g^{(n)}(y,dz)dy + \int_{\frac{\lfloor n\theta_1 \rfloor}{n}}^{\frac{\lfloor n\theta_2 \rfloor}{n}} \int_{-\infty}^{t_2} f^2(z) g^{(n)}(y,dz)dy + O(\Delta_n^\alpha) \\
&= \left(d_f((\theta_1,t_1);(\theta_2,t_2))\right)^2 + O(\Delta_n^\alpha)
\end{aligned}$$

for an appropriately small $\alpha > 0$ such that $1/n = o(\Delta_n^\alpha)$ and $(n\Delta_n)^{-1/2} = O(\Delta_n^\alpha)$ according to Assumption 6.12(b7). Moreover, the O-term is uniform in $(\theta_1, t_1), (\theta_2, t_2) \in [0,1] \times \mathbb{R}$. As a consequence,

$$\left| d_f^{(n)}((\theta_1,t_1);(\theta_2,t_2)) - d_f((\theta_1,t_1);(\theta_2,t_2)) \right| = O(\Delta_n^{\alpha/2})$$

uniformly as well, because

$$\left| \sqrt{a} - \sqrt{b} \right| \leq \sqrt{|a-b|}$$

holds for arbitrary $a, b \geq 0$. This uniform convergence implies immediately that for deterministic sequences $((\theta_n^{(1)}, t_n^{(1)}))_{n\in\mathbb{N}}, ((\theta_n^{(2)}, t_n^{(2)}))_{n\in\mathbb{N}} \subset [0,1] \times \mathbb{R}$ with $d_f((\theta_n^{(1)}, t_n^{(1)}); (\theta_n^{(2)}, t_n^{(2)})) \to 0$ we also have $d_f^{(n)}((\theta_n^{(1)}, t_n^{(1)}); (\theta_n^{(2)}, t_n^{(2)})) \to 0$.

Finally, d_f is in fact a semimetric: Define for $(\theta, t) \in [0,1] \times \mathbb{R}$ the random vectors $g_n(\theta,t) = (g_{n1}(\theta,t), \ldots, g_{nn}(\theta,t)) \in \mathbb{R}^n$ and apply first the triangle inequality in $\mathbb{R}^n$ and afterwards the Minkowski inequality to obtain

$$\begin{aligned}
d_f^{(n)}((\theta_1,t_1);(\theta_2,t_2)) &= \left\{\mathbb{E}\|g_n(\theta_1,t_1) - g_n(\theta_2,t_2)\|_2^2\right\}^{1/2} \\
&\leq \left\{\mathbb{E}\left(\|g_n(\theta_1,t_1) - g_n(\theta_3,t_3)\|_2 + \|g_n(\theta_3,t_3) - g_n(\theta_2,t_2)\|_2\right)^2\right\}^{1/2} \\
&\leq \left\{\mathbb{E}\|g_n(\theta_1,t_1) - g_n(\theta_3,t_3)\|_2^2\right\}^{1/2} + \left\{\mathbb{E}\|g_n(\theta_3,t_3) - g_n(\theta_2,t_2)\|_2^2\right\}^{1/2} \\
&= d_f^{(n)}((\theta_1,t_1);(\theta_3,t_3)) + d_f^{(n)}((\theta_3,t_3);(\theta_2,t_2)),
\end{aligned}$$

for $(\theta_1,t_1), (\theta_2,t_2), (\theta_3,t_3) \in [0,1] \times \mathbb{R}$ and $n \in \mathbb{N}$, where $\|\cdot\|_2$ denotes the Euclidean norm in $\mathbb{R}^n$. The triangle inequality for d_f follows immediately. $\square$

The decomposition below is similar to Step 5 in the proof of Theorem 13.1.1 in Jacod and Protter (2012) and it will occur frequently in the sequel. With the constants from Assumption 6.12 let $\ell \in \mathbb{R}$ satisfy

$$1 < \ell < \frac{1}{2\beta\overline{w}} \wedge (1+\epsilon) \quad \text{and also} \quad \ell < \frac{2(p-1)\overline{w} - 1}{2(\beta-1)\overline{w}} \text{ if } \beta > 1,$$

with an $\epsilon > 0$ for which Assumption 6.12(b6) holds. Then we set

$$u_n = (v_n)^\ell \quad \text{and} \quad F_n = \{z \colon |z| > u_n\}$$

as well as

$$\begin{aligned}
\tilde{X}''^n &= (z\mathbb{1}_{F_n}(z)) \star \mu^{(n)}, \\
\tilde{X}''(\alpha)^n &= (z\mathbb{1}_{F_n \cap \{|z| \le \alpha/4\}}(z)) \star \mu^{(n)}, \qquad \text{for } \alpha > 0 \\
\hat{X}''(\alpha)^n &= (z\mathbb{1}_{\{|z| > \alpha/4\}}) \star \mu^{(n)}, \qquad \text{for } \alpha > 0 \\
N_t^n &= (\mathbb{1}_{F_n} \star \mu^{(n)})_t, \\
\tilde{X}_t'^n &= X_t^{(n)} - \tilde{X}_t''^n \\
&= X_0^{(n)} + \int_0^t b_s^{(n)} ds + \int_0^t \sigma_s^{(n)} dW_s^{(n)} + \\
&\qquad\qquad + (z\mathbb{1}_{F_n^C}(z)) \star (\mu^{(n)} - \bar{\mu}^{(n)})_t - (z\mathbb{1}_{\{|z| \le 1\} \cap F_n}(z)) \star \bar{\mu}_t^{(n)}, \\
A_i^n &= \{|\Delta_i^n \tilde{X}'^n| \le v_n/2\} \cap \{\Delta_i^n N^n \le 1\}.
\end{aligned} \tag{6.23}$$

In the following proofs it is necessary to ensure that with high probability at most one large jump occurs and the increments of the remaining part, that is the quantities $\Delta_i^n \tilde{X}'^n$, are small. To this end, we show in Lemma A.4 that $\mathbb{P}(Q_n) \to 1$ as $n \to \infty$ for the sets

$$Q_n = \bigcap_{i=1}^n A_i^n. \tag{6.24}$$

Proof of Lemma 6.2. Let $\alpha > 0$ be fixed and recall the definition of the processes $L^{(n)} = (z\mathbb{1}_{\{|z| > v_n\}}) \star \mu^{(n)}$ in (6.4). Due to Proposition 6.3 and Proposition 6.4 the processes

$$\tilde{Y}_{\rho_\alpha}^{(n)}(\theta, t) = \sqrt{n\Delta_n}\Big\{\frac{1}{n\Delta_n}\sum_{i=1}^{\lfloor n\theta \rfloor} \chi_t^{(\alpha)}(\Delta_i^n L^{(n)}) - N_{\rho_\alpha}(g^{(n)}; \theta, t)\Big\}$$

converge weakly to $\mathbb{G}_{\rho_\alpha}$ in $\ell^\infty([0,1] \times \mathbb{R})$, because by Assumption 6.12(a1) we have $\lambda_1 - \operatorname{ess\,sup}_{y \in [0,1]} \int_{-\infty}^t \rho_\alpha(z) g^{(n)}(y, dz) \le K$ for all $n \in \mathbb{N}$ and thus

$$\begin{aligned}
&\sup_{(\theta,t) \in [0,1] \times \mathbb{R}} \Big| \tilde{Y}_{\rho_\alpha}^{(n)}(\theta, t) - Y_{\rho_\alpha}^{(n)}(\theta, t) \Big| \\
&\qquad \le K\sqrt{n\Delta_n}\Big(\int_{\lfloor n\theta \rfloor/n}^{\theta} \int_{-\infty}^{t} \rho_\alpha(z) g^{(n)}(y, dz) dy + \frac{\lfloor n\theta \rfloor}{n\Delta_n}\big(\Delta_n^2 v_n^{-2((\beta+\delta)\wedge 2)} + \Delta_n v_n^{p-((\beta+\delta)\wedge 2)}\big)\Big) \\
&\qquad = O\Big(\sqrt{\Delta_n/n} + \sqrt{n\Delta_n^{3-4\beta\overline{w}-\delta}} + \sqrt{n\Delta_n^{1+2\overline{w}(p-\beta)-\delta}}\Big) = o(1)
\end{aligned} \tag{6.25}$$

holds for some small $\delta > 0$. The final equality in the display above follows using $1 - 2\beta\overline{w} > 0$, $p - \beta > 1$, as well as Assumption 6.12(b4) and (b6). As a consequence, it suffices to show

$$V_\alpha^{(n)} := \frac{1}{\sqrt{n\Delta_n}} \sup_{(\theta,t) \in [0,1] \times \mathbb{R}} \Big| \sum_{i=1}^{\lfloor n\theta \rfloor} \Big\{\chi_t^{(\alpha)}(\Delta_i^n X^{(n)}) \mathbb{1}_{\{|\Delta_i^n X^{(n)}| > v_n\}} - \chi_t^{(\alpha)}(\Delta_i^n L^{(n)})\Big\} \Big| \xrightarrow{\mathbb{P}} 0. \tag{6.26}$$

According to Lemma A.9 we have for $n \in \mathbb{N}$ large enough such that $v_n \leq \alpha/4$

$$V_\alpha^{(n)} \leq C_n(\alpha) + D_n(\alpha),$$

on Q_n with

$$C_n(\alpha) = \frac{K}{\sqrt{n\Delta_n}} \sup_{t\in\mathbb{R}} \sum_{i=1}^n \big|\mathbb{1}_{(-\infty,t]}(\Delta_i^n \tilde{X}'^n + \Delta_i^n \hat{X}''(\alpha)^n) - \mathbb{1}_{(-\infty,t]}(\Delta_i^n \hat{X}''(\alpha)^n)\big| \times \\ \times \mathbb{1}_{\{|\Delta_i^n \hat{X}''(\alpha)^n| > \alpha/4\}} \mathbb{1}_{\{|\Delta_i^n \tilde{X}'^n| \leq v_n/2\}},$$

$$D_n(\alpha) = \frac{1}{\sqrt{n\Delta_n}} \sum_{i=1}^n \big|\rho_\alpha(\Delta_i^n \tilde{X}'^n + \Delta_i^n \hat{X}''(\alpha)^n)\mathbb{1}_{\{|\Delta_i^n \tilde{X}'^n + \Delta_i^n \hat{X}''(\alpha)^n| > v_n\}} - \\ - \rho_\alpha(\Delta_i^n \hat{X}''(\alpha)^n)\mathbb{1}_{\{|\Delta_i^n \hat{X}''(\alpha)^n| > v_n\}}\big| \mathbb{1}_{\{|\Delta_i^n \tilde{X}'^n| \leq v_n/2\}}, \tag{6.27}$$

where the processes in the display above are defined in (6.23) and where $K > 0$ denotes a bound for ρ. Therefore, due to $\mathbb{P}(Q_n) \to 1$ it is enough to show $C_n(\alpha) = o_{\mathbb{P}}(1)$ and $D_n(\alpha) = o_{\mathbb{P}}(1)$ in order to verify (6.26) and to complete the proof of Lemma 6.2.

First, we consider $D_n(\alpha)$. For later reasons, we let f be either ρ_α or ρ_α°. Then there exists a constant $K > 0$ which depends only on α, such that we have for $x, z \in \mathbb{R}$ and $v > 0$:

$$\big|f(x+z)\mathbb{1}_{\{|x+z|>v\}} - f(x)\mathbb{1}_{\{|x|>v\}}\big|\mathbb{1}_{\{|z|\leq v/2\}} \leq K\big(|x|^p \mathbb{1}_{\{|x|\leq 2v\}} + |x|^{p-1}|z|\mathbb{1}_{\{|z|\leq v/2\}}\big). \tag{6.28}$$

Note that for $|x+z| > v$ and $|x| > v$ we use the mean value theorem and $|z| \leq |x|$ as well as $|\frac{df}{dx}(x)| \leq K|x|^{p-1}$ for all $x \in \mathbb{R}$ by the assumptions on ρ and because the derivatives of Ψ_α and Ψ_α° have a compact support, which is bounded away from 0. In all other cases in which the left hand side does not vanish we have $|z| \leq |x| \leq 2v$ as well as $|f(x)| \leq K|x|^p$ for all $x \in \mathbb{R}$ by another application of the mean value theorem and the assumptions on ρ. Consequently,

$$\mathbb{E}D_n(\alpha) \leq a_n(\alpha) + b_n(\alpha) \tag{6.29}$$

holds for

$$a_n(\alpha) = \frac{1}{\sqrt{n\Delta_n}} \sum_{i=1}^n \mathbb{E}\Big\{\big|\Delta_i^n \hat{X}''(\alpha)^n\big|^p \mathbb{1}_{\{|\Delta_i^n \hat{X}''(\alpha)^n| \leq 2v_n\}}\Big\},$$
$$b_n(\alpha) = \frac{v_n}{2\sqrt{n\Delta_n}} \sum_{i=1}^n \mathbb{E}\big|\Delta_i^n \hat{X}''(\alpha)^n\big|^{p-1},$$

and we conclude $D_n(\alpha) = o_{\mathbb{P}}(1)$ because of Lemma A.17.

Finally, we show $C_n(\alpha) = o_{\mathbb{P}}(1)$. To this end, we define for $1 \leq i, j \leq n$ with $i \neq j$ and the constant $\overline{r}$ in Assumption 6.12

$$R_{i,j}^{(n)}(\alpha) = \Big\{\big|\Delta_i^n \hat{X}''(\alpha)^n - \Delta_j^n \hat{X}''(\alpha)^n\big| \leq \Delta_n^{\overline{r}}\Big\} \cap \Big\{\big|\Delta_i^n \hat{X}''(\alpha)^n\big| > \alpha/4\Big\} \cap Q_n, \tag{6.30}$$

as well as the sets $J_n^{(1)}(\alpha)$ by:

$$J_n^{(1)}(\alpha)^C = \bigcup_{\substack{i,j=1\\ i\neq j}}^{n} R_{i,j}^{(n)}(\alpha). \tag{6.31}$$

Then according to Lemma A.6 we have $\mathbb{P}\big(J_n^{(1)}(\alpha)\big) \to 1$. Moreover, Lemma A.8 shows that for all $n \in \mathbb{N}$, $\omega \in J_n^{(1)}(\alpha) \cap Q_n$ and $t \in \mathbb{R}$ the random set

$$\begin{aligned}\tilde{A}_1(\omega;\alpha,n,t) = \big\{i \in \{1,\dots,n\} \mid |\Delta_i^n \hat{X}''(\alpha)^n(\omega)| > \alpha/4 \text{ and}\\ \mathbb{1}_{(-\infty,t]}(\Delta_i^n \tilde{X}'^n(\omega) + \Delta_i^n \hat{X}''(\alpha)^n(\omega)) \neq \mathbb{1}_{(-\infty,t]}(\Delta_i^n \hat{X}''(\alpha)^n(\omega))\big\}\end{aligned}$$

has at most $c_n := \lceil (v_n/\Delta_n^{\overline{r}}) + 1\rceil$ elements. Consequently, on $J_n^{(1)}(\alpha) \cap Q_n$ for each $t \in \mathbb{R}$ at most c_n summands in the sum of the definition of $C_n(\alpha)$ can be equal to 1 and we conclude

$$C_n(\alpha) \leq K/\sqrt{n\Delta_n^{(1+2(\overline{r}-\overline{w}))\vee 1}},$$

on $J_n^{(1)}(\alpha)\cap Q_n$. Thus, $C_n = o_{\mathbb{P}}(1)$ follows using Assumption 6.12(b7) as well as $\mathbb{P}\big(J_n^{(1)}(\alpha)\cap Q_n\big) \to 1$. □

Proof of Lemma 6.7. For $\alpha > 0$ define the following processes:

$$\tilde{Y}_{\rho_\alpha^\circ}^{(n)}(\theta,t) = \sqrt{n\Delta_n}\Big\{\frac{1}{n\Delta_n}\sum_{i=1}^{\lfloor n\theta\rfloor} \chi_t^{\circ(\alpha)}(\Delta_i^n L^{(n)}) - N_{\rho_\alpha^\circ}(g^{(n)};\theta,t)\Big\}.$$

Similar to (6.25) we obtain with Proposition 6.3 and Proposition 6.4 that for $n \to \infty$ the processes in the display above converge weakly in $\ell^\infty([0,1]\times\mathbb{R})$, that is

$$\tilde{Y}_{\rho_\alpha^\circ}^{(n)} \rightsquigarrow \mathbb{G}_{\rho_\alpha^\circ}.$$

On the other hand, we have weak convergence

$$\mathbb{G}_{\rho_\alpha^\circ} \rightsquigarrow 0$$

in $\ell^\infty([0,1]\times\mathbb{R})$ as $\alpha \to 0$, by Proposition 6.6. Therefore, by using the Portmanteau theorem (Theorem 1.3.4 in Van der Vaart and Wellner (1996)) twice, we obtain for arbitrary $\eta > 0$:

$$\limsup_{\alpha\to 0}\limsup_{n\to\infty}\mathbb{P}\big(\sup_{(\theta,t)\in[0,1]\times\mathbb{R}}\big|\tilde{Y}_{\rho_\alpha^\circ}^{(n)}(\theta,t)\big| \geq \eta\big) \leq \limsup_{\alpha\to 0}\mathbb{P}\big(\sup_{(\theta,t)\in[0,1]\times\mathbb{R}}\big|\mathbb{G}_{\rho_\alpha^\circ}(\theta,t)\big| \geq \eta\big) = 0.$$

Thus, it suffices to show $V_\alpha^{\circ(n)} = o_{\mathbb{P}}(1)$ as $n \to \infty$ for each $\alpha > 0$ in a neighbourhood of 0, where

$$V_\alpha^{\circ(n)} = \frac{1}{\sqrt{n\Delta_n}}\sup_{(\theta,t)\in[0,1]\times\mathbb{R}}\Big|\sum_{i=1}^{\lfloor n\theta\rfloor}\Big\{\chi_t^{\circ(\alpha)}(\Delta_i^n X^{(n)})\mathbb{1}_{\{|\Delta_i^n X^{(n)}|>v_n\}} - \chi_t^{\circ(\alpha)}(\Delta_i^n L^{(n)})\Big\}\Big|. \tag{6.32}$$

Due to Lemma A.10 we have for $\alpha > 0$, $\omega \in Q_n$ and $n \in \mathbb{N}$ large enough such that $v_n \leq \alpha$ with the processes defined in (6.23)

$$V_\alpha^{\circ(n)} \leq C_n^\circ(\alpha) + D_n^\circ(\alpha) + E_n^\circ(\alpha),$$

where

$$C_n^\circ(\alpha) = \frac{K}{\sqrt{n\Delta_n}} \sup_{t\in\mathbb{R}} \sum_{i=1}^n \big|\mathbb{1}_{(-\infty,t]}(\varsigma_i^n(\alpha)) - \mathbb{1}_{(-\infty,t]}(\Delta_i^n \tilde{X}''(8\alpha)^n)\big| \times \\ \times \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n| > \Delta_n^{\overline{v}}\}} \mathbb{1}_{\{|\Delta_i^n \tilde{X}'^n| \leq v_n/2\}},$$

with $K > 0$ a bound for ρ and

$$D_n^\circ(\alpha) = \frac{1}{\sqrt{n\Delta_n}} \sum_{i=1}^n \big|\rho_\alpha^\circ(\varsigma_i^n(\alpha))\mathbb{1}_{\{|\varsigma_i^n(\alpha)|>v_n\}} - \rho_\alpha^\circ(\Delta_i^n \tilde{X}''(8\alpha)^n)\mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n|>v_n\}}\big| \times \\ \times \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n| > \Delta_n^{\overline{v}}\}} \mathbb{1}_{\{|\Delta_i^n \tilde{X}'^n| \leq v_n/2\}}, \tag{6.33}$$

$$E_n^\circ(\alpha) = \frac{1}{\sqrt{n\Delta_n}} \sup_{t\in\mathbb{R}} \sum_{i=1}^n \big|\rho_\alpha^\circ(\varsigma_i^n(\alpha))\mathbb{1}_{\{|\varsigma_i^n(\alpha)|>v_n\}}\mathbb{1}_{(-\infty,t]}(\varsigma_i^n(\alpha)) - \rho_\alpha^\circ(\Delta_i^n \tilde{X}''(8\alpha)^n)\times \\ \times \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n|>v_n\}}\mathbb{1}_{(-\infty,t]}(\Delta_i^n \tilde{X}''(8\alpha)^n)\big|\mathbb{1}_{\{|\Delta_i^n \tilde{X}'^n| \leq v_n/2\}}\mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n| \leq \Delta_n^{\overline{v}}\}}\mathbb{1}_{Q_n},$$

where $\varsigma_i^n(\alpha) = \Delta_i^n \tilde{X}'^n + \Delta_i^n \tilde{X}''(8\alpha)^n$ and $\overline{v} > 0$ is the constant from Assumption 6.12(a(4)I). Thus, as a consequence of $\mathbb{P}(Q_n) \to 1$ it suffices to show for each $\eta > 0$ and each $\alpha > 0$ in a neighbourhood of zero:

$$\lim_{n\to\infty} \mathbb{P}\big(C_n^\circ(\alpha) > \eta\big) = 0, \tag{6.34}$$

$$\lim_{n\to\infty} \mathbb{P}\big(D_n^\circ(\alpha) > \eta\big) = 0, \tag{6.35}$$

$$\lim_{n\to\infty} \mathbb{P}\big(E_n^\circ(\alpha) > \eta\big) = 0. \tag{6.36}$$

Concerning (6.34), similar to (6.30) we define for $1 \leq i,j \leq n$ with $i \neq j$ and the constants $\overline{v} < \overline{r}$ in Assumption 6.12:

$$S_{i,j}^{(n)}(\alpha) = \Big\{\big|\Delta_i^n \tilde{X}''(8\alpha)^n - \Delta_j^n \tilde{X}''(8\alpha)^n\big| \leq \Delta_n^{\overline{r}}\Big\} \cap \Big\{\big|\Delta_i^n \tilde{X}''(8\alpha)^n\big| > \Delta_n^{\overline{v}}\Big\} \cap Q_n.$$

as well as $J_n^{(2)}(\alpha)$ by

$$\big(J_n^{(2)}(\alpha)\big)^C = \bigcup_{\substack{i,j=1 \\ i\neq j}}^n S_{i,j}^{(n)}(\alpha). \tag{6.37}$$

Then Lemma A.7 shows $\mathbb{P}\big(J_n^{(2)}(\alpha)\big) \to 1$ for all $\alpha \in (0, \alpha_0/2)$ with α_0 from Assumption 6.12 and according to Lemma A.8 the random set

$$\tilde{A}_2(\omega;\alpha,n,t) = \big\{i \in \{1,\ldots,n\} \mid |\Delta_i^n \tilde{X}''(8\alpha)^n(\omega)| > \Delta_n^{\overline{v}} \text{ and} \\ \mathbb{1}_{(-\infty,t]}(\Delta_i^n \tilde{X}'^n(\omega) + \Delta_i^n \tilde{X}''(8\alpha)^n(\omega)) \neq \mathbb{1}_{(-\infty,t]}(\Delta_i^n \tilde{X}''(8\alpha)^n(\omega))\big\}$$

has at most $c_n = \lceil (v_n/\Delta_n^{\overline{r}}) + 1\rceil$ elements for all $\omega \in J_n^{(2)}(\alpha) \cap Q_n$, $n \in \mathbb{N}$, $t \in \mathbb{R}$ and $\alpha > 0$. So for each $t \in \mathbb{R}$ at most c_n summands in $C_n^\circ(\alpha)$ can be equal to 1, and we have

$$C_n^\circ(\alpha) \leq K/\sqrt{n\Delta_n^{(1+2(\overline{r}-\overline{w}))\vee 1}}$$

on $J_n^{(2)}(\alpha) \cap Q_n$. Consequently, (6.34) follows by Assumption 6.12(b7) for all $\alpha \in (0, \alpha_0/2)$.

Furthermore, because of (6.28) we have

$$\mathbb{E} D_n^\circ(\alpha) \le c_n(\alpha) + d_n(\alpha), \tag{6.38}$$

for

$$c_n(\alpha) = \frac{1}{\sqrt{n\Delta_n}} \sum_{i=1}^n \mathbb{E}\Big\{ \big|\Delta_i^n \tilde{X}''(8\alpha)^n\big|^p \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n| \le 2v_n\}} \Big\},$$
$$d_n(\alpha) = \frac{v_n}{2\sqrt{n\Delta_n}} \sum_{i=1}^n \mathbb{E}\big|\Delta_i^n \tilde{X}''(8\alpha)^n\big|^{p-1}.$$

Thus, Lemma A.18 yields (6.35).

Concerning (6.36), let $\alpha > 0$ be fixed. Because of the triangle inequality and $|\rho_\alpha^\circ(z)| \le K|z|^p$ for all $z \in \mathbb{R}$, an upper bound for $E_n^\circ(\alpha)$ is clearly given by

$$\frac{K}{\sqrt{n\Delta_n}} \sum_{i=1}^n \Big(|\varsigma_i^n(\alpha)|^p \mathbb{1}_{\{|\varsigma_i^n(\alpha)| > v_n\}} + |\Delta_i^n \tilde{X}''(8\alpha)^n|^p \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n| > v_n\}} \Big) \times$$
$$\times \mathbb{1}_{\{|\Delta_i^n \tilde{X}'^n| \le v_n/2\}} \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n| \le \Delta_n^{\overline{v}}\}} \mathbb{1}_{Q_n},$$

with $\varsigma_i^n(\alpha) = \Delta_i^n \tilde{X}'^n + \Delta_i^n \tilde{X}''(8\alpha)^n$. As a consequence, we have $\mathbb{E} E_n^\circ(\alpha) \le K\big(y_n^{(\alpha)} + 2z_n^{(\alpha)}\big)$ for

$$y_n^{(\alpha)} = \frac{1}{\sqrt{n\Delta_n}} \sum_{i=1}^n \mathbb{E}\Big\{ \big| |\varsigma_i^n(\alpha)|^p - |\Delta_i^n \tilde{X}''(8\alpha)^n|^p \big| \times$$
$$\times \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n| \le \Delta_n^{\overline{v}}\}} \mathbb{1}_{\{|\varsigma_i^n(\alpha)| > v_n\}} \mathbb{1}_{\{|\Delta_i^n \tilde{X}'^n| \le v_n/2\}} \mathbb{1}_{Q_n} \Big\},$$
$$z_n^{(\alpha)} = \frac{1}{\sqrt{n\Delta_n}} \sum_{i=1}^n \mathbb{E}\Big\{ \big|\Delta_i^n \tilde{X}''(8\alpha)^n\big|^p \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n| \le \Delta_n^{\overline{v}}\}} \mathbb{1}_{Q_n} \Big\}.$$

Therefore, we obtain (6.36) by Lemma A.19. □

Proof of Theorem 6.1. In order to establish weak convergence we use Theorem 1.12.2 in Van der Vaart and Wellner (1996). It is sufficient to prove

$$\mathbb{E}^* h(G_\rho^{(n)}) \to \mathbb{E} h(\mathbb{G}_\rho)$$

for each bounded Lipschitz function $h \in \mathrm{BL}_1(\ell^\infty([0,1] \times \mathbb{R}))$, where $\mathrm{BL}_1(\mathbb{D})$ for a metric space $(\mathbb{D}, d)$ was introduced in Definition 3.4. Here, we use that the tight process $\mathbb{G}_\rho$ is also separable (see Lemma 1.3.2 in Van der Vaart and Wellner (1996)).

Thus, let $h \in \mathrm{BL}_1(\ell^\infty([0,1] \times \mathbb{R}))$ and $\delta > 0$. Then we choose $\alpha > 0$ with

$$\limsup_{n\to\infty} \mathbb{P}\Big(\sup_{(\theta,t)\in[0,1]\times\mathbb{R}} |G_{\rho,n}^{\circ(\alpha)}(\theta,t)| > \delta/6 \Big) < \delta/12 \tag{6.39}$$

and

$$|\mathbb{E} h(\mathbb{G}_{\rho_\alpha}) - \mathbb{E} h(\mathbb{G}_\rho)| \le \delta/3. \tag{6.40}$$

(6.39) is possible using Lemma 6.7, and Lemma 6.5 allows (6.40). For this $\alpha > 0$ choose an $N \in \mathbb{N}$ with

$$\left|\mathbb{E}^* h(G_{\rho,n}^{(\alpha)}) - \mathbb{E}h(\mathbb{G}_{\rho_\alpha})\right| \leq \delta/3,$$

for $n \geq N$. This is possible due to Lemma 6.2. Now, because of the previous inequalities and the Lipschitz property of h, we have for $n \in \mathbb{N}$ large enough:

$$\begin{aligned}\left|\mathbb{E}^* h(G_\rho^{(n)}) - \mathbb{E}h(\mathbb{G}_\rho)\right| &\leq \\ &\leq \mathbb{E}^*\left|h(G_\rho^{(n)}) - h(G_{\rho,n}^{(\alpha)})\right| + \left|\mathbb{E}^* h(G_{\rho,n}^{(\alpha)}) - \mathbb{E}h(\mathbb{G}_{\rho_\alpha})\right| + \left|\mathbb{E}h(\mathbb{G}_{\rho_\alpha}) - \mathbb{E}h(\mathbb{G}_\rho)\right| < \delta.\end{aligned}$$

□

6.5 Proof of Theorem 6.8

Proof of Proposition 6.11. Recall the triangular array $\{g_{ni}(\theta,t) \mid n \in \mathbb{N}, i = 1,\dots,n; (\theta,t) \in [0,1]\times\mathbb{R}\}$ in the proof of Proposition 6.4 given by

$$g_{ni}(\omega;(\theta,t)) = \frac{1}{\sqrt{n\Delta_n}} f(\Delta_i^n L^{(n)}(\omega))\mathbb{1}_{(-\infty,t]}(\Delta_i^n L^{(n)}(\omega))\mathbb{1}_{\{i\leq\lfloor n\theta\rfloor\}},$$

for $(\theta,t) \in [0,1]\times\mathbb{R}$ and let $\mu_{ni}(\theta,t) = \mathbb{E}\big(g_{ni}(\theta,t)\big)$. Moreover, for $n \in \mathbb{N}$, $i = 1,\dots,n$ and $(\theta,t) \in [0,1]\times\mathbb{R}$ let

$$\hat{\mu}_{ni}(\theta,t) = \frac{1}{\sqrt{n\Delta_n}}\mathbb{1}_{\{i\leq\lfloor n\theta\rfloor\}}\tilde{\eta}_f^{(n)}(t),$$

with

$$\tilde{\eta}_f^{(n)}(t) = \frac{1}{n}\sum_{j=1}^n f(\Delta_j^n L^{(n)})\mathbb{1}_{(-\infty,t]}(\Delta_j^n L^{(n)}),$$

be an estimator for $\mu_{ni}(\theta,t)$. Then we proceed in two steps:

(a)

$$\hat{Y}_{f,0}^{(n)}(\theta,t) := \sum_{i=1}^n \xi_i\big(g_{ni}(\theta,t) - \hat{\mu}_{ni}(\theta,t)\big) \rightsquigarrow_\xi \mathbb{G}_f, \tag{6.41}$$

(b)

$$\sup_{(\theta,t)\in[0,1]\times\mathbb{R}}\left|\hat{Y}_f^{(n)}(\theta,t) - \hat{Y}_{f,0}^{(n)}(\theta,t)\right| = \sup_{(\theta,t)\in[0,1]\times\mathbb{R}}\left|\sum_{i=1}^n \xi_i\hat{\mu}_{ni}(\theta,t)\right| = o_{\mathbb{P}}(1),$$

then the claim follows using Lemma C.1.

Proof of Step (a). $\{g_{ni}\}$ satisfies conditions (A)-(F) in the proof of Proposition 6.4. Thus, the conditional weak convergence (6.41) holds by Theorem 11.18 in Kosorok (2008), if we can show the following four conditions of the triangular array $\{\hat{\mu}_{ni}(\theta,t) \mid n \in \mathbb{N}, i = 1,\dots,n; (\theta,t) \in [0,1]\times\mathbb{R}\}$:

(G) $\{\hat{\mu}_{ni}\}$ is AMS;

(H)

$$\sup_{(\theta,t)\in[0,1]\times\mathbb{R}} \sum_{i=1}^{n} \left[\hat{\mu}_{ni}(\omega;(\theta,t)) - \mu_{ni}(\theta,t)\right]^2 = o_{\mathbb{P}}(1);$$

(I) The processes $\{\hat{\mu}_{ni}\}$ are manageable with envelopes $\{\hat{F}_{ni}\}$ given by

$$\hat{F}_{ni} = \frac{K}{n\sqrt{n\Delta_n}} \sum_{j=1}^{n} 1 \wedge |\Delta_j^n L^{(n)}|^p,$$

for $n \in \mathbb{N}$ and $i = 1, \ldots, n$, with a $K > 0$ such that $|f(z)| \leq K(1 \wedge |z|^p)$ for all $z \in \mathbb{R}$;

(J) There exists an $M \in \mathbb{R}_+$ such that

$$M \vee \sum_{i=1}^{n} [\hat{F}_{ni}]^2 \xrightarrow{\mathbb{P}} M.$$

Proof of (G). For each $n \in \mathbb{N}$ define the countable set $S_n = ([0,1] \times \mathbb{R}) \cap \mathbb{Q}^2$ to obtain

$$\mathbb{P}^*\Bigg(\sup_{(\theta_1,t_1)\in[0,1]\times\mathbb{R}} \inf_{(\theta_2,t_2)\in S_n} \sum_{i=1}^{n} (\hat{\mu}_{ni}(\omega;(\theta_1,t_1)) - \hat{\mu}_{ni}(\omega;(\theta_2,t_2)))^2 > 0\Bigg) = 0.$$

As a consequence, the triangular array $\{\hat{\mu}_{ni}\}$ is separable and therefore AMS by Lemma 11.15 in Kosorok (2008).

Proof of (H). Simple calculations show

$$\begin{aligned}
A_n :=& \sup_{(\theta,t)\in[0,1]\times\mathbb{R}} \sum_{i=1}^{n} \left[\hat{\mu}_{ni}(\omega;(\theta,t)) - \mu_{ni}(\theta,t)\right]^2 \\
=& \frac{1}{n^3\Delta_n} \sup_{t\in\mathbb{R}} \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} \Big(f(\Delta_j^n L^{(n)}) \mathbb{1}_{(-\infty,t]}(\Delta_j^n L^{(n)}) \\
&\qquad\qquad - \mathbb{E}\big(f(\Delta_i^n L^{(n)}) \mathbb{1}_{(-\infty,t]}(\Delta_i^n L^{(n)})\big)\Big) \times \\
&\qquad \times \Big(f(\Delta_k^n L^{(n)}) \mathbb{1}_{(-\infty,t]}(\Delta_k^n L^{(n)}) - \mathbb{E}\big(f(\Delta_i^n L^{(n)}) \mathbb{1}_{(-\infty,t]}(\Delta_i^n L^{(n)})\big)\Big) \\
=& \sup_{t\in\mathbb{R}} \Bigg\{ \frac{1}{n^2\Delta_n} \sum_{j=1}^{n}\sum_{k=1}^{n} f(\Delta_j^n L^{(n)}) \mathbb{1}_{(-\infty,t]}(\Delta_j^n L^{(n)}) f(\Delta_k^n L^{(n)}) \mathbb{1}_{(-\infty,t]}(\Delta_k^n L^{(n)}) \\
&\quad - \frac{1}{n^2\Delta_n} \sum_{i=1}^{n}\sum_{k=1}^{n} f(\Delta_k^n L^{(n)}) \mathbb{1}_{(-\infty,t]}(\Delta_k^n L^{(n)}) \mathbb{E}\big(f(\Delta_i^n L^{(n)}) \mathbb{1}_{(-\infty,t]}(\Delta_i^n L^{(n)})\big) \\
&\quad - \frac{1}{n^2\Delta_n} \sum_{i=1}^{n}\sum_{j=1}^{n} f(\Delta_j^n L^{(n)}) \mathbb{1}_{(-\infty,t]}(\Delta_j^n L^{(n)}) \mathbb{E}\big(f(\Delta_i^n L^{(n)}) \mathbb{1}_{(-\infty,t]}(\Delta_i^n L^{(n)})\big) \\
&\quad + \frac{1}{n\Delta_n} \sum_{i=1}^{n} \big(\mathbb{E}\big(f(\Delta_i^n L^{(n)}) \mathbb{1}_{(-\infty,t]}(\Delta_i^n L^{(n)})\big)\big)^2 \Bigg\}.
\end{aligned} \tag{6.42}$$

Furthermore, by Lemma A.20 and the assumptions on f we obtain

$$\sup_{i\in\{1,\dots,n\}} \mathbb{E}\big(\sup_{t\in\mathbb{R}} \big|f(\Delta_i^n L^{(n)})\mathbb{1}_{(-\infty,t]}(\Delta_i^n L^{(n)})\big|\big) = O(\Delta_n), \tag{6.43}$$

$$\mathbb{E}\sup_{t\in\mathbb{R}}\Big|\frac{1}{n^2\Delta_n}\sum_{i=1}^n\sum_{k=1}^n f(\Delta_k^n L^{(n)})\mathbb{1}_{(-\infty,t]}(\Delta_k^n L^{(n)})\mathbb{E}\big(f(\Delta_i^n L^{(n)})\mathbb{1}_{(-\infty,t]}(\Delta_i^n L^{(n)})\big)\Big|$$
$$\leq \frac{K}{n}\sum_{k=1}^n \mathbb{E}\big(1\wedge|\Delta_k^n L^{(n)}|^p\big) = O(\Delta_n). \tag{6.44}$$

Thus (6.42), (6.43) and (6.44) give

$$0 \leq A_n \leq \sup_{t\in\mathbb{R}}\Big\{\frac{1}{n^2\Delta_n}\sum_{j=1}^n\sum_{k=1}^n |f(\Delta_j^n L^{(n)})|\mathbb{1}_{(-\infty,t]}(\Delta_j^n L^{(n)})\times$$
$$\times|f(\Delta_k^n L^{(n)})|\mathbb{1}_{(-\infty,t]}(\Delta_k^n L^{(n)})\Big\} + o_{\mathbb{P}}(1)$$
$$= \frac{1}{n}\sup_{t\in\mathbb{R}}\big(Y_{|f|}^{(n)}(1,t)\big)^2 + o_{\mathbb{P}}(1) = o_{\mathbb{P}}(1),$$

because $Y_{|f|}^{(n)}$ converges weakly to the tight process $\mathbb{G}_{|f|}$ in $\ell^\infty([0,1]\times\mathbb{R})$ by Proposition 6.4.

Proof of (I). According to Theorem 11.17 in Kosorok (2008), it suffices to verify that the triangular arrays

$$\{\tilde{\mu}_{ni}(\omega;t) := \frac{1}{\sqrt{n\Delta_n}}\tilde{\eta}_f^{(n)}(t) \mid n\in\mathbb{N}; i=1,\dots,n; t\in\mathbb{R}\},$$

and

$$\{\tilde{h}_{ni}(\omega;\theta) := \mathbb{1}_{\{i\leq\lfloor n\theta\rfloor\}} \mid n\in\mathbb{N}; i=1,\dots,n;\theta\in[0,1]\}$$

are manageable with envelopes $\{\hat{F}_{ni} \mid n\in\mathbb{N}; i=1,\dots,n\}$ and $\{\tilde{H}_{ni}(\omega) :\equiv 1 \mid n\in\mathbb{N}; i=1,\dots,n\}$, respectively. The manageability of the triangular array $\{\tilde{h}_{ni}\}$ has already been shown in the proof of Proposition 6.4. Concerning the triangular array $\{\tilde{\mu}_{ni}\}$ we consider for $n\in\mathbb{N}$ and $\omega\in\Omega$ the sets

$$\mathcal{F}_{n\omega} = \big\{\big(\tilde{\mu}_{n1}(\omega;t),\dots,\tilde{\mu}_{nn}(\omega;t)\big) \mid t\in\mathbb{R}\big\} \subset \mathbb{R}^n,$$

which are bounded with envelope vector $(\hat{F}_{n1}(\omega),\dots,\hat{F}_{nn}(\omega))$. But $\tilde{\mu}_{ni}$ does not depend on i, such that every coordinate projection of $\mathcal{F}_{n\omega}$ onto two coordinates $i_1,i_2\in\{1,\dots,n\}$ is a subset of the straight line $\{(x,y)\in\mathbb{R}^2 \mid x=y\}$. Consequently, in the sense of Definition 4.2 in Pollard (1990), for every $s\in\mathbb{R}^2$ no proper coordinate projection of $\mathcal{F}_{n\omega}$ can surround s and therefore $\mathcal{F}_{n\omega}$ has a pseudo dimension of at most 1 (Definition 4.3 in Pollard (1990)). Now the managebility of the triangular array $\{\tilde{\mu}_{ni}\}$ follows with the same reasoning as in the verification of (B) in the proof of Proposition 6.4.

Proof of (J). The envelopes $\{\hat{F}_{ni}\}$ are independent of i as well. Therefore, with Lemma A.20 we obtain

$$\mathbb{E}\Big\{\sum_{i=1}^n[\hat{F}_{ni}]^2\Big\} = n\mathbb{E}[\hat{F}_{n1}]^2 = \frac{1}{n^2\Delta_n}\mathbb{E}\Big\{\sum_{i=1}^n\sum_{j=1}^n\big(1\wedge|\Delta_i^n L^{(n)}|^p\big)\big(1\wedge|\Delta_j^n L^{(n)}|^p\big)\Big\}$$
$$= O\big(\Delta_n + 1/n\big),$$

because the processes $L^{(n)}$ have independent increments. As a consequence, we have in fact $\sum_{i=1}^n [\hat{F}_{ni}]^2 = o_{\mathbb{P}}(1)$, which proves the claim.

Proof of Step (b). We have

$$\sum_{i=1}^n \xi_i \hat{\mu}_{ni}(\theta, t) = U_n(t) \frac{1}{\sqrt{n}} \sum_{i=1}^{\lfloor n\theta \rfloor} \xi_i,$$

where

$$U_n(t) = \frac{1}{n\sqrt{\Delta_n}} \sum_{j=1}^n f(\Delta_j^n L^{(n)}) \mathbb{1}_{(-\infty,t]}(\Delta_j^n L^{(n)}).$$

As an immediate consequence of Lemma A.20 we obtain $\sup_{t\in\mathbb{R}} |U_n(t)| = o_{\mathbb{P}}(1)$. Furthermore, the $(\xi_i)_{i\in\mathbb{N}}$ are i.i.d. with mean zero and variance one, so it is well known from empirical process theory (see for instance Theorem 2.5.2 and Theorem 2.12.1 in Van der Vaart and Wellner (1996)) that $1/\sqrt{n} \times \sum_{i=1}^{\lfloor n\theta \rfloor} \xi_i$ converges weakly to a Brownian motion in $\ell^\infty([0,1])$. The law of a Brownian motion is tight in $\ell^\infty([0,1])$ (see for example Section 8 in Billingsley (1999)) and thus $U_n(t)/\sqrt{n} \times \sum_{i=1}^{\lfloor n\theta \rfloor} \xi_i$ converges to 0 in $\ell^\infty([0,1]\times\mathbb{R})$ in outer probability. □

Proof of Lemma 6.9. By Proposition 6.11 we have $\hat{Y}_{\rho_\alpha}^{(n)} \rightsquigarrow_\xi \mathbb{G}_{\rho_\alpha}$ for each fixed $\alpha > 0$ and therefore due to Lemma C.1 it only remains to show that the term

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

for $(\theta,t) \in [0,1]\times\mathbb{R}$, converges to 0 in $\ell^\infty([0,1]\times\mathbb{R})$ in outer probability. Consequently, it suffices to show

$$\hat{V}_\alpha^{(n)} \xrightarrow{\mathbb{P}} 0,$$

for each fixed $\alpha > 0$, where

$$\hat{V}_\alpha^{(n)} = \frac{1}{\sqrt{n\Delta_n}} \sup_{(\theta,t)\in[0,1]\times\mathbb{R}} \Big| \sum_{i=1}^{\lfloor n\theta \rfloor} \xi_i \big(\chi_t^{(\alpha)}(\Delta_i^n X^{(n)}) \mathbb{1}_{\{|\Delta_i^n X^{(n)}| > v_n\}} - \chi_t^{(\alpha)}(\Delta_i^n L^{(n)})\big) \Big|. \tag{6.45}$$

To this end, Lemma A.11 yields the estimate

$$\hat{V}_\alpha^{(n)} \le \hat{D}_n(\alpha) + \hat{E}_n(\alpha) + \hat{F}_n(\alpha),$$

on $J_n^{(1)}(\alpha) \cap Q_n$ for $n \in \mathbb{N}$ large enough such that $v_n \le \alpha/4$, where Q_n is defined in (6.24), $J_n^{(1)}(\alpha)$ is defined in (6.31) and with

$$\hat{D}_n(\alpha) = \frac{1}{\sqrt{n\Delta_n}} \sum_{i=1}^n |\xi_i| \big| \rho_\alpha(\Delta_i^n \tilde{X}'^n + \Delta_i^n \hat{X}''(\alpha)^n) \mathbb{1}_{\{|\Delta_i^n \tilde{X}'^n + \Delta_i^n \hat{X}''(\alpha)^n| > v_n\}} - \rho_\alpha(\Delta_i^n \hat{X}''(\alpha)^n) \mathbb{1}_{\{|\Delta_i^n \hat{X}''(\alpha)^n| > v_n\}} \big| \mathbb{1}_{\{|\Delta_i^n \tilde{X}'^n| \le v_n/2\}},$$

$$\hat{E}_n(\alpha) = \sup_{A\in\mathfrak{S}_n} \Big| \sum_{i\in A} \xi_i a_i^n(\alpha) \Big|,$$

$$\hat{F}_n(\alpha) = \sup_{A \in \mathfrak{S}_n} \Big| \sum_{i \in A} \xi_i b_i^n(\alpha) \Big|,$$

for

$$a_i^n(\alpha) = \frac{1}{\sqrt{n\Delta_n}} \rho_\alpha(\Delta_i^n \hat{X}''(\alpha)^n) \mathbb{1}_{\{|\Delta_i^n \hat{X}''(\alpha)^n| > \alpha/4\}},$$

$$b_i^n(\alpha) = \frac{1}{\sqrt{n\Delta_n}} \rho_\alpha(\Delta_i^n \tilde{X}'^n + \Delta_i^n \hat{X}''(\alpha)^n) \mathbb{1}_{\{|\Delta_i^n \tilde{X}'^n + \Delta_i^n \hat{X}''(\alpha)^n| > v_n\}} \mathbb{1}_{\{|\Delta_i^n \hat{X}''(\alpha)^n| > \alpha/4\}},$$

where the quantities in the display above are introduced in (6.23) and where $\mathfrak{S}_n = \{M \subset \{1, \ldots, n\} \mid \#M \leq c_n\}$ with $c_n = \lceil (v_n/\Delta_n^{\overline{r}}) + 1 \rceil$. Lemma A.4 and Lemma A.6 show $\mathbb{P}(J_n^{(1)}(\alpha) \cap Q_n) \to 1$ and thus it is further enough to verify

$$\hat{D}_n(\alpha) = o_{\mathbb{P}}(1), \tag{6.46}$$

$$\hat{E}_n(\alpha) = o_{\mathbb{P}}(1), \tag{6.47}$$

$$\hat{F}_n(\alpha) = o_{\mathbb{P}}(1), \tag{6.48}$$

for each $\alpha > 0$ as $n \to \infty$.
Recall the quantity $D_n(\alpha)$ introduced in (6.27). (6.29) and Lemma A.17 yield $\mathbb{E}D_n(\alpha) \to 0$. Moreover, the bootstrap multipliers have variance 1 and satisfy therefore $\mathbb{E}|\xi_i| \leq 1$ for all $i \in \mathbb{N}$. Thus because of the independence of the multipliers and the other involved processes we obtain $0 \leq \mathbb{E}\hat{D}_n(\alpha) \leq \mathbb{E}D_n(\alpha) \to 0$, which proves (6.46).
Concerning (6.47) we have for $n \in \mathbb{N}$ large enough

$$\mathbb{E}|a_i^n(\alpha)|^m \leq \Big(\frac{K(\alpha)}{\sqrt{n\Delta_n}} \Big)^m \Delta_n,$$

for some $K(\alpha) > 0$, all $m \in \mathbb{N}$ and all $i = 1, \ldots, n$ by virtue of Lemma A.14. Thus, using Assumption 3.6 as well as independence of ξ_i and $a_i^n(\alpha)$ we obtain for every integer $m \geq 2$ and $n \in \mathbb{N}$ large enough

$$\mathbb{E}|Z_i^n(\alpha)|^m \leq m! \Big(\frac{C_1}{\sqrt{n\Delta_n}} \Big)^{m-2} \frac{C_2}{n},$$

for some constants $C_1, C_2 > 0$ and where

$$Z_i^n(\alpha) = \xi_i a_i^n(\alpha)$$

for $n \in \mathbb{N}$, $\alpha > 0$, $i = 1, \ldots, n$. Furthermore, due to the definition of $\hat{X}''(\alpha)^n$ in (6.23) the variables $(Z_i^n(\alpha))_{i=1,\ldots,n}$ are independent with mean zero. Consequently, Lemma A.16 shows

$$\mathbb{E}\hat{E}_n(\alpha) = \mathbb{E}\Big\{ \sup_{A \in \mathfrak{S}_n} \Big| \sum_{i \in A} Z_i^n(\alpha) \Big| \Big\} = o(1),$$

which proves (6.47).
In order to show (6.48) observe first that for $n \in \mathbb{N}$ large enough

$$\mathbb{1}_{\{|\Delta_i^n \tilde{X}'^n + \Delta_i^n \hat{X}''(\alpha)^n| > v_n\}} \mathbb{1}_{\{|\Delta_i^n \hat{X}''(\alpha)^n| > \alpha/4\}} = \mathbb{1}_{\{|\Delta_i^n \hat{X}''(\alpha)^n| > \alpha/4\}}$$

holds for each $i = 1, \ldots, n$ on the set Q_n, because by (6.24) we have $|\Delta_i^n \tilde{X}'^n| \leq v_n/2$ on Q_n. Therefore, we obtain from the mean value theorem for large n on the set Q_n

$$b_i^n(\alpha) = a_i^n(\alpha) + \frac{1}{\sqrt{n\Delta_n}} \mathbb{1}_{\{|\Delta_i^n \hat{X}''(\alpha)^n| > \alpha/4\}} \mathbb{1}_{\{|\Delta_i^n \tilde{X}'^n| \leq v_n/2\}} \Delta_i^n \tilde{X}'^n \Big(\frac{d\rho_\alpha}{dx}\Big)(\zeta_i^n(\alpha))$$

for some $\zeta_i^n(\alpha)$ between $\Delta_i^n \hat{X}''(\alpha)^n$ and $\Delta_i^n \tilde{X}'^n + \Delta_i^n \hat{X}''(\alpha)^n$. Thus, the indicators and Assumption 6.12(a2) show for large $n \in \mathbb{N}$

$$\begin{aligned} \big|\hat{E}_n(\alpha) - \hat{F}_n(\alpha)\big| \mathbb{1}_{Q_n} &\leq \sup_{A \in \mathfrak{S}_n} \frac{K}{\sqrt{n\Delta_n}} \sum_{i \in A} |\xi_i| |\Delta_i^n \hat{X}''(\alpha)^n|^{p-1} |\Delta_i^n \tilde{X}'^n| \times \\ &\qquad\qquad \times \mathbb{1}_{\{|\Delta_i^n \hat{X}''(\alpha)^n| > \alpha/4\}} \mathbb{1}_{\{|\Delta_i^n \tilde{X}'^n| \leq v_n/2\}} \\ &\leq \frac{K v_n}{\sqrt{n\Delta_n}} \sum_{i=1}^n |\xi_i| |\Delta_i^n \hat{X}''(\alpha)^n|^{p-1}. \end{aligned} \tag{6.49}$$

The bootstrap multipliers are defined on a distinct probability space and satisfy $\mathbb{E}|\xi_i| \leq 1$ for all $i \in \mathbb{N}$. As a consequence, (6.49) gives

$$\mathbb{E}\big|\hat{E}_n(\alpha) - \hat{F}_n(\alpha)\big| \mathbb{1}_{Q_n} \leq K b_n(\alpha),$$

with

$$b_n(\alpha) = \frac{v_n}{2\sqrt{n\Delta_n}} \sum_{i=1}^n \mathbb{E}\big|\Delta_i^n \hat{X}''(\alpha)^n\big|^{p-1}.$$

Therefore, (6.48) follows from (6.47), Lemma A.4 and Lemma A.17. $\square$

Proof of Lemma 6.10. Due to Proposition 6.11 we have $\hat{Y}^{(n)}_{\rho_\alpha^\circ} \rightsquigarrow_\xi \mathbb{G}_{\rho_\alpha^\circ}$ in $\ell^\infty([0,1] \times \mathbb{R})$ for every $\alpha > 0$. Thus, according to Lemma C.1 it suffices to show

$$\sup_{(\theta,t) \in [0,1] \times \mathbb{R}} \big|\hat{G}^{\circ(\alpha)}_{\rho,n} - \hat{Y}^{(n)}_{\rho_\alpha^\circ}\big| = o_{\mathbb{P}}(1)$$

for each $\alpha > 0$ in a neighbourhood of zero. Simple manipulations give

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

for $(\theta,t) \in [0,1] \times \mathbb{R}$. As a consequence, it only remains to show that

$$\hat{V}_\alpha^{\circ(n)} \xrightarrow{\mathbb{P}} 0, \tag{6.50}$$

for each $\alpha > 0$ in a neighbourhood of 0 with

$$\hat{V}_\alpha^{\circ(n)} = \frac{1}{\sqrt{n\Delta_n}} \sup_{(\theta,t) \in [0,1] \times \mathbb{R}} \Big| \sum_{i=1}^{\lfloor n\theta \rfloor} \xi_i \big(\chi_t^{\circ(\alpha)}(\Delta_i^n X^{(n)}) \mathbb{1}_{\{|\Delta_i^n X^{(n)}| > v_n\}} - \chi_t^{\circ(\alpha)}(\Delta_i^n L^{(n)})\big)\Big|.$$

Due to Lemma A.12 for each $\alpha > 0$ and $n \in \mathbb{N}$ large enough such that $v_n \leq \alpha$ we have the bound

$$\hat{V}_\alpha^{\circ(n)} \leq \hat{C}_n^\circ(\alpha) + \hat{D}_n^\circ(\alpha) + \hat{E}_n^\circ(\alpha) + \hat{F}_n^\circ(\alpha)$$

on $J_n^{(2)}(\alpha) \cap Q_n$, where Q_n is defined in (6.24), $J_n^{(2)}(\alpha)$ is defined in (6.37) and with

$$\hat{C}_n^\circ(\alpha) = \frac{1}{\sqrt{n\Delta_n}} \sup_{t\in\mathbb{R}} \sum_{i=1}^n |\xi_i| \big| \rho_\alpha^\circ(\varsigma_i^n(\alpha)) \mathbb{1}_{\{|\varsigma_i^n(\alpha)|>v_n\}} \mathbb{1}_{(-\infty,t]}(\varsigma_i^n(\alpha)) - \\ - \rho_\alpha^\circ(\Delta_i^n \tilde{X}''(8\alpha)^n) \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n|>v_n\}} \mathbb{1}_{(-\infty,t]}(\Delta_i^n \tilde{X}''(8\alpha)^n) \big| \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n| \le \Delta_n^{\overline{v}}\}} \mathbb{1}_{\{|\Delta_i^n \tilde{X}'^n| \le v_n/2\}},$$

$$\hat{D}_n^\circ(\alpha) = \frac{1}{\sqrt{n\Delta_n}} \sum_{i=1}^n |\xi_i| \big| \rho_\alpha^\circ(\varsigma_i^n(\alpha)) \mathbb{1}_{\{|\varsigma_i^n(\alpha)|>v_n\}} - \rho_\alpha^\circ(\Delta_i^n \tilde{X}''(8\alpha)^n) \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n|>v_n\}} \big| \times \\ \times \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n| > \Delta_n^{\overline{v}}\}} \mathbb{1}_{\{|\Delta_i^n \tilde{X}'^n| \le v_n/2\}},$$

$$\hat{E}_n^\circ(\alpha) = \sup_{A\in\mathfrak{S}_n} \Big| \sum_{i\in A} \xi_i \bar{a}_i^n(\alpha) \Big|,$$

$$\hat{F}_n^\circ(\alpha) = \sup_{A\in\mathfrak{S}_n} \Big| \sum_{i\in A} \xi_i \bar{b}_i^n(\alpha) \Big|,$$

where the processes involved in the display above are introduced in (6.23), $\overline{v} > 0$ is the constant from Assumption 6.12(a(4)I), $\varsigma_i^n(\alpha) = \Delta_i^n \tilde{X}'^n + \Delta_i^n \tilde{X}''(8\alpha)^n$, $\mathfrak{S}_n = \{M \subset \{1,\ldots,n\} \mid \#M \le c_n\}$ for $c_n = \lceil (v_n/\Delta_n^{\overline{r}}) + 1 \rceil$ and with

$$\bar{a}_i^n(\alpha) = \frac{1}{\sqrt{n\Delta_n}} \rho_\alpha^\circ(\Delta_i^n \tilde{X}''(8\alpha)^n) \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n| > v_n \vee \Delta_n^{\overline{v}}\}},$$
$$\bar{b}_i^n(\alpha) = \frac{1}{\sqrt{n\Delta_n}} \rho_\alpha^\circ(\varsigma_i^n(\alpha)) \mathbb{1}_{\{|\varsigma_i^n(\alpha)|>v_n\}} \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n| > \Delta_n^{\overline{v}}\}}.$$

Lemma A.4 and Lemma A.7 show $\mathbb{P}(J_n^{(2)}(\alpha) \cap Q_n) \to 1$ for each $\alpha > 0$ small enough and consequently it suffices to verify

$$\hat{C}_n^\circ(\alpha) = o_{\mathbb{P}}(1), \tag{6.51}$$
$$\hat{D}_n^\circ(\alpha) = o_{\mathbb{P}}(1), \tag{6.52}$$
$$\hat{E}_n^\circ(\alpha) = o_{\mathbb{P}}(1), \tag{6.53}$$
$$\hat{F}_n^\circ(\alpha) = o_{\mathbb{P}}(1), \tag{6.54}$$

for all $\alpha > 0$ as $n \to \infty$.

Concerning (6.51), we have due to the triangle inequality and $|\rho_\alpha^\circ(z)| \le K|z|^p$ for all $z \in \mathbb{R}$

$$\hat{C}_n^\circ(\alpha) \le \frac{K}{\sqrt{n\Delta_n}} \sum_{i=1}^n |\xi_i| \Big(|\varsigma_i^n(\alpha)|^p \mathbb{1}_{\{|\varsigma_i^n(\alpha)|>v_n\}} + |\Delta_i^n \tilde{X}''(8\alpha)^n|^p \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n|>v_n\}} \Big) \times \\ \times \mathbb{1}_{\{|\Delta_i^n \tilde{X}'^n| \le v_n/2\}} \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n| \le \Delta_n^{\overline{v}}\}} \mathbb{1}_{Q_n}$$

for fixed $\alpha > 0$ on the set Q_n and where $\varsigma_i^n(\alpha) = \Delta_i^n \tilde{X}'^n + \Delta_i^n \tilde{X}''(8\alpha)^n$. Consequently, because of $\mathbb{E}|\xi_i| \le 1$ for each $i \in \mathbb{N}$ and the fact that the multipliers are defined on a distinct probability space we obtain

$$\mathbb{E}\big(\hat{C}_n^\circ(\alpha) \mathbb{1}_{Q_n}\big) \le K\big(y_n^{(\alpha)} + 2z_n^{(\alpha)}\big),$$

with

$$y_n^{(\alpha)} = \frac{1}{\sqrt{n\Delta_n}} \sum_{i=1}^n \mathbb{E}\Big\{\big|\big|\Delta_i^n \tilde{X}'^n + \Delta_i^n \tilde{X}''(8\alpha)^n\big|^p - \big|\Delta_i^n \tilde{X}''(8\alpha)^n\big|^p\big| \times \\ \times \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n| \leq \Delta_n^{\overline{v}}\}} \mathbb{1}_{\{|\Delta_i^n \tilde{X}'^n + \Delta_i^n \tilde{X}''(8\alpha)^n| > v_n\}} \mathbb{1}_{\{|\Delta_i^n \tilde{X}'^n| \leq v_n/2\}} \mathbb{1}_{Q_n}\Big\},$$

and

$$z_n^{(\alpha)} = \frac{1}{\sqrt{n\Delta_n}} \sum_{i=1}^n \mathbb{E}\Big\{\big|\Delta_i^n \tilde{X}''(8\alpha)^n\big|^p \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n| \leq \Delta_n^{\overline{v}}\}} \mathbb{1}_{Q_n}\Big\}.$$

Thus, (6.51) follows using $\mathbb{P}(Q_n) \to 1$ and Lemma A.19.
Recall the quantity $D_n^\circ(\alpha)$ introduced in (6.33). Because of (6.38), Lemma A.18 and the fact that the multipliers $(\xi_i)_{i \in \mathbb{N}}$ are independent of the other involved quantities and satisfy $\mathbb{E}|\xi_i| \leq 1$ we obtain $0 \leq \mathbb{E}\hat{D}_n^\circ(\alpha) \leq \mathbb{E}D_n^\circ(\alpha) \to 0$, which proves (6.52).
In order to show (6.53) we have for $\alpha > 0$ and $n \in \mathbb{N}$ large enough

$$\mathbb{E}|\bar{a}_i^n(\alpha)|^m \leq \Big(\frac{K}{\sqrt{n\Delta_n}}\Big)^m \Delta_n,$$

for some $K > 0$, all $m \in \mathbb{N}$ and all $i = 1, \ldots, n$ using Lemma A.15. Thus, with Assumption 3.6 as well as independence of ξ_i and $\bar{a}_i^n(\alpha)$ we obtain for every integer $m \geq 2$ and $n \in \mathbb{N}$ large enough

$$\mathbb{E}|\bar{Z}_i^n(\alpha)|^m \leq m!\Big(\frac{C_1}{\sqrt{n\Delta_n}}\Big)^{m-2}\frac{C_2}{n},$$

for some constants $C_1, C_2 > 0$, where

$$\bar{Z}_i^n(\alpha) = \xi_i \bar{a}_i^n(\alpha)$$

for $n \in \mathbb{N}$, $\alpha > 0$ and $i = 1, \ldots, n$. Furthermore, due to the definition of $\tilde{X}''(8\alpha)^n$ in (6.23) the variables $(\bar{Z}_i^n(\alpha))_{i=1,\ldots,n}$ are independent with mean zero. Consequently, Lemma A.16 shows

$$\mathbb{E}\hat{E}_n^\circ(\alpha) = \mathbb{E}\Big\{\sup_{A \in \mathfrak{S}_n}\Big|\sum_{i \in A} \bar{Z}_i^n(\alpha)\Big|\Big\} = o(1),$$

which proves (6.53).
Concerning (6.54) notice first of all that we have $\bar{a}_i^n(\alpha) = \bar{b}_i^n(\alpha) = 0$ on the set $\{|\Delta_i^n \tilde{X}''(8\alpha)^n| \leq v_n/2\} \cap Q_n$, because of the indicators in the definition of these terms and the fact that $|\Delta_i^n \tilde{X}'^n| \leq v_n/2$ holds for each $i = 1, \ldots, n$ on Q_n according to (6.24). Therefore, we have for arbitrary $i \in \{1, \ldots, n\}$

$$|\bar{a}_i^n(\alpha) - \bar{b}_i^n(\alpha)|\mathbb{1}_{Q_n} = |\bar{a}_i^n(\alpha) - \bar{b}_i^n(\alpha)|\mathbb{1}_{Q_n}\mathbb{1}_{\{v_n/2 < |\Delta_i^n \tilde{X}''(8\alpha)^n| \leq 2v_n\}} + \\ + |\bar{a}_i^n(\alpha) - \bar{b}_i^n(\alpha)|\mathbb{1}_{Q_n}\mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n| > 2v_n\}}. \quad (6.55)$$

For the first summand on the right-hand side of (6.55) the triangle inequality gives

$$|\bar{a}_i^n(\alpha) - \bar{b}_i^n(\alpha)|\mathbb{1}_{Q_n}\mathbb{1}_{\{v_n/2 < |\Delta_i^n \tilde{X}''(8\alpha)^n| \leq 2v_n\}} \leq \\ \frac{1}{\sqrt{n\Delta_n}}(|\rho_\alpha^\circ(\Delta_i^n \tilde{X}''(8\alpha)^n)| + |\rho_\alpha^\circ(\Delta_i^n \tilde{X}'^n + \Delta_i^n \tilde{X}''(8\alpha)^n)|)\mathbb{1}_{Q_n}\mathbb{1}_{\{v_n/2 < |\Delta_i^n \tilde{X}''(8\alpha)^n| \leq 2v_n\}}.$$

Due to Assumption 6.12(a2) we have $|\rho(z)| \leq K|z|^p$ for all $z \in \mathbb{R}$ and some $K > 0$. Thus, because $|\Delta_i^n \tilde{X}'^n| \leq |\Delta_i^n \tilde{X}''(8\alpha)^n|$ holds on $\{v_n/2 < |\Delta_i^n \tilde{X}''(8\alpha)^n|\} \cap Q_n$ we further obtain

$$\begin{aligned}
&|\bar{a}_i^n(\alpha) - \bar{b}_i^n(\alpha)|\mathbb{1}_{Q_n}\mathbb{1}_{\{v_n/2<|\Delta_i^n\tilde{X}''(8\alpha)^n|\leq 2v_n\}} \\
&\qquad \leq \frac{1}{\sqrt{n\Delta_n}}\big(K|\Delta_i^n\tilde{X}''(8\alpha)^n|^p + K2^p|\Delta_i^n\tilde{X}''(8\alpha)^n|^p\big)\mathbb{1}_{Q_n}\mathbb{1}_{\{v_n/2<|\Delta_i^n\tilde{X}''(8\alpha)^n|\leq 2v_n\}} \\
&\qquad \leq \frac{K}{\sqrt{n\Delta_n}}|\Delta_i^n\tilde{X}''(8\alpha)^n|^p\mathbb{1}_{Q_n}\mathbb{1}_{\{|\Delta_i^n\tilde{X}''(8\alpha)^n|\leq 2v_n\}} \\
&\qquad \leq \frac{Kv_n}{\sqrt{n\Delta_n}}\big|\Delta_i^n\tilde{X}''(8\alpha)^n\big|^{p-1}. \qquad (6.56)
\end{aligned}$$

The fact that $|\Delta_i^n \tilde{X}'^n| \leq v_n/2$ for all $i = 1, \ldots, n$ on Q_n yields for the second summand in (6.55)

$$\begin{aligned}
&|\bar{a}_i^n(\alpha) - \bar{b}_i^n(\alpha)|\mathbb{1}_{Q_n}\mathbb{1}_{\{|\Delta_i^n\tilde{X}''(8\alpha)^n|>2v_n\}} \\
&= \frac{1}{\sqrt{n\Delta_n}}\big|\rho_\alpha^\circ(\Delta_i^n\tilde{X}''(8\alpha)^n) - \rho_\alpha^\circ(\Delta_i^n\tilde{X}'^n + \Delta_i^n\tilde{X}''(8\alpha)^n)\big|\mathbb{1}_{Q_n}\mathbb{1}_{\{|\Delta_i^n\tilde{X}''(8\alpha)^n|>2v_n\vee\Delta_n^{\overline{v}}\}}.
\end{aligned}$$

The derivative of the function Ψ_α° from (6.3) is supported by a compact set which is bounded away from the origin. Therefore, by Assumption 6.12(a2) there exists a constant $K > 0$, which may depend on α, such that the derivative satisfies $|\frac{d}{dz}\rho_\alpha^\circ(z)| \leq K|z|^{p-1}$. As a consequence, we have due to the mean value theorem and $|\Delta_i^n \tilde{X}'^n| \leq |\Delta_i^n \tilde{X}''(8\alpha)^n|$ on $\{|\Delta_i^n \tilde{X}''(8\alpha)^n| > 2v_n\} \cap Q_n$

$$\begin{aligned}
&|\bar{a}_i^n(\alpha) - \bar{b}_i^n(\alpha)|\mathbb{1}_{Q_n}\mathbb{1}_{\{|\Delta_i^n\tilde{X}''(8\alpha)^n|>2v_n\}} \\
&\qquad \leq \frac{K}{\sqrt{n\Delta_n}}|\Delta_i^n\tilde{X}'^n||\Delta_i^n\tilde{X}''(8\alpha)^n|^{p-1}\mathbb{1}_{Q_n}\mathbb{1}_{\{|\Delta_i^n\tilde{X}''(8\alpha)^n|>2v_n\vee\Delta_n^{\overline{v}}\}} \\
&\qquad \leq \frac{Kv_n}{\sqrt{n\Delta_n}}\big|\Delta_i^n\tilde{X}''(8\alpha)^n\big|^{p-1}. \qquad (6.57)
\end{aligned}$$

Finally we conclude with (6.55), (6.56) and (6.57)

$$\begin{aligned}
\mathbb{E}\big(|\hat{E}_n^\circ(\alpha) - \hat{F}_n^\circ(\alpha)|\mathbb{1}_{Q_n}\big) &\leq \mathbb{E}\big(\mathbb{1}_{Q_n}\sup_{A\in\mathfrak{S}_n}\sum_{i\in A}|\xi_i||\bar{a}_i^n(\alpha) - \bar{b}_i^n(\alpha)|\big) \\
&\leq \mathbb{E}\big(\mathbb{1}_{Q_n}\sum_{i=1}^n|\xi_i||\bar{a}_i^n(\alpha) - \bar{b}_i^n(\alpha)|\big) \\
&\leq \frac{Kv_n}{\sqrt{n\Delta_n}}\sum_{i=1}^n\mathbb{E}\big|\Delta_i^n\tilde{X}''(8\alpha)^n\big|^{p-1} \to 0, \qquad (6.58)
\end{aligned}$$

because the multipliers are defined on a distinct probability space and satisfy $\mathbb{E}|\xi_i| \leq 1$ for all $i \in \mathbb{N}$. The final convergence in the display above holds due to Lemma A.18. (6.58) together with $\mathbb{P}(Q_n) \to 1$ (see Lemma A.4) shows $\hat{F}_n^\circ(\alpha) \xrightarrow{\mathbb{P}} 0$, because in the previous part of the proof we have already verified $\hat{E}_n^\circ(\alpha) \xrightarrow{\mathbb{P}} 0$. □

Proof of Theorem 6.8. According to Definition 3.4 we have to show

$$\sup_{h\in \mathrm{BL}_1(\ell^\infty([0,1]\times\mathbb{R}))}\big|\mathbb{E}_\xi h(\hat{G}_\rho^{(n)}) - \mathbb{E}h(\mathbb{G}_\rho)\big| \xrightarrow{\mathbb{P}^*} 0, \qquad (6.59)$$

and

$$\mathbb{E}_\xi h(\hat{G}_\rho^{(n)})^* - \mathbb{E}_\xi h(\hat{G}_\rho^{(n)})_* \xrightarrow{\mathbb{P}} 0 \text{ for all } h \in \mathrm{BL}_1(\ell^\infty([0,1]\times\mathbb{R})), \tag{6.60}$$

where $h(\hat{G}_\rho^{(n)})^*$ and $h(\hat{G}_\rho^{(n)})_*$ denote a minimal measurable majorant and a maximal measurable minorant with respect to the joint data, respectively.
In order to show (6.59) observe that by the properties of bounded Lipschitz functions we have for each $h \in \mathrm{BL}_1(\ell^\infty([0,1]\times\mathbb{R}))$

$$\begin{aligned}\big|\mathbb{E}_\xi h(\hat{G}_\rho^{(n)}) - \mathbb{E}h(\mathbb{G}_\rho)\big| \le \big|\mathbb{E}_\xi h(\hat{G}_{\rho,n}^{(\alpha)}) - \mathbb{E}h(\mathbb{G}_{\rho_\alpha})\big| + \\ + \big|\mathbb{E}h(\mathbb{G}_{\rho_\alpha}) - \mathbb{E}h(\mathbb{G}_\rho)\big| + |\mathbb{E}_\xi Y_n^{(\alpha)} - \mathbb{E}\mathbb{Y}^{(\alpha)}| + \mathbb{E}\mathbb{Y}^{(\alpha)},\end{aligned}$$

for every $\alpha > 0$, where

$$Y_n^{(\alpha)} = \sup_{(\theta,t)\in[0,1]\times\mathbb{R}} |\hat{G}_{\rho,n}^{\circ(\alpha)}(\theta,t)| \wedge 2 \quad \text{and} \quad \mathbb{Y}^{(\alpha)} = \sup_{(\theta,t)\in[0,1]\times\mathbb{R}} |\mathbb{G}_{\rho_\alpha^\circ}(\theta,t)| \wedge 2.$$

Thus, due to Lemma 1.2.2(i) in Van der Vaart and Wellner (1996) we obtain

$$\Big(\sup_{h\in\mathrm{BL}_1(\ell^\infty([0,1]\times\mathbb{R}))} \big|\mathbb{E}_\xi h(\hat{G}_\rho^{(n)}) - \mathbb{E}h(\mathbb{G}_\rho)\big|\Big)^* \le q_n^{(\alpha)} + p(\alpha) + |\mathbb{E}_\xi Y_n^{(\alpha)} - \mathbb{E}\mathbb{Y}^{(\alpha)}| + \mathbb{E}\mathbb{Y}^{(\alpha)}, \tag{6.61}$$

for each $\alpha > 0$, where

$$\begin{aligned}q_n^{(\alpha)} &= \Big(\sup_{h\in\mathrm{BL}_1(\ell^\infty([0,1]\times\mathbb{R}))} \big|\mathbb{E}_\xi h(\hat{G}_{\rho,n}^{(\alpha)}) - \mathbb{E}h(\mathbb{G}_{\rho_\alpha})\big|\Big)^* \\ p(\alpha) &= \sup_{h\in\mathrm{BL}_1(\ell^\infty([0,1]\times\mathbb{R}))} \big|\mathbb{E}h(\mathbb{G}_{\rho_\alpha}) - \mathbb{E}h(\mathbb{G}_\rho)\big|.\end{aligned}$$

Notice that the supremum in the definition of $Y_n^{(\alpha)}$ is measurable, because the process $\hat{G}_{\rho,n}^{\circ(\alpha)}$ depends only via $\lfloor n\theta \rfloor$ on $\theta \in [0,1]$ and is right-continuous in $t \in \mathbb{R}$. Let $\varepsilon > 0$ be arbitrary. Then due to Proposition B.4 and monotonicity of the integral we obtain

$$\mathbb{E}\mathbb{Y}^{(\alpha)} \le \varepsilon/4, \tag{6.62}$$

for all $\alpha > 0$ in a neighbourhood of 0. Moreover, because of Lemma 6.5 and Theorem 1.12.1 in Van der Vaart and Wellner (1996) we have

$$p(\alpha) \le \varepsilon/4, \tag{6.63}$$

for $\alpha > 0$ small enough. Thus, choose an $\alpha > 0$ such that (6.62), (6.63) and Lemma 6.10 hold. Then Lemma 6.9 yields $q_n^{(\alpha)} \xrightarrow{\mathbb{P}} 0$ as $n \to \infty$ and due to Lemma 6.10 we have $|\mathbb{E}_\xi Y_n^{(\alpha)} - \mathbb{E}\mathbb{Y}^{(\alpha)}| \xrightarrow{\mathbb{P}} 0$ as $n \to \infty$, because $Y_n^{(\alpha)} = h_0(\hat{G}_{\rho,n}^{\circ(\alpha)})$ and $\mathbb{Y}^{(\alpha)} = h_0(\mathbb{G}_{\rho_\alpha^\circ})$ for the bounded Lipschitz function $h_0 : \ell^\infty([0,1]\times\mathbb{R}) \to \mathbb{R}$ given by $h_0(f) = \sup_{(\theta,t)\in[0,1]\times\mathbb{R}} |f(\theta,t)| \wedge 2$. As a consequence, we obtain (6.59) with (6.61):

$$\begin{aligned}\mathbb{P}\Big(\Big(\sup_{h\in\mathrm{BL}_1(\ell^\infty([0,1]\times\mathbb{R}))} \big|\mathbb{E}_\xi h(\hat{G}_\rho^{(n)}) - \mathbb{E}h(\mathbb{G}_\rho)\big|\Big)^* > \varepsilon\Big) \\ \le \mathbb{P}(q_n^{(\alpha)} > \varepsilon/4) + \mathbb{P}(|\mathbb{E}_\xi Y_n^{(\alpha)} - \mathbb{E}\mathbb{Y}^{(\alpha)}| > \varepsilon/4) \to 0.\end{aligned}$$

In order to show (6.60) we have for each $h \in \mathrm{BL}_1(\ell^\infty([0,1]\times\mathbb{R}))$ and each $\alpha > 0$

$$h(\hat{G}_{\rho,n}^{(\alpha)}) - Y_n^{(\alpha)} \le h(\hat{G}_\rho^{(n)}) \le h(\hat{G}_{\rho,n}^{(\alpha)}) + Y_n^{(\alpha)}.$$

Therefore, applying Lemma 1.2.2(i) in Van der Vaart and Wellner (1996) and the relation $-Z_* = (-Z)^*$ between the minimal measurable majorant and the maximal measurable minorant of a random element Z several times yields

$$\begin{aligned}|\mathbb{E}_\xi h(\hat{G}_\rho^{(n)})^* - \mathbb{E}_\xi h(\hat{G}_\rho^{(n)})_*| &= \mathbb{E}_\xi h(\hat{G}_\rho^{(n)})^* - \mathbb{E}_\xi h(\hat{G}_\rho^{(n)})_* \\ &\le |\mathbb{E}_\xi h(\hat{G}_{\rho,n}^{(\alpha)})^* - \mathbb{E}_\xi h(\hat{G}_{\rho,n}^{(\alpha)})_*| + 2\mathbb{E}_\xi Y_n^{(\alpha)} \\ &\le |\mathbb{E}_\xi h(\hat{G}_{\rho,n}^{(\alpha)})^* - \mathbb{E}_\xi h(\hat{G}_{\rho,n}^{(\alpha)})_*| + 2|\mathbb{E}_\xi Y_n^{(\alpha)} - \mathbb{E}\mathbb{Y}^{(\alpha)}| + 2\mathbb{E}\mathbb{Y}^{(\alpha)},\end{aligned} \tag{6.64}$$

for every $h \in \mathrm{BL}_1(\ell^\infty([0,1]\times\mathbb{R}))$ and each $\alpha > 0$. For arbitrary $\varepsilon > 0$ we choose $\alpha > 0$ such that Lemma 6.10 holds and $\mathbb{E}\mathbb{Y}^{(\alpha)} \le \varepsilon/8$. Then as above we see $|\mathbb{E}_\xi Y_n^{(\alpha)} - \mathbb{E}\mathbb{Y}^{(\alpha)}| \xrightarrow{\mathbb{P}} 0$ for $n \to \infty$ and furthermore we have $|\mathbb{E}_\xi h(\hat{G}_{\rho,n}^{(\alpha)})^* - \mathbb{E}_\xi h(\hat{G}_{\rho,n}^{(\alpha)})_*| \xrightarrow{\mathbb{P}} 0$ by Lemma 6.9 as $n \to \infty$. These facts together with (6.64) give (6.60):

$$\begin{aligned}&\mathbb{P}(|\mathbb{E}_\xi h(\hat{G}_\rho^{(n)})^* - \mathbb{E}_\xi h(\hat{G}_\rho^{(n)})_*| > \varepsilon) \le \\ &\qquad \mathbb{P}(|\mathbb{E}_\xi h(\hat{G}_{\rho,n}^{(\alpha)})^* - \mathbb{E}_\xi h(\hat{G}_{\rho,n}^{(\alpha)})_*| > \varepsilon/2) + \mathbb{P}(|\mathbb{E}_\xi Y_n^{(\alpha)} - \mathbb{E}\mathbb{Y}^{(\alpha)}| > \varepsilon/8) \to 0.\end{aligned}$$

□

6.6 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 $\mathbf{H}_1^{(loc)}$

$$\begin{aligned}\mathbb{T}_\rho^{(n)}(\theta,t) &= h_n\big(G_\rho^{(n)}\big)(\theta,t) + \sqrt{n\Delta_n}\big(N_\rho(g^{(n)};\theta,t) - \frac{\lfloor n\theta\rfloor}{n}N_\rho(g^{(n)};1,t)\big) \\ &= h_n\big(G_\rho^{(n)}\big)(\theta,t) + \sqrt{n\Delta_n}\Big(\theta - \frac{\lfloor n\theta\rfloor}{n}\Big)\int_{-\infty}^t \rho(z)\nu_0(dz) + \\ &\quad + \big(N_\rho(g_1;\theta,t) - \frac{\lfloor n\theta\rfloor}{n}N_\rho(g_1;1,t)\big) + \sqrt{n\Delta_n}\big(N_\rho(\mathcal{R}_n;\theta,t) - \frac{\lfloor n\theta\rfloor}{n}N_\rho(\mathcal{R}_n;1,t)\big),\end{aligned}$$

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

$$h_n(f)(\theta,t) = f(\theta,t) - \frac{\lfloor n\theta\rfloor}{n}f(1,t), \quad (n \in \mathbb{N}), \quad h_0(f)(\theta,t) = f(\theta,t) - \theta f(1,t). \tag{6.65}$$

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. (6.12) in the proof of Proposition 6.13 shows that Assumption 6.12(a3) is also valid for $2p$ instead of p. Thus, Theorem 6.1 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 (6.65). Then we have $\hat{\mathbb{T}}_\rho^{(n)} = h_n(\hat{G}_\rho^{(n)})$ and Proposition 10.7 in Kosorok (2008) together with Theorem 6.8 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 C.1 yields $\hat{\mathbb{T}}_\rho^{(n)} \rightsquigarrow_\xi h_0(\mathbb{G}_\rho)$. The covariance structure (3.7) of $h_0(\mathbb{G}_\rho) = \mathbb{T}_\rho$ can be obtained using (6.1). □

Proof of Proposition 3.10. First, we show (3.11) 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 C.2 and the continuous mapping theorem we have $(T_\rho^{(n)}, \hat{T}_{\rho,\xi^{(1)}}^{(n)}, \ldots, \hat{T}_{\rho,\xi^{(B)}}^{(n)}) \rightsquigarrow (T_{\rho,g_1}, T_{\rho,(1)}, \ldots, T_{\rho,(B)})$ in $(\mathbb{R}^{B+1}, \mathbb{B}^{B+1})$ for fixed $B \in \mathbb{N}$, where $T_{\rho,(1)}, \ldots, T_{\rho,(B)}$ 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 $\hat{T}_{\rho,\xi^{(1)}}^{(n)}, \ldots, \hat{T}_{\rho,\xi^{(B)}}^{(n)}$ and let L_B be the empirical c.d.f. calculated from $T_{\rho,(1)}, \ldots, T_{\rho,(B)}$. Due to the right continuity of $L_{n,B}$ we have

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

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

$\Psi_{(B)}(x_0, x_1, \ldots, x_B) = B^{-1}\sum_{i=1}^{B} \mathbb{1}(x_i \le x_0)$ is almost surely continuous with respect to the image measure $(T_{\rho,g_1}, T_{\rho,(1)}, \ldots, T_{\rho,(B)})(\mathbb{P})$. 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

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

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

$$\mathbb{P}\big(\sup_{x\in\mathbb{R}} |L_B(x) - L_\rho(x)| \ge \varepsilon\big) \le \varepsilon, \tag{6.66}$$

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

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

and we obtain

$$\limsup_{B\to\infty} \mathbb{P}\big(L_B(T_{\rho,g_1}) \ge 1-\alpha\big) \le \mathbb{P}\big(L_\rho(T_{\rho,g_1}) \ge 1-\alpha\big). \tag{6.67}$$

The terms on both sides of inequality (6.67) are increasing in α and the right-hand side is right continuous in α. As a consequence, (6.67) is also valid for each $\alpha \in (0,1) \cap \mathbb{Q}$. Furthermore, we have

$$\liminf_{B\to\infty} \mathbb{P}\big(L_B(T_{\rho,g_1}) \ge 1-\alpha\big) \ge \mathbb{P}\big(L_\rho(T_{\rho,g_1}) > 1-\alpha\big), \tag{6.68}$$

because according to (6.66)

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

holds. Both sides of (6.68) are increasing in α and the right-hand side is left continuous in α. Thus, (6.68) is also true for $\alpha \in (0,1) \cap \mathbb{Q}$. Finally, (3.13) can be shown by exactly the same steps as above and (3.12) 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 $\nu_0 \ne 0$, 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 $\mathbb{P}\big(L_\rho(T_\rho) > 1-\alpha\big) = \mathbb{P}\big(L_\rho(T_\rho) \ge 1-\alpha\big) = \alpha$ for all $\alpha \in (0,1)$. Hence, (3.14) follows from (3.11) and the claim (3.15) can be obtained by a similar reasoning using (3.12) as well as (3.13). □

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. □

6.7 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,

$$\int_{-\infty}^t \rho(z)g_0(y,dz) = A_\theta(t) \tag{6.69}$$

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 (6.69) are right-continuous in $t \in \mathbb{R}$. As a consequence, (6.69) 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

$$\Lambda(f)(\zeta,\theta,t) := f(\zeta,t)-\frac{\zeta}{\theta}f(\theta,t). \tag{6.70}$$

As $\|\Lambda(f_1)-\Lambda(f_2)\|_{C\times\mathbb{R}}\le 2\|f_1-f_2\|_{[0,1]\times\mathbb{R}}$ the mapping Λ is Lipschitz continuous. Thus, due to Theorem 6.1 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 $\mathbb{H}_\rho:=\Lambda(\mathbb{G}_\rho)$, where a simple calculation shows that $\mathbb{H}_\rho$ has the 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 (6.70). Thus, the assertion follows from Proposition 10.7 in Kosorok (2008). □

Proof of Theorem 4.7. The assertion follows with the same reasoning as in the proof of Theorem 4.2 in Hoffmann et al. (2017). □

Proof of Theorem 4.8. Basically the same steps as in the proof of Theorem 4.3 in Hoffmann et al. (2017) yield the claim. □

Proof of Theorem 4.9. The claim of Theorem 4.9 follows by the same reasoning as in the proof of Theorem 4.4 in Hoffmann et al. (2017). □

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

$$\mathbb{P}\big(\hat{\varkappa}_{n,B_n}^{(\alpha_n,\rho)}(r)/\sqrt{n\Delta_n}\le x\big) = \mathbb{P}\Big(\frac{1}{B_n}\sum_{i=1}^{B_n}\mathbb{1}_{\{\hat{\mathbb{H}}_{\rho,*}^{(n,i)}(\hat\theta_n^*)\le(\sqrt{n\Delta_n}x)^{1/r}\}}\ge 1-\alpha_n\Big)\to 1, \tag{6.71}$$

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

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

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

$$\mathbb{P}\Big(\Big|\frac{1}{B_n}\sum_{i=1}^{B_n}\mathbb{1}_{\{\hat{\mathbb{H}}_{\rho,*}^{(n,i)}(\hat\theta_n^*)\le(\sqrt{n\Delta_n}x)^{1/r}\}} - \mathbb{P}_\xi\Big(\hat{\mathbb{H}}_{\rho,*}^{(n)}(\hat\theta_n^*)\le(\sqrt{n\Delta_n}x)^{1/r}\Big)\Big| > \alpha_n/2\Big)\le 4\alpha_n^{-2}B_n^{-1}\to 0. \tag{6.72}$$

Therefore, in order to prove (6.71), it suffices to verify

$$\mathbb{P}\Big(\mathbb{P}_\xi\Big(\hat{\mathbb{H}}_{\rho,*}^{(n)}(\hat{\theta}_n^*) \le (\sqrt{n\Delta_n}x)^{1/r}\Big) < 1-\alpha_n/2\Big) \le \frac{2}{\alpha_n}\mathbb{P}\Big(\hat{\mathbb{H}}_{\rho,*}^{(n)}(\hat{\theta}_n^*) > (\sqrt{n\Delta_n}x)^{1/r}\Big)$$
$$\le 2\alpha_n^{-1}\mathbb{P}(Q_n^C) + 2\alpha_n^{-1}\mathbb{P}\Big(\Big\{2\sup_{t\in\mathbb{R}}\sup_{\theta\in[0,1]}|\hat{G}_\rho^{(n)}(\theta,t)| > (\sqrt{n\Delta_n}x)^{1/r}\Big\}\cap Q_n\Big) \to 0, \quad (6.73)$$

with Q_n the set defined in (6.24). The first inequality in the above display follows with the Markov inequality and the last inequality in (6.73) 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 A.5 we obtain $\mathbb{P}(Q_n^C) \le Kn\Delta_n^{1+\tau}$ and consequently $\alpha_n^{-1}\mathbb{P}(Q_n^C) \to 0$. For the second summand on the right-hand side of (6.73) the definition of $\hat{G}_\rho^{(n)}$ in (6.7) gives

$$\mathbb{E}\Big\{\sup_{t\in\mathbb{R}}\sup_{\theta\in[0,1]}|\hat{G}_\rho^{(n)}(\theta,t)|\mathbb{1}_{Q_n}\Big\} \le \frac{1}{\sqrt{n\Delta_n}}\sum_{i=1}^n \mathbb{E}\big(|\xi_i||\rho(\Delta_i^n X^{(n)})|\mathbb{1}_{\{|\Delta_i^n X^{(n)}|>v_n\}}\mathbb{1}_{Q_n}\big) \le K\sqrt{n\Delta_n}.$$

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

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

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

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

for each $x > 0$. With the same considerations as for (6.72) it is sufficient to show

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

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

$$N_{\rho^2}(\bar{\zeta},t_0) > 0 \quad (6.74)$$

and because of

$$\hat{\mathbb{H}}_\rho^{(n)}(\bar{\zeta},\bar{\theta},t_0) \le \hat{\mathbb{H}}_{\rho,*}^{(n)}(\hat{\theta}_n^*) \Longrightarrow \mathbb{P}_\xi\Big(\hat{\mathbb{H}}_\rho^{(n)}(\bar{\zeta},\bar{\theta},t_0) > x^{1/r}\Big) \le \mathbb{P}_\xi\Big(\hat{\mathbb{H}}_{\rho,*}^{(n)}(\hat{\theta}_n^*) > x^{1/r}\Big)$$

on the set $\{\bar{\theta} < \hat{\theta}_n^*\}$ and consistency of the preliminary estimate it further suffices to prove

$$\mathbb{P}\Big(\mathbb{P}_\xi\Big(\hat{\mathbb{H}}_{\rho,*}^{(n)}(\hat{\theta}_n^*) > x^{1/r}\Big) \le 2\alpha_n \text{ and } \bar{\theta} < \hat{\theta}_n^*\Big) \le \mathbb{P}\Big(\mathbb{P}_\xi\Big(\hat{\mathbb{H}}_\rho^{(n)}(\bar{\zeta},\bar{\theta},t_0) > x^{1/r}\Big) \le 2\alpha_n\Big) \to 0. \quad (6.75)$$

In order to show (6.75) we want to use a Berry-Esseen type result. Recall

$$\hat{\mathbb{H}}_\rho^{(n)}(\bar{\zeta},\bar{\theta},t_0) = \frac{1}{\sqrt{n\Delta_n}}\sum_{j=1}^n \hat{B}_j^n \xi_j$$

from (4.12) with $\hat{B}_j^n = \Big(\mathbb{1}_{\{j\leq \lfloor n\bar{\zeta}\rfloor\}} - \frac{\bar{\zeta}}{\bar{\theta}}\mathbb{1}_{\{j\leq\lfloor n\bar{\theta}\rfloor\}}\Big)\hat{A}_j^n$, where

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

By the assumptions on the multiplier sequence it is immediate 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

$$\begin{aligned}\sup_{x\in\mathbb{R}}\Big|\mathbb{P}_\xi\Big(\hat{\mathbb{H}}_\rho^{(n)}(\bar{\zeta},\bar{\theta},t_0)>x\Big) - (1-\Phi(x/\hat{W}_n))\Big| \\ \leq K\Big\{\sum_{i=1}^n \mathbb{E}_\xi \hat{U}_{i,n}^2\mathbb{1}_{\{|\hat{U}_{i,n}|>1\}} + \sum_{i=1}^n\mathbb{E}_\xi|\hat{U}_{i,n}|^3\mathbb{1}_{\{|\hat{U}_{i,n}|\leq 1\}}\Big\},\end{aligned} \tag{6.76}$$

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 (6.75), we first show

$$\frac{1}{\hat{W}_n^2} = \frac{n\Delta_n}{\sum\limits_{j=1}^n(\hat{B}_j^n)^2} = O_{\mathbb{P}}(1), \tag{6.77}$$

that is

$$\lim_{M\to\infty}\limsup_{n\to\infty}\mathbb{P}\Big(n\Delta_n > M\sum_{j=1}^n(\hat{B}_j^n)^2\Big) = 0.$$

Let $M>0$. Then a straightforward calculation gives

$$\begin{aligned}\mathbb{P}\Big(n\Delta_n > M\sum_{j=1}^n(\hat{B}_j^n)^2\Big) &\leq \mathbb{P}\Big(n\Delta_n > M'\sum_{j=1}^{\lfloor n\bar{\zeta}\rfloor}(\hat{A}_j^n)^2\Big) \\ &= \mathbb{P}\Big(1/M' > N_{\rho^2}^{(n)}(\bar{\zeta},t_0)\Big),\end{aligned}$$

with $M' = M(1-\bar{\zeta}/\bar{\theta})^2$. Consequently, with (6.74) we obtain (6.77) due to

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

because Theorem 6.1 also holds for ρ^2 since Assumption 6.12 is also valid for $2p$ instead of p (cf. (6.12) in the proof of Proposition 6.13). Recall that our main objective is to show (6.75) and thus we consider the Berry-Esseen bound on the right-hand side of (6.76). 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

$$|\hat{U}_{i,n}| \leq \frac{\sqrt{M}K}{\sqrt{n\Delta_n}}$$

for all $i = 1, \dots, n$ on the set $\{1/\hat{W}_n^2 \leq M\}$, since $|\hat{B}_i^n|$ is bounded. As a consequence,

$$\sum_{i=1}^n \mathbb{E}_\xi \hat{U}_{i,n}^2 \mathbb{1}_{\{|\hat{U}_{i,n}|>1\}} = 0 \tag{6.78}$$

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

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

$$\mathbb{E}_\xi \xi^2 \mathbb{1}_{\{|\xi|>y\}} = \frac{2}{\sqrt{2\pi}} \int_y^\infty z^2 e^{-z^2/2} dz \leq K\mathbb{P}(\mathcal{N}(0,2) > y) \leq K_1 \exp(-K_2 y^2). \tag{6.79}$$

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 (6.79). 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

$$\sum_{i=1}^n \mathbb{E}_\xi \hat{U}_{i,n}^2 \mathbb{1}_{\{|\hat{U}_{i,n}|>1\}} \leq K_1 \Delta_n^{-1} (n\Delta_n)^{-\bar{q}} = K_1 \big(n\Delta_n^{1+\bar{\delta}}\big)^{-\bar{q}}. \tag{6.80}$$

We now consider the second term on the right-hand side of (6.76), for which

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

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

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

Consequently,

$$\mathbb{P}\Big(\Big\{1/\hat{W}_n^2 \le M \text{ and } K\sum_{i=1}^n \mathbb{E}_\xi |\hat{U}_{i,n}|^3 \mathbb{1}_{\{|\hat{U}_{i,n}|\le 1\}} > (n\Delta_n)^{-1/4}\Big\} \cap Q_n\Big)$$
$$\le \mathbb{P}\Big(\Big\{\frac{K}{(n\Delta_n)^{3/2}}\sum_{i=1}^n |\hat{B}_i^n| > (n\Delta_n)^{-1/4}\Big\} \cap Q_n\Big) \le K(n\Delta_n)^{-1/4} \quad (6.81)$$

follows. Thus, from (6.78), (6.80) and (6.81) we see that with $K > 0$ from (6.76) for each $M > 0$ there exists a $K_3 > 0$ such that

$$\mathbb{P}\Big(1/\hat{W}_n^2 \le M \text{ and } K\Big\{\sum_{i=1}^n \mathbb{E}_\xi \hat{U}_{i,n}^2 \mathbb{1}_{\{|\hat{U}_{i,n}|>1\}} + \sum_{i=1}^n \mathbb{E}_\xi |\hat{U}_{i,n}|^3 \mathbb{1}_{\{|\hat{U}_{i,n}|\le 1\}}\Big\}$$
$$> K_3((n\Delta_n)^{-1/4} + (n\Delta_n^{1+\bar{\delta}})^{-\bar{q}})\Big) \to 0. \quad (6.82)$$

Now we can show (6.75). Let $\eta > 0$ and according to (6.77) 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 (6.82) 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 (6.76) 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 we are only left with proving (4.17). Let

$$K = \Big((1+\varepsilon)/c\Big)^{1/\varpi} > \Big(1/c\Big)^{1/\varpi}$$

for some $\varepsilon > 0$. Then

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

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 $\varphi_n^* \xrightarrow{\mathbb{P}} 0$

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

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

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

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

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

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

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

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

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

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

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

Consequently, there exists a $K > 0$ such that $|\rho'_{L,\hat{p}}(z)| \le K|z|^{\hat{p}-1}$ 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 $\rho'_{L,\hat{p}}(z) \ge 0$ 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

$$\Upsilon_{L,\hat{p}}(a,\beta,p,t) = La \times \begin{cases} \frac{2+\hat{p}}{p-1}|t|^{1-p}, & \text{for } t \leq -2 \\ \frac{2+\hat{p}}{p-1}2^{1-p} + 4 - \frac{2\hat{p}}{3} - 2\hat{p}t^2 - \frac{\hat{p}}{3}t^3 + (2-3\hat{p})t, & \text{for } -2 \leq t \leq -1 \\ \frac{2+\hat{p}}{p-1}2^{1-p} + 2 + \frac{2\hat{p}}{3} + \frac{2}{\hat{p}-\beta}(1 + \operatorname{sign}(t)|t|^{\hat{p}-\beta}), & \text{for } -1 \leq t \leq 1 \\ \frac{2+\hat{p}}{p-1}2^{1-p} + 2\hat{p} + \frac{4}{\hat{p}-\beta} + 2\hat{p}t^2 - \frac{\hat{p}}{3}t^3 + 2t - 3\hat{p}t, & \text{for } 1 \leq t \leq 2 \\ \frac{4+2\hat{p}}{p-1}2^{1-p} + \frac{4}{\hat{p}-\beta} + \frac{4\hat{p}}{3} + 4 + \frac{2+\hat{p}}{1-p}t^{1-p}, & \text{for } t \geq 2. \end{cases} \tag{6.83}$$

Furthermore, it is well known from complex analysis that there is a domain $U \subset U^* \subset \mathbb{C}$ with holomorphic functions $A^* : U^* \to \mathbb{C}$, $\beta^* : U^* \to \mathbb{C}^{\hat{p}-} := \{u \in \mathbb{C} \mid \operatorname{Re}(u) < \hat{p}\}$ and $p^* : U^* \to \mathbb{C}^{1+} := \{u \in \mathbb{C} \mid \operatorname{Re}(u) > 1\}$ 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 (6.83) 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

$$\bar{N}(y,t) = \sum_{k=0}^{\infty} \frac{N_k(t)}{k!}(y-\theta_0)^k, \tag{6.84}$$

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 $K > 0$

$$\Psi_{L,\hat{p}}(y) + K\frac{\bar{A}(y)}{1-\bar{p}(y)}t^{1-\bar{p}(y)} = N_0(t) \tag{6.85}$$

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

$$\Psi_{L,\hat{p}}(y) = L\bar{A}(y)\Big(\frac{4+2\hat{p}}{\bar{p}(y)-1}2^{1-\bar{p}(y)} + \frac{4}{\hat{p}-\bar{\beta}(y)} + \frac{4\hat{p}}{3} + 4\Big). \tag{6.86}$$

Taking the derivative with respect to $y \in U$ on both sides of (6.85) yields

$$\Psi'_{L,\hat{p}}(y) + K\frac{\bar{A}'(y)(1-\bar{p}(y)) + \bar{A}(y)\bar{p}'(y)}{(1-\bar{p}(y))^2}t^{1-\bar{p}(y)} - \bar{p}'(y)\frac{K\bar{A}(y)}{1-\bar{p}(y)}\log(t)t^{1-\bar{p}(y)} = 0, \tag{6.87}$$

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 (6.87) 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 (6.86) 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

$$\begin{aligned} \sup_{t \in \mathbb{R}} |(f_1 f_2)^{(\ell)}(\theta_0, t)| &= \sup_{t \in \mathbb{R}} \Big| \sum_{j=0}^{\ell} \binom{\ell}{j} f_1^{(j)}(\theta_0, t) f_2^{(\ell-j)}(\theta_0, t) \Big| \\ &\leq K^2 \sum_{j=0}^{\ell} \binom{\ell}{j} (Kj)^j (K(\ell-j))^{\ell-j} \\ &\leq K^2 (K\ell)^\ell \sum_{j=0}^{\ell} \binom{\ell}{j} \leq K(K\ell)^\ell, \end{aligned} \tag{6.88}$$

for some $K > 0$ which does not depend on ℓ. (6.83) and (6.88) yield a constant $K > 0$ such that

$$\sup_{t \in \mathbb{R}} |N_\ell(t)| \leq K(K\ell)^\ell \tag{6.89}$$

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

$$|\bar{A}^{(\ell)}(\theta_0)| \leq K(K\ell)^\ell, \tag{6.90}$$

$$\Big| \Big(\frac{1}{\bar{p}(y) - 1} \Big)^{(\ell)} (\theta_0) \Big| \leq K(K\ell)^\ell, \tag{6.91}$$

$$\Big| \Big(\frac{1}{\hat{p} - \bar{\beta}(y)} \Big)^{(\ell)} (\theta_0) \Big| \leq K(K\ell)^\ell, \tag{6.92}$$

$$\sup_{t \geq 2} \Big| \Big(t^{1 - \bar{p}(y)} \Big)^{(\ell)} (\theta_0) \Big| \leq K(K\ell)^\ell, \tag{6.93}$$

$$\sup_{t \in [0,1]} \Big| \Big(t^{\hat{p} - \bar{\beta}(y)} \Big)^{(\ell)} (\theta_0) \Big| \leq K(K\ell)^\ell. \tag{6.94}$$

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 $|A_\ell| \leq K^{\ell+1}$ for each $\ell \in \mathbb{N}_0$. Thus, because of $\bar{A}^{(\ell)}(\theta_0) = \ell! A_\ell$ for each $\ell \in \mathbb{N}_0$, (6.90) follows. By assumption in Example 4.3 we have $\bar{\beta}(y) \leq \hat{\beta} \leq 1 \vee \hat{\beta} < \hat{p} < \bar{p}(y)$ for each $y \in U$. As a consequence, the functions

$$y \mapsto \frac{1}{\bar{p}(y) - 1} \quad \text{and} \quad y \mapsto \frac{1}{\hat{p} - \bar{\beta}(y)}$$

are real analytic on U as concatenations of real analytic functions. So the same reasoning as above yields (6.91) and (6.92). 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

$$\left(t^{1-\bar{p}(y)}\right)^{(\ell)}(\theta_0) = t^{1-p_0}(-p_1 \log(t))^\ell.$$

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 $t \geq 2$ 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 (6.93)

$$\sup_{t \geq 2} \left| \left(t^{1-p(y)}\right)^{(\ell)}(\theta_0) \right| \leq |p_1|^\ell \max\left\{2^{1-p_0} \log(2)^\ell, \left(\frac{\ell}{p_0 - 1}\right)^\ell e^{-\ell}\right\} \leq K(K\ell)^\ell$$

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

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

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 (6.94)

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

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 (6.84) and (6.89) 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. □

A Technical details in the proofs of Theorem 6.1 and Theorem 6.8

In this appendix we give the details of the proofs of Theorem 6.1 and Theorem 6.8. Here and also in the appendices B and C K or $K(\alpha)$ denote generic constants which sometimes depend on a further quantity α and may change from place to place.

A.1 Moments of functionals of integer-valued random measures

Hoffmann and Vetter (2017) used Lemma 2.1.5 and Lemma 2.1.7 of Jacod and Protter (2012) frequently in order to achieve their weak convergence results. However, in Jacod and Protter (2012) these results are only proved for Poisson random measures with a predictable compensator of the form $ds \otimes F(dz)$ with a Lévy measure F. Therefore, using tools from Jacod (1979) we prove the generalized versions stated below. First, we introduce some notations. Let $\mathfrak{p}$ be an integer-valued random measure on $\mathbb{R}_+ \times \mathbb{R}$ with predictable compensator $\mathfrak{q}(\omega; ds, dz) = \nu_s(\omega; dz)ds$ for a transition kernel $\nu_s(\omega; dz)$ from $(\Omega \times \mathbb{R}_+, \mathcal{P})$ into $(\mathbb{R}, \mathbb{B})$, where $\mathcal{P}$ is the predictable σ-algebra on $\Omega \times \mathbb{R}_+$ (with respect to some prespecified filtration $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$) and where an integer-valued random measure is a random measure which satisfies the requirements of Definition II.1.3 and Definition II.1.13 in Jacod and Shiryaev (2002). Furthermore, we set $\Omega' = \Omega \times \mathbb{R}_+ \times \mathbb{R}$ and $\mathcal{P}' = \mathcal{P} \otimes \mathbb{B}$ is the predictable σ-algebra on Ω'. Then, for a real-valued $\mathcal{P}'$-measurable function δ on Ω' and $p, t \in \mathbb{R}_+$, $u > 0$ let

$$\hat{\delta}(p)_{t,u}(\omega) = \frac{1}{u} \int_t^{t+u} \int |\delta(\omega, s, z)|^p \nu_s(\omega; dz) ds.$$

Lemma A.1. *Suppose that $\hat{\delta}(2)_{0,u} < \infty$ almost surely for all $u > 0$. Then the process $Y = \delta \star (\mathfrak{p} - \mathfrak{q})$ is a locally square integrable martingale, and for all finite stopping times T and $u > 0$ we have for $p \in [1, 2]$*

$$\mathbb{E}\Big(\sup_{0 \le v \le u} |Y_{T+v} - Y_T|^p \mid \mathcal{F}_T \Big) \le K_p u \mathbb{E}\big(\hat{\delta}(p)_{T,u} \mid \mathcal{F}_T\big),$$

and also for $p \geq 2$

$$\mathbb{E}\Big(\sup_{0 \le v \le u} |Y_{T+v} - Y_T|^p \mid \mathcal{F}_T \Big) \le K_p \big(u \mathbb{E}\big(\hat{\delta}(p)_{T,u} \mid \mathcal{F}_T\big) + u^{p/2} \mathbb{E}\big(\hat{\delta}(2)_{T,u}^{p/2} \mid \mathcal{F}_T\big)\big).$$

Lemma A.2. *Suppose that $\hat{\delta}(1)_{0,u} < \infty$ almost surely for all $u > 0$. Then the process $Y = \delta \star \mathfrak{p}$ is of locally integrable variation. Furthermore, for all finite stopping times T and $u > 0$ we have for $p \in (0, 1]$*

$$\mathbb{E}\Big(\sup_{0 \le v \le u} |Y_{T+v} - Y_T|^p \mid \mathcal{F}_T \Big) \le K_p u \mathbb{E}\big(\hat{\delta}(p)_{T,u} \mid \mathcal{F}_T\big)$$

and for $p \geq 1$

$$\mathbb{E}\Big(\sup_{0 \le v \le u} |Y_{T+v} - Y_T|^p \mid \mathcal{F}_T \Big) \le K_p \big(u \mathbb{E}\big(\hat{\delta}(p)_{T,u} \mid \mathcal{F}_T\big) + u^{p} \mathbb{E}\big(\hat{\delta}(1)_{T,u}^{p} \mid \mathcal{F}_T\big)\big).$$

Proof of Lemma A.1. $\delta^2 \star \mathfrak{q}$ is a continuous increasing process and we have $\delta^2 \star \mathfrak{q}_t = t\hat{\delta}(2)_{0,t}$ for all $t > 0$. Thus, for $n \in \mathbb{N}$ let M_n be a null set such that $\delta^2 \star \mathfrak{q}_n$ is finite on M_n^C. Such a set exists by the assumption on δ. Then the increasing process $\delta^2 \star \mathfrak{q}_t$ is finite for all $t \in \mathbb{R}_+$ on M^C with $M = \bigcup_{n \in \mathbb{N}} M_n$. Therefore, $(T_n)_{n \in \mathbb{N}}$ defined via $T_n = \inf\{t > 0 \mid \delta^2 \star \mathfrak{q}_t \geq n\}$ is a localizing sequence of stopping times and the stopped continuous processes satisfy $(\delta^2 \star \mathfrak{q})_t^{T_n} \leq n$

for all $t \in \mathbb{R}_+$. Consequently, $\delta^2 \star \mathfrak{q}$ is locally bounded and in particular locally integrable. Thus, by Theorem II.1.33(a) in Jacod and Shiryaev (2002) the process Y is well-defined and a locally square integrable martingale.

In order to show the claimed inequalities we want to reduce our setup to the situation of Lemma 2.1.5 in Jacod and Protter (2012). To this end, let F be a Lévy measure on $(\mathbb{R}, \mathbb{B})$ without atoms and $F(\mathbb{R}) = \infty$. Furthermore, let $x_0 \notin \mathbb{R}$ be an exterior point of $\mathbb{R}$ and let $(\mathbb{R}_{x_0}, \mathbb{B}_{x_0})$ denote the measurable one point extension of $(\mathbb{R}, \mathbb{B})$, that is $\mathbb{R}_{x_0} = \mathbb{R} \cup \{x_0\}$ and $\mathbb{B}_{x_0} = \{B, B \cup \{x_0\} \mid B \in \mathbb{B}\}$. Then according to Theorem 14.53 in Jacod (1979) there exist a measurable function $h\colon (\Omega', \mathcal{P}') \to (\mathbb{R}_{x_0}, \mathbb{B}_{x_0})$ and a $\mathbb{P}$-null set N such that

$$\mathfrak{q}(\omega; A) = \int \int \mathbb{1}_A(s, h(\omega, s, z)) F(dz) ds \tag{A.1}$$

for each $A \in \mathbb{B}(\mathbb{R}_+) \otimes \mathbb{B}$ and $\omega \notin N$. Additionally, by Theorem 14.56 in Jacod (1979) there exists a filtered measurable space $(\Omega^\circ, \mathcal{F}^\circ, (\mathcal{F}_t^\circ)_{t \in \mathbb{R}_+})$ and a transition probability $Q(\omega, d\omega^\circ)$ from $(\Omega, \mathcal{F})$ into $(\Omega^\circ, \mathcal{F}^\circ)$ such that on the extended filtered probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_t)_{t \in \mathbb{R}_+}, \tilde{\mathbb{P}})$, which is given by $\tilde{\Omega} = \Omega \times \Omega^\circ$, $\tilde{\mathcal{F}} = \mathcal{F} \otimes \mathcal{F}^\circ$, $\tilde{\mathcal{F}}_t = \bigcap_{s > t} \mathcal{F}_s \otimes \mathcal{F}_s^\circ$ and $\tilde{\mathbb{P}}(d(\omega, \omega^\circ)) = Q(\omega, d\omega^\circ)\mathbb{P}(d\omega)$, there exists a Poisson random measure $\tilde{\mathfrak{p}}$ with predictable compensator $\tilde{\mathfrak{q}}(ds, dz) = F(dz)ds$ such that for $\tilde{\mathbb{P}}$-almost every $\tilde{\omega} = (\omega, \omega^\circ)$ we have

$$\mathfrak{p}(\omega; A) = \int \int \mathbb{1}_A(s, h(\omega, s, z)) \tilde{\mathfrak{p}}((\omega, \omega^\circ), ds, dz) \tag{A.2}$$

for all $A \in \mathbb{B}(\mathbb{R}_+) \otimes \mathbb{B}$. Furthermore, we identify the filtration $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$ on $(\Omega, \mathcal{F})$ with the induced filtration $\mathcal{F}_t \otimes \{\emptyset, \Omega^\circ\}$ on $(\tilde{\Omega}, \tilde{\mathcal{F}})$, which we denote by $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$ as well. Any random variable X on $(\Omega, \mathcal{F})$ will be identified with the induced mapping $X(\omega, \omega^\circ) = X(\omega)$. Then we have for every $A \in \tilde{\mathcal{F}}$ and every stopping time T on $(\Omega, \mathcal{F})$

$$\begin{aligned} A \in \mathcal{F}_T(\Omega) \otimes \{\emptyset, \Omega^\circ\} &\iff A = A_1 \times \Omega^\circ \text{ for some } A_1 \in \mathcal{F}_T(\Omega) \\ &\iff A \cap \{T \le t\} \in \mathcal{F}_t \otimes \{\emptyset, \Omega^\circ\} \text{ for every } t \in \mathbb{R}_+ \\ &\iff A \in \mathcal{F}_T(\tilde{\Omega}), \end{aligned}$$

where for the sake of a clear notation we denote by $\mathcal{F}_T(\Omega)$ and $\mathcal{F}_T(\tilde{\Omega})$, respectively, the σ-algebra of events up to time T with respect to $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}_+})$ and $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\mathcal{F}_t)_{t \in \mathbb{R}_+})$, respectively. Consequently, for $A = A_1 \times \Omega^\circ \in \mathcal{F}_T(\tilde{\Omega})$ with $A_1 \in \mathcal{F}_T(\Omega)$ and a random variable X on $(\Omega, \mathcal{F})$ we have by the definition of the conditional expectation

$$\begin{aligned} \int\limits_{A_1 \times \Omega^\circ} X d\tilde{\mathbb{P}} &= \int\limits_{A_1} \int\limits_{\Omega^\circ} X(\omega) Q(\omega, d\omega^\circ) \mathbb{P}(d\omega) \\ &= \int\limits_{A_1} X d\mathbb{P} = \int\limits_{A_1} \mathbb{E}_{\mathbb{P}}(X \mid \mathcal{F}_T) d\mathbb{P} = \int\limits_{A_1 \times \Omega^\circ} \mathbb{E}_{\mathbb{P}}(X \mid \mathcal{F}_T) d\tilde{\mathbb{P}} \end{aligned}$$

and thus

$$\mathbb{E}_{\tilde{\mathbb{P}}}(X \mid \mathcal{F}_T) = \mathbb{E}_{\mathbb{P}}(X \mid \mathcal{F}_T) \quad \tilde{\mathbb{P}} - \text{almost surely}. \tag{A.3}$$

Let $\mathcal{O}$ be the optional σ-algebra on $\Omega \times \mathbb{R}_+$ and let $\tilde{\mathcal{O}}$ denote the optional σ-algebra on $\tilde{\Omega} \times \mathbb{R}_+$. Then by Proposition II.1.14 there exist a thin random set $D \in \mathcal{O}$, an optional process $(\beta_s)_{s\in\mathbb{R}_+}$on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P})$, a thin random set $\tilde{D} \in \tilde{\mathcal{O}}$ and an optional process $(\tilde{\beta}_s)_{s\in\mathbb{R}_+}$ on $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_t)_{t\in\mathbb{R}_+}, \tilde{\mathbb{P}})$ such that

$$\begin{aligned}\mathfrak{p}(\omega; ds, dz) &= \sum_{t\geq 0} \mathbb{1}_D(\omega, t)\epsilon_{(t,\beta_t(\omega))}(ds, dz)\\ \tilde{\mathfrak{p}}((\omega, \omega^\circ); ds, dz) &= \sum_{t\geq 0} \mathbb{1}_{\tilde{D}}((\omega, \omega^\circ), t)\epsilon_{(t,\tilde{\beta}_t(\omega,\omega^\circ))}(ds, dz),\end{aligned}$$

for every $(\omega, \omega^\circ) \in \tilde{\Omega}$, where $\epsilon_{(x,y)}$ is the Dirac measure on $\mathbb{R}_+ \times \mathbb{R}$ with mass in (x, y). As a consequence, we obtain from (A.2)

$$\begin{aligned}\delta(\omega, t, \beta_t(\omega))\mathbb{1}_D(\omega, t) &= \int\int \delta(\omega, s, z)\mathbb{1}_{\{s=t\}}\mathfrak{p}(\omega; ds, dz)\\ &= \int\int \delta(\omega, s, h(\omega, s, z))\mathbb{1}_{\{s=t\}}\tilde{\mathfrak{p}}((\omega, \omega^\circ); ds, dz)\\ &= \delta(\omega, t, h(\omega, t, \tilde{\beta}_t(\omega, \omega^\circ)))\mathbb{1}_{\tilde{D}}((\omega, \omega^\circ), t),\end{aligned}$$

for every $t \geq 0$ and $\tilde{\mathbb{P}}$-almost every (ω, ω°), where we set $f(\omega, s, h(\omega, s, z)) = 0$ if $h(\omega, s, z) = x_0$ for a real-valued predictable function f on Ω'. Thus, the processes $\delta(\omega, t, \beta_t(\omega))\mathbb{1}_D(\omega, t)$ and $\delta(\omega, t, h(\omega, t, \tilde{\beta}_t(\omega, \omega^\circ)))\mathbb{1}_{\tilde{D}}((\omega, \omega^\circ), t)$ are $\tilde{\mathbb{P}}$-indistinguishable on $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_t)_{t\in\mathbb{R}_+}, \tilde{\mathbb{P}})$ and the stochastic integrals $\delta \star (\mathfrak{p} - \mathfrak{q})$ and $(\delta \circ h) \star (\tilde{\mathfrak{p}} - \tilde{\mathfrak{q}})$ are $\tilde{\mathbb{P}}$-indistinguishable as well (cf. Definition II.1.27 in Jacod and Shiryaev (2002)), where for the sake of brevity $(\delta \circ h)$ denotes the predictable map $(\omega, s, z) \mapsto \delta(\omega, s, h(\omega, s, z))$ on Ω'. Notice that $\delta^2 \star \mathfrak{q} = (\delta^2 \circ h) \star \tilde{\mathfrak{q}}$ outside a null set due to (A.1). Thus, the same reasoning as at the beginning of the proof shows that $\tilde{Y}_t := (\delta \circ h) \star (\tilde{\mathfrak{p}} - \tilde{\mathfrak{q}})_t$ is well-defined and a locally square integrable martingale. Finally, for every finite stopping time T and all $u > 0$, $p \geq 1$ the variables $\sup_{0\leq v\leq u} |Y_{T+v} - Y_T|^p$ and $\sup_{0\leq v\leq u} |\tilde{Y}_{T+v} - \tilde{Y}_T|^p$ coincide $\tilde{\mathbb{P}}$-almost surely. Consequently, using (A.3) we obtain

$$\begin{aligned}\mathbb{E}_{\mathbb{P}}\Big(\sup_{0\leq v\leq u} |Y_{T+v} - Y_T|^p \mid \mathcal{F}_T\Big) &= \mathbb{E}_{\tilde{\mathbb{P}}}\Big(\sup_{0\leq v\leq u} |Y_{T+v} - Y_T|^p \mid \mathcal{F}_T\Big)\\ &= \mathbb{E}_{\tilde{\mathbb{P}}}\Big(\sup_{0\leq v\leq u} |\tilde{Y}_{T+v} - \tilde{Y}_T|^p \mid \mathcal{F}_T\Big). \qquad \text{(A.4)}\end{aligned}$$

Now, Lemma 2.1.5 in Jacod and Protter (2012), (A.1) and (A.3) give for $p \in [1, 2]$

$$\begin{aligned}\mathbb{E}_{\mathbb{P}}\Big(\sup_{0\leq v\leq u} |Y_{T+v} - Y_T|^p \mid \mathcal{F}_T\Big) &\leq K_p u \mathbb{E}_{\tilde{\mathbb{P}}}\Big(\frac{1}{u} \int_T^{T+u} \int |\delta(\omega, s, h(\omega, s, z))|^p F(dz)ds \mid \mathcal{F}_T\Big)\\ &= K_p u \mathbb{E}_{\mathbb{P}}\Big(\frac{1}{u} \int_T^{T+u} \int |\delta(\omega, s, h(\omega, s, z))|^p F(dz)ds \mid \mathcal{F}_T\Big)\\ &= K_p u \mathbb{E}\big(\hat{\delta}(p)_{T,u} \mid \mathcal{F}_T\big). \qquad \text{(A.5)}\end{aligned}$$

The second asserted inequality follows with exactly the same reasoning. □

Proof of Lemma A.2. In the same way as at the beginning of the proof of Lemma A.1 we see that the increasing, continuous and finite-valued process $|\delta| \star \mathfrak{q}$ is locally bounded. Hence, by the definition of the predictable compensator (Theorem II.1.8 in Jacod and Shiryaev (2002)) the process $|\delta| \star \mathfrak{p}$ is locally integrable and thus Y is of locally integrable variation.

With the same quantities as in the proof of Lemma A.1 we obtain from (A.2)

$$\begin{aligned} Y_t &= \int \int \mathbb{1}_{[0,t]}(s)\delta(\omega, s, z)\mathfrak{p}(\omega; ds, dz) \\ &= \int \int \mathbb{1}_{[0,t]}(s)\delta(\omega, s, h(\omega, s, z))\tilde{\mathfrak{p}}((\omega, \omega^\circ); ds, dz) \\ &= (\delta \circ h) \star \tilde{\mathfrak{p}}_t =: \tilde{Y}_t, \end{aligned}$$

for all $t \in \mathbb{R}_+$ $\tilde{\mathbb{P}}(d(\omega, \omega^\circ))$-almost surely. Thus, we have

$$\sup_{0 \le v \le u} |Y_{T+v} - Y_T|^p = \sup_{0 \le v \le u} |\tilde{Y}_{T+v} - \tilde{Y}_T|^p$$

$\tilde{\mathbb{P}}$-almost surely for all finite stopping times T and $p, u > 0$. Now, the same reasoning as in (A.4) and (A.5), but using Lemma 2.1.7 of Jacod and Protter (2012) instead, yields the desired inequalities. □

Remark A.3. In the proofs below integral processes of the form $Y_{(1)} = \delta \star \mu^{(n)}$ and $Y_{(2)} = \delta \star (\mu^{(n)} - \bar{\mu}^{(n)})$ occur frequently, where $\mu^{(n)}$ is the random measure associated with the jumps of the underlying process, $\bar{\mu}^{(n)}$ denotes its predictable compensator and δ is some suitable $\mathcal{P}'$-measurable function on Ω'. When we want to apply Lemma A.1 and Lemma A.2 to these processes the question is whether the condition $\hat{\delta}(2)_{0,u} < \infty$ almost surely or $\hat{\delta}(1)_{0,u} < \infty$ almost surely is satisfied for all $u > 0$, respectively. However, due to the observation scheme $\{X^{(n)}_{i\Delta_n} \mid i = 0, 1, \ldots, n\}$ only values of the processes $Y_{(1),t}$, $Y_{(2),t}$ for $t \le n\Delta_n$ are relevant and we can consider the stopped processes $Y^{T_n}_{(1)}$, $Y^{T_n}_{(2)}$ instead, where $T_n \equiv n\Delta_n$ is the constant stopping time. According to Definition II.1.27 and Proposition II.1.30 in Jacod and Shiryaev (2002) we have $Y^{T_n}_{(1)} = \delta \star \eta^{(n)}$ and $Y^{T_n}_{(2)} = \delta \star (\eta^{(n)} - \bar{\eta}^{(n)})$, where $\eta^{(n)}$ denotes the restriction of $\mu^{(n)}$ to the set $[0, n\Delta_n] \times \mathbb{R}$. Obviously, the predictable compensator $\bar{\eta}^{(n)}$ of $\eta^{(n)}$ is the restriction of $\bar{\mu}^{(n)}$ to $[0, n\Delta_n] \times \mathbb{R}$. As a consequence, we have $n\Delta_n\hat{\delta}(2)_{0,n\Delta_n} = n\Delta_n\hat{\delta}(2, \bar{\eta})_{0,n\Delta_n} = u\hat{\delta}(2, \bar{\eta})_{0,u}$ and $n\Delta_n\hat{\delta}(1)_{0,n\Delta_n} = n\Delta_n\hat{\delta}(1, \bar{\eta})_{0,n\Delta_n} = u\hat{\delta}(1, \bar{\eta})_{0,u}$ for all $u \ge n\Delta_n$, where $\hat{\delta}(2, \bar{\eta})$ and $\hat{\delta}(1, \bar{\eta})$ denote the function $\hat{\delta}(2)$ and $\hat{\delta}(1)$, respectively, calculated with respect to $\bar{\eta}$. Thus, Lemma A.1 and Lemma A.2 can be applied, if $T + u \le n\Delta_n$ and $\hat{\delta}(2)_{0,n\Delta_n} < \infty$ almost surely or $\hat{\delta}(1)_{0,n\Delta_n} < \infty$ almost surely, respectively. This is always satisfied when we apply these lemmas.

A.2 Results on the crucial decomposition

Recall the quantities defined in (6.23) which are used frequently in the proof of Theorem 6.1 and Theorem 6.8. With the constants from Assumption 6.12 let $\ell \in \mathbb{R}$ have the properties

$$1 < \ell < \frac{1}{2\beta\overline{w}} \wedge (1 + \epsilon) \quad \text{and also} \quad \ell < \frac{2(p-1)\overline{w} - 1}{2(\beta - 1)\overline{w}} \text{ if } \beta > 1, \tag{A.6}$$

with an $\epsilon > 0$ for which Assumption 6.12(b6) holds. Then we have

$$u_n = (v_n)^\ell \quad \text{and} \quad F_n = \{z \colon |z| > u_n\} \tag{A.7}$$

as well as

$$\begin{aligned}
\tilde{X}''^n &= (z\mathbb{1}_{F_n}(z)) \star \mu^{(n)}, \\
\tilde{X}''(\alpha)^n &= (z\mathbb{1}_{F_n \cap \{|z| \le \alpha/4\}}(z)) \star \mu^{(n)}, \quad \text{for } \alpha > 0 \\
\hat{X}''(\alpha)^n &= (z\mathbb{1}_{\{|z| > \alpha/4\}}) \star \mu^{(n)}, \quad \text{for } \alpha > 0 \\
N_t^n &= (\mathbb{1}_{F_n} \star \mu^{(n)})_t, \\
\tilde{X}_t'^n &= X_t^{(n)} - \tilde{X}_t''^n \\
&= X_0^{(n)} + \int_0^t b_s^{(n)} ds + \int_0^t \sigma_s^{(n)} dW_s^{(n)} + \\
&\qquad + (z\mathbb{1}_{F_n^C}(z)) \star (\mu^{(n)} - \bar{\mu}^{(n)})_t - (z\mathbb{1}_{\{|z| \le 1\} \cap F_n}(z)) \star \bar{\mu}_t^{(n)}, \\
A_i^n &= \{|\Delta_i^n \tilde{X}'^n| \le v_n/2\} \cap \{\Delta_i^n N^n \le 1\}.
\end{aligned} \tag{A.8}$$

The following lemma ensures that with high probability at most one large jump occurs and the remaining part is appropriately small.

Lemma A.4. *Let Assumption 6.12 be satisfied. Then for the sets*

$$Q_n = \bigcap_{i=1}^n A_i^n \tag{A.9}$$

we have $\mathbb{P}(Q_n) \to 1$ as $n \to \infty$.

Proof. Choose some $m' \in \mathbb{R}$ with

$$m' > \frac{2 + \beta\ell}{\ell - 1} \vee \frac{1 + 2\overline{w}}{1/2 - \overline{w}}.$$

Then by Lemma A.1 and Assumption 6.12(a1) we obtain for $1 \le i \le n$ and any sufficiently small $0 < \delta < 1$

$$\begin{aligned}
&\mathbb{E}\big|\Delta_i^n \big(z\mathbb{1}_{F_n^C}(z)\big) \star \big(\mu^{(n)} - \bar{\mu}^{(n)}\big)\big|^{m'} \\
&\le K\Bigg(\int_{(i-1)\Delta_n}^{i\Delta_n} \int_{\{|z| \le u_n\}} |z|^{m'} \nu_s^{(n)}(dz) ds + \Bigg\{ \int_{(i-1)\Delta_n}^{i\Delta_n} \int_{\{|z| \le u_n\}} |z|^2 \nu_s^{(n)}(dz) ds \Bigg\}^{m'/2} \Bigg) \\
&= K\Bigg(n\Delta_n \int_{(i-1)/n}^{i/n} \int_{\{|z| \le u_n\}} |z|^{m'} g^{(n)}(y, dz) dy + \Bigg\{ n\Delta_n \int_{(i-1)/n}^{i/n} \int_{\{|z| \le u_n\}} |z|^2 g^{(n)}(y, dz) dy \Bigg\}^{m'/2} \Bigg) \\
&\le K(\delta)\big(\Delta_n^{1 + (m' - \beta - \delta)\ell\overline{w}} + \Delta_n^{m'/2}\big).
\end{aligned}$$

Note that $m' > 2$ always so the lemma quoted above can be applied. Furthermore, $\bar{\mu}^{(n)}(ds, dz) = \nu_s^{(n)}(dz)ds$ yields for $1 \leq i \leq n$ and arbitrary $\delta > 0$ small enough

$$\begin{aligned}\left|\Delta_i^n\big(z\mathbb{1}_{\{|z|\leq 1\}\cap F_n}(z) \star \bar{\mu}^{(n)}\big)\right| &= \left|\int\limits_{(i-1)\Delta_n}^{i\Delta_n} \int\limits_{\{u_n<|z|\leq 1\}} z\nu_s^{(n)}(dz)ds\right| \\ &\leq u_n^{-(\beta+\delta-1)_+} n\Delta_n \int\limits_{(i-1)/n}^{i/n} \int\limits_{\{u_n<|z|\leq 1\}} |z|^{\beta+\delta} g^{(n)}(y, dz)dy \\ &\leq K(\delta)\Delta_n^{1-\ell\overline{w}(\beta+\delta-1)_+}.\end{aligned}$$

Let $m_b, m_\sigma \in \mathbb{R}$ be the constants in Assumption 6.12(c). Because of $m_b > 1$ and $m_\sigma > 2$ we can apply Hölder's inequality and the Burkholder-Davis-Gundy inequalities (see page 39 in Jacod and Protter (2012)) to obtain due to Assumption 6.12(c) for $1 \leq i \leq n$:

$$\begin{aligned}\mathbb{E}\left|\int\limits_{(i-1)\Delta_n}^{i\Delta_n} b_s^{(n)} ds\right|^{m_b} &\leq \Delta_n^{m_b}\mathbb{E}\bigg(\frac{1}{\Delta_n}\int\limits_{(i-1)\Delta_n}^{i\Delta_n} |b_s^{(n)}|^{m_b} ds\bigg) \\ &= \Delta_n^{m_b}\bigg(\frac{1}{\Delta_n}\int\limits_{(i-1)\Delta_n}^{i\Delta_n} \mathbb{E}|b_s^{(n)}|^{m_b} ds\bigg) \leq K\Delta_n^{m_b}\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}\left|\int\limits_{(i-1)\Delta_n}^{i\Delta_n} \sigma_s^{(n)} dW_s^{(n)}\right|^{m_\sigma} &\leq K\mathbb{E}\bigg(\int\limits_{(i-1)\Delta_n}^{i\Delta_n} |\sigma_s^{(n)}|^2 ds\bigg)^{m_\sigma/2} \\ &\leq K\Delta_n^{m_\sigma/2}\mathbb{E}\bigg(\frac{1}{\Delta_n}\int\limits_{(i-1)\Delta_n}^{i\Delta_n} |\sigma_s^{(n)}|^{m_\sigma} ds\bigg) \\ &= K\Delta_n^{m_\sigma/2}\bigg(\frac{1}{\Delta_n}\int\limits_{(i-1)\Delta_n}^{i\Delta_n} \mathbb{E}|\sigma_s^{(n)}|^{m_\sigma} ds\bigg) \leq K\Delta_n^{m_\sigma/2},\end{aligned}$$

where the equalities in the above displays hold according to Fubini's theorem. Additionally, according to Lemma A.13 we have for any $1 \leq i \leq n$ and some $K(\delta) > 0$

$$\mathbb{P}(\Delta_i^n N^n \geq 2) \leq K(\delta)\Delta_n^{2-2(\beta+\delta)\ell\overline{w}},$$

for $n \in \mathbb{N}$ large enough. Let us now choose $\delta > 0$ in such a way that $1 - \ell\overline{w}(\beta + \delta - 1)_+ > \overline{w}$. Then, for n large enough we have $\Delta_n^{1-\ell\overline{w}(\beta+\delta-1)_+} \leq Kv_n$, and the Markov inequality gives

$$\begin{aligned}\sum_{i=1}^n \mathbb{P}\big((A_i^n)^C\big) \leq K(\delta)n\big\{\Delta_n^{2-2(\beta+\delta)\ell\overline{w}} + \Delta_n^{1+(m'-\beta-\delta)\ell\overline{w}-m'\overline{w}} + \\ + \Delta_n^{m'/2-m'\overline{w}} + \Delta_n^{m_\sigma/2-m_\sigma\overline{w}} + \Delta_n^{m_b-m_b\overline{w}}\big\}. \quad \text{(A.10)}\end{aligned}$$

From the choice of the constants we further have

$$2 - 2(\beta+\delta)\ell\overline{w} \geq 2 - 2\beta(1+\epsilon)\overline{w} \tag{A.11}$$

and

$$\big(1 + (m' - \beta - \delta)\ell\overline{w} - m'\overline{w}\big) \wedge \big(m'/2 - m'\overline{w}\big) \wedge \big(m_\sigma/2 - m_\sigma\overline{w}\big) \wedge \big(m_b - m_b\overline{w}\big) \geq 1 + 2\overline{w}, \tag{A.12}$$

again for $\delta > 0$ small enough. Thus, the right hand side of (A.10) converges to zero for this choice of δ, using Assumption 6.12(b4) and (b6). □

If moreover Assumption 2.3 is valid, we can even give a rate for the convergence $\mathbb{P}(Q_n) \to 1$.

Lemma A.5. *If Assumption 2.3 is satisfied for some* $0 < \beta < 2$, $0 < \tau < (1/5 \wedge \frac{2-\beta}{2+5\beta})$ *and* $p > \beta + ((\frac{1}{2} + \frac{3}{2}\beta) \vee \frac{2}{1+5\tau})$, *we have*

$$\mathbb{P}\big(Q_n^C\big) \leq Kn\Delta_n^{1+\tau},$$

for some $K > 0$.

Proof. If Assumption 2.3 holds, then according to (6.11) in the proof of Proposition 6.13 Assumption 6.12 is valid for constants satisfying

$$1 + \tau = 2(1 - \beta\overline{w}(1+\epsilon)) < (1 + 2\overline{w}).$$

Comparing this fact with (A.10), (A.11) and (A.12) yields the assertion. □

In the next auxiliary lemma we consider for $\alpha > 0$, $1 \leq i, j \leq n$ with $i \neq j$ and the constant $\overline{r}$ in Assumption 6.12 the sets

$$R_{i,j}^{(n)}(\alpha) = \Big\{\big|\Delta_i^n \hat{X}''(\alpha)^n - \Delta_j^n \hat{X}''(\alpha)^n\big| \leq \Delta_n^{\overline{r}}\Big\} \cap \Big\{\big|\Delta_i^n \hat{X}''(\alpha)^n\big| > \alpha/4\Big\} \cap Q_n, \tag{A.13}$$

with the pure jump Itō semimartingale $\hat{X}''(\alpha)^n$ from (A.8). Furthermore, for $\alpha > 0$ let the sets $J_n^{(1)}(\alpha)$ be defined by their complements:

$$J_n^{(1)}(\alpha)^C = \bigcup_{\substack{i,j=1 \\ i\neq j}}^{n} R_{i,j}^{(n)}(\alpha). \tag{A.14}$$

Lemma A.6. *Grant Assumption 6.12. Then for each* $\alpha > 0$ *the sets* $J_n^{(1)}(\alpha)$ *defined in* (A.14) *satisfy*

$$\lim_{n\to\infty} \mathbb{P}\big(J_n^{(1)}(\alpha)\big) = 1.$$

Proof. Let x be arbitrary and either $z = 0$ or $|z| > \alpha/4$. Then, for n large enough we have

$$\mathbb{1}_{\{|x-z|\leq\Delta_n^{\overline{r}}\}}\mathbb{1}_{\{|x|>\alpha/4\}} \leq \mathbb{1}_{\{|x-z|\leq\Delta_n^{\overline{r}}\}}\mathbb{1}_{\{|x|>\alpha/4\}}\mathbb{1}_{\{|z|>\alpha/4\}}.$$

Furthermore, using the fact that for large $n \in \mathbb{N}$ on Q_n there is at most one jump of $\hat{X}''(\alpha)^n$ on an interval $((k-1)\Delta_n, k\Delta_n]$ with $1 \le k \le n$, we thus obtain

$$\mathbb{P}\big(R_{i,j}^{(n)}(\alpha)\big) \le \int\int\int\int\int \mathbb{1}_{\{|x-z|\le\Delta_n^{\overline{r}}\}}\mathbb{1}_{((j-1)\Delta_n,j\Delta_n]}(t)\mathbb{1}_{\{|z|>\alpha/4\}}\times \\ \times \mathbb{1}_{Q_n}(\omega)\mu^{(n)}(\omega;dt,dz)\mathbb{1}_{\{|x|>\alpha/4\}}\mathbb{1}_{((i-1)\Delta_n,i\Delta_n]}(s)\mu^{(n)}(\omega;ds,dx)\mathbb{P}(d\omega). \quad \text{(A.15)}$$

Now, forget about the indicator involving Q_n and assume $j < i$. If $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ denotes the underlying filtration, the inner integral in (A.15) with respect to $\mu^{(n)}(\omega;dt,dz)$ is an $(\mathcal{F}_{j\Delta_n} \otimes \mathbb{B})$-measurable function in (ω, x). Accordingly, the integrand in the integral with respect to $\mu^{(n)}(\omega;ds,dx)$ is in fact $\mathcal{P}'$-measurable. Therefore, Fubini's theorem and the definition of the predictable compensator of an optional $\mathcal{P}'$-σ-finite random measure (see Theorem II.1.8 in Jacod and Shiryaev (2002)) yield for n large enough:

$$\begin{aligned}\mathbb{P}\big(R_{i,j}^{(n)}(\alpha)\big) &\le \int_{(i-1)\Delta_n}^{i\Delta_n}\int_{(j-1)\Delta_n}^{j\Delta_n}\int\int \mathbb{1}_{\{|x-z|\le\Delta_n^{\overline{r}}\}}\mathbb{1}_{\{|x|>\alpha/4\}}\times \\ &\qquad\qquad \times \mathbb{1}_{\{|z|>\alpha/4\}}\nu_{s_1}^{(n)}(dz)\nu_{s_2}^{(n)}(dx)ds_1ds_2 \\ &\le n^2\Delta_n^2 \int_{(i-1)/n}^{i/n}\int_{(j-1)/n}^{j/n}\int\int \mathbb{1}_{\{|x-z|\le\Delta_n^{\overline{r}}\}}\mathbb{1}_{\{|x|>\alpha/4\}}\times \\ &\qquad\qquad \times \mathbb{1}_{\{|z|>\alpha/4\}}g^{(n)}(y_1,dz)g^{(n)}(y_2,dx)dy_1dy_2.\end{aligned} \quad \text{(A.16)}$$

Thus, we have $\mathbb{P}\big(J_n^{(1)}(\alpha)\big) \to 1$, because (A.16), Assumption 6.12(a(4)II) and Assumption 6.12(b3) show that there is a constant $K > 0$ such that

$$\mathbb{P}\big(J_n^{(1)}(\alpha)^C\big) \le Kn^2\Delta_n^{2+q} \to 0.$$

□

Similar to (A.13) for $\alpha > 0$, $1 \le i,j \le n$ with $i \ne j$ and the constants $\overline{v} < \overline{r}$ in Assumption 6.12 let

$$S_{i,j}^{(n)}(\alpha) = \Big\{\big|\Delta_i^n\tilde{X}''(8\alpha)^n - \Delta_j^n\tilde{X}''(8\alpha)^n\big| \le \Delta_n^{\overline{r}}\Big\} \cap \Big\{\big|\Delta_i^n\tilde{X}''(8\alpha)^n\big| > \Delta_n^{\overline{v}}\Big\} \cap Q_n.$$

and define the sets $J_n^{(2)}(\alpha)$ by

$$J_n^{(2)}(\alpha)^C = \bigcup_{\substack{i,j=1\\ i\ne j}}^{n} S_{i,j}^{(n)}(\alpha). \quad \text{(A.17)}$$

Lemma A.7. *Grant Assumption 6.12. Then for each $\alpha \in (0, \alpha_0/2)$, with α_0 the constant in Assumption 6.12(a(4)I), the sets $J_n^{(2)}(\alpha)$ defined in (A.17) satisfy*

$$\lim_{n\to\infty}\mathbb{P}\big(J_n^{(2)}(\alpha)\big) = 1.$$

Proof. The same considerations as for (A.15) and (A.16) yield

$$\begin{aligned}\mathbb{P}(S_{i,j}^{(n)}(\alpha)) &\le \int_{(i-1)\Delta_n}^{i\Delta_n}\int_{(j-1)\Delta_n}^{j\Delta_n}\int\int \mathbb{1}_{\{|x-z|\le\Delta_n^{\overline{r}}\}}\mathbb{1}_{\{\Delta_n^{\overline{v}}/2<|x|\le\alpha_0\}}\times \\ &\qquad\qquad\times \mathbb{1}_{\{\Delta_n^{\overline{v}}/2<|z|\le\alpha_0\}}\nu_{s_1}^{(n)}(dz)\nu_{s_2}^{(n)}(dx)ds_1ds_2 \\ &\le n^2\Delta_n^2\int_{(i-1)/n}^{i/n}\int_{(j-1)/n}^{j/n}\int\int \mathbb{1}_{\{|x-z|\le\Delta_n^{\overline{r}}\}}\mathbb{1}_{\{\Delta_n^{\overline{v}}/2<|x|\le\alpha_0\}}\times \\ &\qquad\qquad\times \mathbb{1}_{\{\Delta_n^{\overline{v}}/2<|z|\le\alpha_0\}}g^{(n)}(y_1,dz)g^{(n)}(y_2,dx)dy_1dy_2 \\ &\le K\Delta_n^{2+q},\end{aligned}$$

for n large enough, because of Assumption 6.12(a(4)I) and $\overline{v}<\overline{r}$. Thus, we obtain

$$\mathbb{P}\big(J_n^{(2)}(\alpha)^C\big)\le Kn^2\Delta_n^{2+q}\to 0,$$

by Assumption 6.12(b3) □

The next lemma yields bounds for the cardinality of the following random sets. For $\alpha>0$, $n\in\mathbb{N}$, $t\in\mathbb{R}$ and $\omega\in\Omega$ let

$$\begin{aligned}\tilde{A}_1(\omega;\alpha,n,t) &= \big\{i\in\{1,\dots,n\}\mid |\Delta_i^n\hat{X}''(\alpha)^n(\omega)|>\alpha/4 \text{ and} \\ &\qquad \mathbb{1}_{(-\infty,t]}(\Delta_i^n\tilde{X}'^n(\omega)+\Delta_i^n\hat{X}''(\alpha)^n(\omega))\ne\mathbb{1}_{(-\infty,t]}(\Delta_i^n\hat{X}''(\alpha)^n(\omega))\big\} \\ \tilde{A}_2(\omega;\alpha,n,t) &= \big\{i\in\{1,\dots,n\}\mid |\Delta_i^n\tilde{X}''(8\alpha)^n(\omega)|>\Delta_n^{\overline{v}} \text{ and} \\ &\qquad \mathbb{1}_{(-\infty,t]}(\Delta_i^n\tilde{X}'^n(\omega)+\Delta_i^n\tilde{X}''(8\alpha)^n(\omega))\ne\mathbb{1}_{(-\infty,t]}(\Delta_i^n\tilde{X}''(8\alpha)^n(\omega))\big\}\end{aligned}$$

Lemma A.8. *Grant Assumption 6.12 and let $c_n:=\lceil(v_n/\Delta_n^{\overline{r}})+1\rceil$. Then for all $\alpha>0$, $n\in\mathbb{N}$ and $t\in\mathbb{R}$ we have*

$$\#\tilde{A}_1(\omega;\alpha,n,t)\le c_n \tag{A.18}$$

for every $\omega\in J_n^{(1)}(\alpha)\cap Q_n$ as well as

$$\#\tilde{A}_2(\omega;\alpha,n,t)\le c_n \tag{A.19}$$

for all $\omega\in J_n^{(2)}(\alpha)\cap Q_n$, where $\#M$ denotes the cardinality of a set M.

Proof. For $\omega\in J_n^{(1)}(\alpha)\cap Q_n$ we have

$$\big|\Delta_i^n\hat{X}''(\alpha)^n-\Delta_j^n\hat{X}''(\alpha)^n\big|>\Delta_n^{\overline{r}}$$

for all $i,j\in\big\{k\in\{1,\dots,n\}\mid|\Delta_k^n\hat{X}''(\alpha)^n|>\alpha/4\big\}=:M_0(\omega;\alpha,n)$ with $i\ne j$ by the definition of the set $J_n^{(1)}(\alpha)$ in (A.14). Consequently, for fixed $t\in\mathbb{R}$

$$\Delta_i^n\hat{X}''(\alpha)^n\in[t-v_n/2,t+v_n/2]$$

can only hold for at most c_n indices $i\in M_0(\omega;\alpha,n)$. Thus, we conclude (A.18), because according to (A.9) we have $|\Delta_i^n\tilde{X}'^n|\le v_n/2$ for all $i\in\{1,\dots,n\}$ and $\omega\in Q_n$.

The assertion (A.19) follows with exactly the same reasoning. □

In the following we gather auxiliary lemmas which have a similar proof and give bounds for crucial quantities in the proof of Theorem 6.1 and Theorem 6.8. The first one is concerned with

$$V_\alpha^{(n)} := \frac{1}{\sqrt{n\Delta_n}} \sup_{(\theta,t)\in[0,1]\times\mathbb{R}} \Big| \sum_{i=1}^{\lfloor n\theta \rfloor} \Big\{ \chi_t^{(\alpha)}(\Delta_i^n X^{(n)}) \mathbb{1}_{\{|\Delta_i^n X^{(n)}|>v_n\}} - \chi_t^{(\alpha)}(\Delta_i^n L^{(n)}) \Big\} \Big|, \qquad \text{(A.20)}$$

from (6.26), where $\alpha > 0$, $\chi_t^{(\alpha)}$ is defined in (6.3) and $L^{(n)} = \big(z\mathbb{1}_{\{|z|>v_n\}}\big) \star \mu^{(n)}$ is the pure jump Itō semimartingale defined in (6.4).

Lemma A.9. *Let Assumption 6.12 be satisfied. Then for $\alpha > 0$, $\omega \in Q_n$ and $n \in \mathbb{N}$ large enough such that $v_n \leq \alpha/4$ we have*

$$V_\alpha^{(n)} \leq C_n(\alpha) + D_n(\alpha),$$

where

$$C_n(\alpha) = \frac{K}{\sqrt{n\Delta_n}} \sup_{t\in\mathbb{R}} \sum_{i=1}^{n} \big| \mathbb{1}_{(-\infty,t]}(\Delta_i^n \tilde{X}'^n + \Delta_i^n \hat{X}''(\alpha)^n) - \mathbb{1}_{(-\infty,t]}(\Delta_i^n \hat{X}''(\alpha)^n) \big| \times \\ \times \mathbb{1}_{\{|\Delta_i^n \hat{X}''(\alpha)^n|>\alpha/4\}} \mathbb{1}_{\{|\Delta_i^n \tilde{X}'^n| \leq v_n/2\}},$$

for $K > 0$ a bound for ρ and

$$D_n(\alpha) = \frac{1}{\sqrt{n\Delta_n}} \sum_{i=1}^{n} \big| \rho_\alpha(\Delta_i^n \tilde{X}'^n + \Delta_i^n \hat{X}''(\alpha)^n) \mathbb{1}_{\{|\Delta_i^n \tilde{X}'^n + \Delta_i^n \hat{X}''(\alpha)^n|>v_n\}} - \\ - \rho_\alpha(\Delta_i^n \hat{X}''(\alpha)^n) \mathbb{1}_{\{|\Delta_i^n \hat{X}''(\alpha)^n|>v_n\}} \big| \mathbb{1}_{\{|\Delta_i^n \tilde{X}'^n| \leq v_n/2\}},$$

with ρ_α defined prior to (6.3) and where the particular processes are defined in (A.8).

Proof. On Q_n, and with n large enough such that $v_n \leq \alpha/4$, one of the following mutually exclusive possibilities holds for $1 \leq i \leq n$:

(i) $\Delta_i^n N^n = 0$.
Then we have $|\Delta_i^n X^{(n)}| = |\Delta_i^n \tilde{X}'^n| \leq v_n/2$ and there is no jump larger than u_n (and v_n) on the interval $((i-1)\Delta_n, i\Delta_n]$. Thus, $\chi_t^{(\alpha)}(\Delta_i^n X^{(n)}) \mathbb{1}_{\{|\Delta_i^n X^{(n)}|>v_n\}} = 0 = \chi_t^{(\alpha)}(\Delta_i^n L^{(n)})$ holds for all $t \in \mathbb{R}$ and the corresponding summand in (A.20) vanishes.

(ii) $\Delta_i^n N^n = 1$ and $\Delta_i^n \tilde{X}''^n = \Delta_i^n \tilde{X}''(\alpha)^n \neq 0$.
So the only jump in $((i-1)\Delta_n, i\Delta_n]$ (of absolute size) larger than u_n is in fact not larger than $\alpha/4$, and because of $v_n \leq \alpha/4$ we have $|\Delta_i^n X^{(n)}| \leq \alpha/2$. Thus, as in the first case, $\chi_t^{(\alpha)}(\Delta_i^n X^{(n)}) \mathbb{1}_{\{|\Delta_i^n X^{(n)}|>v_n\}} = 0 = \chi_t^{(\alpha)}(\Delta_i^n L^{(n)})$ is true for all $t \in \mathbb{R}$ and the corresponding summand in (A.20) is equal to zero.

(iii) $\Delta_i^n N^n = 1$ and $\Delta_i^n \tilde{X}''^n \neq 0$, but $\Delta_i^n \tilde{X}''(\alpha)^n = 0$.
So the only jump in $((i-1)\Delta_n, i\Delta_n]$ larger than u_n is also larger than $\alpha/4$. If $\hat{X}''(\alpha)^n$ is the quantity defined in (A.8), we get

$$\Delta_i^n X^{(n)} = \Delta_i^n \tilde{X}'^n + \Delta_i^n \hat{X}''(\alpha)^n \quad \text{and}$$
$$\chi_t^{(\alpha)}(\Delta_i^n L^{(n)}) = \rho_\alpha(\Delta_i^n \hat{X}''(\alpha)^n) \mathbb{1}_{\{|\Delta_i^n \hat{X}''(\alpha)^n|>v_n\}} \mathbb{1}_{(-\infty,t]}(\Delta_i^n \hat{X}''(\alpha)^n).$$

Thus, we obtain an upper bound for $V_\alpha^{(n)}$ on Q_n, as soon as $v_n \le \alpha/4$:

$$
\begin{aligned}
V_\alpha^{(n)} \le \frac{1}{\sqrt{n\Delta_n}} \sup_{(\theta,t)\in[0,1]\times\mathbb{R}} \Big| \sum_{i=1}^{\lfloor n\theta \rfloor} \Big\{ &\rho_\alpha(\Delta_i^n X^{(n)}) \mathbb{1}_{\{|\Delta_i^n X^{(n)}|>v_n\}} \mathbb{1}_{(-\infty,t]}(\Delta_i^n X^{(n)}) - \rho_\alpha(\Delta_i^n \hat{X}''(\alpha)^n)\times \\
&\times \mathbb{1}_{\{|\Delta_i^n \hat{X}''(\alpha)^n|>v_n\}} \mathbb{1}_{(-\infty,t]}(\Delta_i^n \hat{X}''(\alpha)^n) \Big\} \mathbb{1}_{\{|\Delta_i^n \hat{X}''(\alpha)^n|>\alpha/4\}} \mathbb{1}_{\{|\Delta_i^n \tilde{X}'^n|\le v_n/2\}} \Big| \\
\le C_n(\alpha) + D_n(\alpha),&
\end{aligned}
$$

where we can substitute $\Delta_i^n X^{(n)} = \Delta_i^n \tilde{X}'^n + \Delta_i^n \hat{X}''(\alpha)^n$ in the first line. □

With a similar reasoning as above we deduce a bound for

$$
V_\alpha^{\circ(n)} = \frac{1}{\sqrt{n\Delta_n}} \sup_{(\theta,t)\in[0,1]\times\mathbb{R}} \Big| \sum_{i=1}^{\lfloor n\theta \rfloor} \Big\{ \chi_t^{\circ(\alpha)}(\Delta_i^n X^{(n)}) \mathbb{1}_{\{|\Delta_i^n X^{(n)}|>v_n\}} - \chi_t^{\circ(\alpha)}(\Delta_i^n L^{(n)}) \Big\} \Big|, \tag{A.21}
$$

from (6.32) in the next lemma, where $\alpha > 0$ and $\chi_t^{\circ(\alpha)}$ is defined in (6.3).

Lemma A.10. *Let Assumption 6.12 be satisfied. Then for $\alpha > 0$, $\omega \in Q_n$ and $n \in \mathbb{N}$ large enough such that $v_n \le \alpha$ we have*

$$
V_\alpha^{\circ(n)} \le C_n^\circ(\alpha) + D_n^\circ(\alpha) + E_n^\circ(\alpha),
$$

where

$$
\begin{aligned}
C_n^\circ(\alpha) = \frac{K}{\sqrt{n\Delta_n}} \sup_{t\in\mathbb{R}} \sum_{i=1}^{n} \big| \mathbb{1}_{(-\infty,t]}(\varsigma_i^n(\alpha)) - \mathbb{1}_{(-\infty,t]}(\Delta_i^n \tilde{X}''(8\alpha)^n) \big| \times& \\
\times \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n|>\Delta_n^{\overline{v}}\}} \mathbb{1}_{\{|\Delta_i^n \tilde{X}'^n|\le v_n/2\}},&
\end{aligned}
$$

with $K > 0$ a bound for ρ and

$$
\begin{aligned}
D_n^\circ(\alpha) = \frac{1}{\sqrt{n\Delta_n}} \sum_{i=1}^{n} \big| \rho_\alpha^\circ(\varsigma_i^n(\alpha)) \mathbb{1}_{\{|\varsigma_i^n(\alpha)|>v_n\}} - \rho_\alpha^\circ(\Delta_i^n \tilde{X}''(8\alpha)^n) \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n|>v_n\}} \big| \times& \\
\times \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n|>\Delta_n^{\overline{v}}\}} \mathbb{1}_{\{|\Delta_i^n \tilde{X}'^n|\le v_n/2\}},&
\end{aligned}
$$

$$
\begin{aligned}
E_n^\circ(\alpha) = \frac{1}{\sqrt{n\Delta_n}} \sup_{t\in\mathbb{R}} \sum_{i=1}^{n} \big| &\rho_\alpha^\circ(\varsigma_i^n(\alpha)) \mathbb{1}_{\{|\varsigma_i^n(\alpha)|>v_n\}} \mathbb{1}_{(-\infty,t]}(\varsigma_i^n(\alpha)) - \rho_\alpha^\circ(\Delta_i^n \tilde{X}''(8\alpha)^n)\times \\
&\times \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n|>v_n\}} \mathbb{1}_{(-\infty,t]}(\Delta_i^n \tilde{X}''(8\alpha)^n) \big| \mathbb{1}_{\{|\Delta_i^n \tilde{X}'^n|\le v_n/2\}} \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n|\le \Delta_n^{\overline{v}}\}} \mathbb{1}_{Q_n},
\end{aligned}
$$

where $\varsigma_i^n(\alpha) = \Delta_i^n \tilde{X}'^n + \Delta_i^n \tilde{X}''(8\alpha)^n$, ρ_α° is defined prior to (6.3), $\overline{v} > 0$ is the constant from Assumption 6.12(a(4)I) and the involved processes are defined in (A.8).

Proof. On the set Q_n, and if $v_n \le \alpha$, we have three mutually exclusive possibilities for $1 \le i \le n$:

(i) $\Delta_i^n N^n = 0$.
Then we have $|\Delta_i^n X^{(n)}| = |\Delta_i^n \tilde{X}'^n| \le v_n/2$ and there is no jump larger than u_n (and v_n) on the interval $((i-1)\Delta_n, i\Delta_n]$. Thus, $\chi_t^{\circ(\alpha)}(\Delta_i^n X^{(n)}) \mathbb{1}_{\{|\Delta_i^n X^{(n)}|>v_n\}} = 0 = \chi_t^{\circ(\alpha)}(\Delta_i^n L^{(n)})$ holds for all $t \in \mathbb{R}$ and the i-th summand in (A.21) vanishes.

(ii) $\Delta_i^n N^n = 1$ and $\Delta_i^n \tilde{X}''^n \neq 0$, but $\Delta_i^n \tilde{X}''(8\alpha)^n = 0$.

So the only jump in $((i-1)\Delta_n, i\Delta_n]$ larger than u_n is also larger than 2α. Because $|\Delta_i^n \tilde{X}'^n| \leq v_n/2 \leq \alpha/2$ holds, we have $|\Delta_i^n X^{(n)}| \geq \alpha$, and consequently $\chi_t^{\circ(\alpha)}(\Delta_i^n X^{(n)}) \mathbb{1}_{\{|\Delta_i^n X^{(n)}|>v_n\}} = 0 = \chi_t^{\circ(\alpha)}(\Delta_i^n L^{(n)})$ using the definition of $\chi_t^{\circ(\alpha)}$.

(iii) $\Delta_i^n N^n = 1$ and $\Delta_i^n \tilde{X}''^n = \Delta_i^n \tilde{X}''(8\alpha)^n \neq 0$.

Here we can write

$$\Delta_i^n X^{(n)} = \Delta_i^n \tilde{X}'^n + \Delta_i^n \tilde{X}''(8\alpha)^n =: \varsigma_i^n(\alpha)$$

and

$$\chi_t^{\circ(\alpha)}(\Delta_i^n L^{(n)}) = \rho_\alpha^\circ(\Delta_i^n \tilde{X}''(8\alpha)^n) \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n|>v_n\}} \mathbb{1}_{(-\infty,t]}(\Delta_i^n \tilde{X}''(8\alpha)^n).$$

Therefore, on Q_n and as soon as $v_n \leq \alpha$, we have

$$\begin{aligned} V_\alpha^{\circ(n)} \leq \frac{1}{\sqrt{n\Delta_n}} \sup_{(\theta,t)\in[0,1]\times\mathbb{R}} \Big| \sum_{i=1}^{\lfloor n\theta \rfloor} \Big\{ & \rho_\alpha^\circ(\varsigma_i^n(\alpha)) \mathbb{1}_{\{|\varsigma_i^n(\alpha)|>v_n\}} \mathbb{1}_{(-\infty,t]}(\varsigma_i^n(\alpha)) - \\ & - \rho_\alpha^\circ(\Delta_i^n \tilde{X}''(8\alpha)^n) \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n|>v_n\}} \mathbb{1}_{(-\infty,t]}(\Delta_i^n \tilde{X}''(8\alpha)^n) \Big\} \mathbb{1}_{\{|\Delta_i^n \tilde{X}'^n|\leq v_n/2\}} \Big| \\ \leq C_n^\circ(\alpha) + D_n^\circ(\alpha) + E_n^\circ(\alpha). & \end{aligned}$$

□

In the following lemma we obtain a bound for

$$\hat{V}_\alpha^{(n)} = \frac{1}{\sqrt{n\Delta_n}} \sup_{(\theta,t)\in[0,1]\times\mathbb{R}} \Big| \sum_{i=1}^{\lfloor n\theta \rfloor} \xi_i \big(\chi_t^{(\alpha)}(\Delta_i^n X^{(n)}) \mathbb{1}_{\{|\Delta_i^n X^{(n)}|>v_n\}} - \chi_t^{(\alpha)}(\Delta_i^n L^{(n)}) \big) \Big|, \qquad \text{(A.22)}$$

from (6.45), where $\alpha > 0$, $\chi_t^{(\alpha)}$ is defined in (6.3), $L^{(n)} = \big(z\mathbb{1}_{\{|z|>v_n\}}\big) \star \mu^{(n)}$ is the pure jump Itō semimartingale defined in (6.4) and $(\xi_i)_{i\in\mathbb{N}}$ is a sequence of multipliers with mean zero and variance one defined on a distinct probability space than the remaining processes. Furthermore, for the claim of the lemma below recall the definition of the sets Q_n and $J_n^{(1)}(\alpha)$ in (A.9) and (A.14), respectively, as well as the definition of ρ_α prior to (6.3) and the quantities defined in (A.8).

Lemma A.11. *Let Assumption 6.12 be satisfied. Then for $\alpha > 0$, $\omega \in J_n^{(1)}(\alpha) \cap Q_n$ and $n \in \mathbb{N}$ large enough such that $v_n \leq \alpha/4$ we have*

$$\hat{V}_\alpha^{(n)} \leq \hat{D}_n(\alpha) + \hat{E}_n(\alpha) + \hat{F}_n(\alpha),$$

with

$$\begin{aligned} \hat{D}_n(\alpha) = \frac{1}{\sqrt{n\Delta_n}} \sum_{i=1}^{n} |\xi_i| \big| & \rho_\alpha(\Delta_i^n \tilde{X}'^n + \Delta_i^n \hat{X}''(\alpha)^n) \mathbb{1}_{\{|\Delta_i^n \tilde{X}'^n + \Delta_i^n \hat{X}''(\alpha)^n|>v_n\}} - \\ & - \rho_\alpha(\Delta_i^n \hat{X}''(\alpha)^n) \mathbb{1}_{\{|\Delta_i^n \hat{X}''(\alpha)^n|>v_n\}} \big| \mathbb{1}_{\{|\Delta_i^n \tilde{X}'^n|\leq v_n/2\}}, \end{aligned}$$

$$\hat{E}_n(\alpha) = \sup_{A \in \mathfrak{S}_n} \Big| \sum_{i \in A} \xi_i a_i^n(\alpha) \Big|,$$

and

$$\hat{F}_n(\alpha) = \sup_{A \in \mathfrak{S}_n} \Big| \sum_{i \in A} \xi_i b_i^n(\alpha) \Big|,$$

for

$$a_i^n(\alpha) = \frac{1}{\sqrt{n\Delta_n}} \rho_\alpha(\Delta_i^n \hat{X}''(\alpha)^n) \mathbb{1}_{\{|\Delta_i^n \hat{X}''(\alpha)^n| > \alpha/4\}},$$
$$b_i^n(\alpha) = \frac{1}{\sqrt{n\Delta_n}} \rho_\alpha(\Delta_i^n \tilde{X}'^n + \Delta_i^n \hat{X}''(\alpha)^n) \mathbb{1}_{\{|\Delta_i^n \tilde{X}'^n + \Delta_i^n \hat{X}''(\alpha)^n| > v_n\}} \mathbb{1}_{\{|\Delta_i^n \hat{X}''(\alpha)^n| > \alpha/4\}},$$

and $\mathfrak{S}_n = \{M \subset \{1, \ldots, n\} \mid \# M \le c_n\}$ *with* $c_n = \lceil (v_n/\Delta_n^{\overline{r}}) + 1 \rceil$.

Proof. Recall the cases which have been distinguished in the proof of Lemma A.9:
On Q_n, and with n large enough such that $v_n \le \alpha/4$, one of the following mutually exclusive possibilities holds for $1 \le i \le n$:

(i) $\Delta_i^n N^n = 0$.
Then we have $|\Delta_i^n X^{(n)}| = |\Delta_i^n \tilde{X}'^n| \le v_n/2$ and there is no jump larger than u_n (and v_n) on the interval $((i-1)\Delta_n, i\Delta_n]$. Thus, $\chi_t^{(\alpha)}(\Delta_i^n X^{(n)}) \mathbb{1}_{\{|\Delta_i^n X^{(n)}| > v_n\}} = 0 = \chi_t^{(\alpha)}(\Delta_i^n L^{(n)})$ holds for all $t \in \mathbb{R}$ and the corresponding summand in (A.22) vanishes.

(ii) $\Delta_i^n N^n = 1$ and $\Delta_i^n \tilde{X}''^n = \Delta_i^n \hat{X}''(\alpha)^n \neq 0$.
So the only jump in $((i-1)\Delta_n, i\Delta_n]$ (of absolute size) larger than u_n is in fact not larger than $\alpha/4$, and because of $v_n \le \alpha/4$ we have $|\Delta_i^n X^{(n)}| \le \alpha/2$. Thus, as in the first case, $\chi_t^{(\alpha)}(\Delta_i^n X^{(n)}) \mathbb{1}_{\{|\Delta_i^n X^{(n)}| > v_n\}} = 0 = \chi_t^{(\alpha)}(\Delta_i^n L^{(n)})$ is true for all $t \in \mathbb{R}$ and the corresponding summand in (A.22) is equal to zero.

(iii) $\Delta_i^n N^n = 1$ and $\Delta_i^n \tilde{X}''^n \neq 0$, but $\Delta_i^n \tilde{X}''(\alpha)^n = 0$.
So the only jump in $((i-1)\Delta_n, i\Delta_n]$ larger than u_n is also larger than $\alpha/4$. If $\hat{X}''(\alpha)^n$ is the quantity defined in (A.8), we get

$$\Delta_i^n X^{(n)} = \Delta_i^n \tilde{X}'^n + \Delta_i^n \hat{X}''(\alpha)^n \quad \text{and}$$
$$\chi_t^{(\alpha)}(\Delta_i^n L^{(n)}) = \rho_\alpha(\Delta_i^n \hat{X}''(\alpha)^n) \mathbb{1}_{\{|\Delta_i^n \hat{X}''(\alpha)^n| > v_n\}} \mathbb{1}_{(-\infty, t]}(\Delta_i^n \hat{X}''(\alpha)^n).$$

Thus, we obtain an upper bound for $\hat{V}_\alpha^{(n)}$ on Q_n, as soon as $v_n \le \alpha/4$:

$$\hat{V}_\alpha^{(n)} \le \frac{1}{\sqrt{n\Delta_n}} \sup_{(\theta,t) \in [0,1] \times \mathbb{R}} \Big| \sum_{i \in A_0(\omega;\alpha,n,(\theta,t))} \xi_i \big(\rho_\alpha(\Delta_i^n X^{(n)}) \mathbb{1}_{\{|\Delta_i^n X^{(n)}| > v_n\}} \mathbb{1}_{(-\infty,t]}(\Delta_i^n X^{(n)}) -$$
$$- \rho_\alpha(\Delta_i^n \hat{X}''(\alpha)^n) \mathbb{1}_{\{|\Delta_i^n \hat{X}''(\alpha)^n| > v_n\}} \mathbb{1}_{(-\infty,t]}(\Delta_i^n \hat{X}''(\alpha)^n) \big) \mathbb{1}_{\{|\Delta_i^n \hat{X}''(\alpha)^n| > \alpha/4\}} \Big|, \quad \text{(A.23)}$$

where we can substitute $\Delta_i^n X^{(n)} = \Delta_i^n \tilde{X}'^n + \Delta_i^n \hat{X}''(\alpha)^n$ in the first line and with the random set

$$A_0(\omega; \alpha, n, (\theta, t)) = \big\{ i \in \{1, \ldots, n\} \mid i \le \lfloor n\theta \rfloor \text{ and } \Delta_i^n N^n = 1, \Delta_i^n \tilde{X}''^n \neq 0, \Delta_i^n \tilde{X}''(\alpha)^n = 0 \big\},$$

for $(\theta,t)\in[0,1]\times\mathbb{R}$, $\alpha>0$, $n\in\mathbb{N}$ and $\omega\in\Omega$. Defining the further random sets

$$\begin{aligned}
A_1(\omega;\alpha,n,(\theta,t)) = \big\{i\in A_0(\omega;\alpha,n,(\theta,t)) \mid |\Delta_i^n\hat{X}''(\alpha)^n|>\alpha/4,\\
\mathbb{1}_{(-\infty,t]}(\Delta_i^n\tilde{X}'^n+\Delta_i^n\hat{X}''(\alpha)^n)=\mathbb{1}_{(-\infty,t]}(\Delta_i^n\hat{X}''(\alpha)^n)=1\big\},\\
A_2(\omega;\alpha,n,(\theta,t)) = \big\{i\in A_0(\omega;\alpha,n,(\theta,t)) \mid |\Delta_i^n\hat{X}''(\alpha)^n|>\alpha/4,\\
\mathbb{1}_{(-\infty,t]}(\Delta_i^n\tilde{X}'^n+\Delta_i^n\hat{X}''(\alpha)^n)=0 \text{ and } \mathbb{1}_{(-\infty,t]}(\Delta_i^n\hat{X}''(\alpha)^n)=1\big\},\\
A_3(\omega;\alpha,n,(\theta,t)) = \big\{i\in A_0(\omega;\alpha,n,(\theta,t)) \mid |\Delta_i^n\hat{X}''(\alpha)^n|>\alpha/4,\\
\mathbb{1}_{(-\infty,t]}(\Delta_i^n\tilde{X}'^n+\Delta_i^n\hat{X}''(\alpha)^n)=1 \text{ and } \mathbb{1}_{(-\infty,t]}(\Delta_i^n\hat{X}''(\alpha)^n)=0\big\},
\end{aligned}$$

for $(\theta,t)\in[0,1]\times\mathbb{R}$, $\alpha>0$, $n\in\mathbb{N}$ and $\omega\in\Omega$, together with (A.23) gives for n large enough and $\omega\in Q_n$

$$\begin{aligned}
\hat{V}_\alpha^{(n)} \le \frac{1}{\sqrt{n\Delta_n}}\sup_{(\theta,t)\in[0,1]\times\mathbb{R}}\sum_{i\in A_1(\omega;\alpha,n,(\theta,t))}|\xi_i|\big|\rho_\alpha(\Delta_i^n\tilde{X}'^n+\Delta_i^n\hat{X}''(\alpha)^n)\mathbb{1}_{\{|\Delta_i^n\tilde{X}'^n+\Delta_i^n\hat{X}''(\alpha)^n|>v_n\}}-\\
-\rho_\alpha(\Delta_i^n\hat{X}''(\alpha)^n)\mathbb{1}_{\{|\Delta_i^n\hat{X}''(\alpha)^n|>v_n\}}\big|\mathbb{1}_{\{|\Delta_i^n\tilde{X}'^n|\le v_n/2\}}\\
+\sup_{(\theta,t)\in[0,1]\times\mathbb{R}}\Big|\sum_{i\in A_2(\omega;\alpha,n,(\theta,t))}\xi_i a_i^n(\alpha)\Big| + \sup_{(\theta,t)\in[0,1]\times\mathbb{R}}\Big|\sum_{i\in A_3(\omega;\alpha,n,(\theta,t))}\xi_i b_i^n(\alpha)\Big|, \qquad \text{(A.24)}
\end{aligned}$$

because $|\Delta_i^n\tilde{X}'^n|\le v_n/2$ holds for each $i=1,\dots,n$ on Q_n by (A.9). Finally, according to Lemma A.8 we have $\#A_2(\omega;\alpha,n,(\theta,t))\le c_n$ and $\#A_3(\omega;\alpha,n,(\theta,t))\le c_n$ on $J_n^{(1)}(\alpha)\cap Q_n$ for all $(\theta,t)\in[0,1]\times\mathbb{R}$. By definition $A_1(\omega;\alpha,n,(\theta,t))\subset\{1,\dots,n\}$ and thus (A.24) yields the assertion. □

In the next lemma we use a similar reasoning as above to obtain a bound for

$$\hat{V}_\alpha^{\circ(n)} = \frac{1}{\sqrt{n\Delta_n}}\sup_{(\theta,t)\in[0,1]\times\mathbb{R}}\Big|\sum_{i=1}^{\lfloor n\theta\rfloor}\xi_i\big(\chi_t^{\circ(\alpha)}(\Delta_i^nX^{(n)})\mathbb{1}_{\{|\Delta_i^nX^{(n)}|>v_n\}}-\chi_t^{\circ(\alpha)}(\Delta_i^nL^{(n)})\big)\Big|. \qquad \text{(A.25)}$$

from (6.50), where $\alpha>0$, $\chi_t^{\circ(\alpha)}$ is defined in (6.3), $L^{(n)}=\big(z\mathbb{1}_{\{|z|>v_n\}}\big)\star\mu^{(n)}$ is the pure jump Itō semimartingale defined in (6.4) and $(\xi_i)_{i\in\mathbb{N}}$ is a sequence of multipliers with mean zero and variance one defined on a distinct probability space than the remaining processes. Furthermore for the claim of the lemma below recall the definition of the sets Q_n and $J_n^{(2)}(\alpha)$ in (A.9) and (A.17), respectively, as well as the definition of ρ_α° prior to (6.3) and the quantities defined in (A.8).

Lemma A.12. *Let Assumption 6.12 be satisfied. Then for $\alpha>0$, $\omega\in J_n^{(2)}(\alpha)\cap Q_n$ and $n\in\mathbb{N}$ large enough such that $v_n\le\alpha$ we have*

$$\hat{V}_\alpha^{\circ(n)} \le \hat{C}_n^\circ(\alpha)+\hat{D}_n^\circ(\alpha)+\hat{E}_n^\circ(\alpha)+\hat{F}_n^\circ(\alpha)$$

with

$$\begin{aligned}
\hat{C}_n^\circ(\alpha) = \frac{1}{\sqrt{n\Delta_n}}\sup_{t\in\mathbb{R}}\sum_{i=1}^n|\xi_i|\big|\rho_\alpha^\circ(\varsigma_i^n(\alpha))\mathbb{1}_{\{|\varsigma_i^n(\alpha)|>v_n\}}\mathbb{1}_{(-\infty,t]}(\varsigma_i^n(\alpha))-\\
-\rho_\alpha^\circ(\Delta_i^n\tilde{X}''(8\alpha)^n)\mathbb{1}_{\{|\Delta_i^n\tilde{X}''(8\alpha)^n|>v_n\}}\mathbb{1}_{(-\infty,t]}(\Delta_i^n\tilde{X}''(8\alpha)^n)\big|\mathbb{1}_{\{|\Delta_i^n\tilde{X}''(8\alpha)^n|\le\Delta_n^{\overline{v}}\}}\mathbb{1}_{\{|\Delta_i^n\tilde{X}'^n|\le v_n/2\}},
\end{aligned}$$

$$\hat{D}_n^\circ(\alpha) = \frac{1}{\sqrt{n\Delta_n}} \sum_{i=1}^n |\xi_i| \big| \rho_\alpha^\circ(\varsigma_i^n(\alpha)) \mathbb{1}_{\{|\varsigma_i^n(\alpha)|>v_n\}} - \rho_\alpha^\circ(\Delta_i^n \tilde{X}''(8\alpha)^n) \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n|>v_n\}} \big| \times$$

$$\times \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n|>\Delta_n^{\overline{v}}\}} \mathbb{1}_{\{|\Delta_i^n \tilde{X}'^n|\le v_n/2\}},$$

$$\hat{E}_n^\circ(\alpha) = \sup_{A\in\mathfrak{S}_n} \Big| \sum_{i\in A} \xi_i \bar{a}_i^n(\alpha) \Big|,$$

$$\hat{F}_n^\circ(\alpha) = \sup_{A\in\mathfrak{S}_n} \Big| \sum_{i\in A} \xi_i \bar{b}_i^n(\alpha) \Big|,$$

where $\overline{v} > 0$ is the constant from Assumption 6.12(a(4)I), $\varsigma_i^n(\alpha) = \Delta_i^n \tilde{X}'^n + \Delta_i^n \tilde{X}''(8\alpha)^n$, $\mathfrak{S}_n = \{M \subset \{1,\ldots,n\} \mid \#M \le c_n\}$ for $c_n = \lceil (v_n/\Delta_n^{\overline{r}}) + 1 \rceil$ and with

$$\bar{a}_i^n(\alpha) = \frac{1}{\sqrt{n\Delta_n}} \rho_\alpha^\circ(\Delta_i^n \tilde{X}''(8\alpha)^n) \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n|>v_n \vee \Delta_n^{\overline{v}}\}},$$
$$\bar{b}_i^n(\alpha) = \frac{1}{\sqrt{n\Delta_n}} \rho_\alpha^\circ(\varsigma_i^n(\alpha)) \mathbb{1}_{\{|\varsigma_i^n(\alpha)|>v_n\}} \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n|>\Delta_n^{\overline{v}}\}}.$$

Proof. Recall the cases which we have distinguished in the proof of Lemma A.10:
On the set Q_n, and if $v_n \le \alpha$, we have three mutually exclusive possibilities for $1 \le i \le n$:

(i) $\Delta_i^n N^n = 0$.
Then we have $|\Delta_i^n X^{(n)}| = |\Delta_i^n \tilde{X}'^n| \le v_n/2$ and there is no jump larger than u_n (and v_n) on the interval $((i-1)\Delta_n, i\Delta_n]$. Thus, $\chi_t^{\circ(\alpha)}(\Delta_i^n X^{(n)}) \mathbb{1}_{\{|\Delta_i^n X^{(n)}|>v_n\}} = 0 = \chi_t^{\circ(\alpha)}(\Delta_i^n L^{(n)})$ holds for all $t \in \mathbb{R}$ and the i-th summand in (A.25) vanishes.

(ii) $\Delta_i^n N^n = 1$ and $\Delta_i^n \tilde{X}''^n \ne 0$, but $\Delta_i^n \tilde{X}''(8\alpha)^n = 0$.
So the only jump in $((i-1)\Delta_n, i\Delta_n]$ larger than u_n is also larger than 2α. Because $|\Delta_i^n \tilde{X}'^n| \le v_n/2 \le \alpha/2$ holds, we have $|\Delta_i^n X^{(n)}| \ge \alpha$, and consequently $\chi_t^{\circ(\alpha)}(\Delta_i^n X^{(n)}) \mathbb{1}_{\{|\Delta_i^n X^{(n)}|>v_n\}} = 0 = \chi_t^{\circ(\alpha)}(\Delta_i^n L^{(n)})$ using the definition of $\chi_t^{\circ(\alpha)}$.

(iii) $\Delta_i^n N^n = 1$ and $\Delta_i^n \tilde{X}''^n = \Delta_i^n \tilde{X}''(8\alpha)^n \ne 0$.
Here we can write

$$\Delta_i^n X^{(n)} = \Delta_i^n \tilde{X}'^n + \Delta_i^n \tilde{X}''(8\alpha)^n =: \varsigma_i^n(\alpha)$$

and

$$\chi_t^{\circ(\alpha)}(\Delta_i^n L^{(n)}) = \rho_\alpha^\circ(\Delta_i^n \tilde{X}''(8\alpha)^n) \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n|>v_n\}} \mathbb{1}_{(-\infty,t]}(\Delta_i^n \tilde{X}''(8\alpha)^n).$$

Thus, we have for all $\alpha > 0$, $(\theta,t) \in [0,1]\times\mathbb{R}$, $\omega \in Q_n$ and $n \in \mathbb{N}$ large enough such that $v_n \le \alpha$

$$\hat{V}_\alpha^{\circ(n)} \le \frac{1}{\sqrt{n\Delta_n}} \sup_{(\theta,t)\in[0,1]\times\mathbb{R}} \Big| \sum_{i\in A_0^\circ(\omega;\alpha,n,(\theta,t))} \xi_i \Big\{ \rho_\alpha^\circ(\varsigma_i^n(\alpha)) \mathbb{1}_{\{|\varsigma_i^n(\alpha)|>v_n\}} \mathbb{1}_{(-\infty,t]}(\varsigma_i^n(\alpha)) -$$
$$- \rho_\alpha^\circ(\Delta_i^n \tilde{X}''(8\alpha)^n) \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n|>v_n\}} \mathbb{1}_{(-\infty,t]}(\Delta_i^n \tilde{X}''(8\alpha)^n) \Big\} \Big|,$$

with the random set

$$A_0^\circ(\omega;\alpha,n,(\theta,t)) = \{i \in \{1,\dots,n\} \mid i \le \lfloor n\theta \rfloor \text{ and } \Delta_i^n N^n = 1, \Delta_i^n \tilde{X}''^n = \Delta_i^n \tilde{X}''(8\alpha)^n \neq 0\},$$

for $(\theta,t) \in [0,1]\times\mathbb{R}$, $\alpha > 0$, $n \in \mathbb{N}$ and $\omega \in \Omega$. Furthermore, we define the random sets

$$\begin{aligned}
A_1^\circ(\omega;\alpha,n,(\theta,t)) = \{i \in A_0^\circ(\omega;\alpha,n,(\theta,t)) \mid |\Delta_i^n \tilde{X}''(8\alpha)^n| > \Delta_n^{\overline{v}}&,\\
\mathbb{1}_{(-\infty,t]}(\varsigma_i^n(\alpha)) = \mathbb{1}_{(-\infty,t]}(\Delta_i^n \tilde{X}''(8\alpha)^n) = 1\}&,\\
A_2^\circ(\omega;\alpha,n,(\theta,t)) = \{i \in A_0^\circ(\omega;\alpha,n,(\theta,t)) \mid |\Delta_i^n \tilde{X}''(8\alpha)^n| > \Delta_n^{\overline{v}}&,\\
\mathbb{1}_{(-\infty,t]}(\varsigma_i^n(\alpha)) = 0 \text{ and } \mathbb{1}_{(-\infty,t]}(\Delta_i^n \tilde{X}''(8\alpha)^n) = 1\}&,\\
A_3^\circ(\omega;\alpha,n,(\theta,t)) = \{i \in A_0^\circ(\omega;\alpha,n,(\theta,t)) \mid |\Delta_i^n \tilde{X}''(8\alpha)^n| > \Delta_n^{\overline{v}}&,\\
\mathbb{1}_{(-\infty,t]}(\varsigma_i^n(\alpha)) = 1 \text{ and } \mathbb{1}_{(-\infty,t]}(\Delta_i^n \tilde{X}''(8\alpha)^n) = 0\}&,
\end{aligned}$$

for $(\theta,t) \in [0,1]\times\mathbb{R}$, $\alpha > 0$, $n \in \mathbb{N}$ and $\omega \in \Omega$, with $\overline{v} > 0$ the constant in Assumption 6.12. As a consequence, we obtain for $\omega \in Q_n$, $\alpha > 0$ and n large enough

$$\begin{aligned}
&\hat{V}_\alpha^{\circ(n)} \le \\
&\le \frac{1}{\sqrt{n\Delta_n}} \sup_{(\theta,t)\in[0,1]\times\mathbb{R}} \sum_{i \in A_0^\circ(\omega;\alpha,n,(\theta,t))} |\xi_i| \big| \rho_\alpha^\circ(\varsigma_i^n(\alpha)) \mathbb{1}_{\{|\varsigma_i^n(\alpha)| > v_n\}} \mathbb{1}_{(-\infty,t]}(\varsigma_i^n(\alpha)) \\
&\quad - \rho_\alpha^\circ(\Delta_i^n \tilde{X}''(8\alpha)^n) \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n| > v_n\}} \mathbb{1}_{(-\infty,t]}(\Delta_i^n \tilde{X}''(8\alpha)^n) \big| \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n| \le \Delta_n^{\overline{v}}\}} \mathbb{1}_{\{|\Delta_i^n \tilde{X}'^n| \le v_n/2\}} \\
&\quad + \frac{1}{\sqrt{n\Delta_n}} \sup_{(\theta,t)\in[0,1]\times\mathbb{R}} \sum_{i \in A_1^\circ(\omega;\alpha,n,(\theta,t))} |\xi_i| \big| \rho_\alpha^\circ(\varsigma_i^n(\alpha)) \mathbb{1}_{\{|\varsigma_i^n(\alpha)| > v_n\}} - \\
&\qquad\qquad - \rho_\alpha^\circ(\Delta_i^n \tilde{X}''(8\alpha)^n) \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n| > v_n\}} \big| \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n| > \Delta_n^{\overline{v}}\}} \mathbb{1}_{\{|\Delta_i^n \tilde{X}'^n| \le v_n/2\}} \\
&\quad + \sup_{(\theta,t)\in[0,1]\times\mathbb{R}} \Big| \sum_{i \in A_2^\circ(\omega;\alpha,n,(\theta,t))} \xi_i \bar{a}_i^n(\alpha) \Big| + \sup_{(\theta,t)\in[0,1]\times\mathbb{R}} \Big| \sum_{i \in A_3^\circ(\omega;\alpha,n,(\theta,t))} \xi_i \bar{b}_i^n(\alpha) \Big|,
\end{aligned} \tag{A.26}$$

because $|\Delta_i^n \tilde{X}'^n| \le v_n/2$ holds for each $i = 1,\dots,n$ on Q_n according to (A.9). As a consequence, of Lemma A.8 we have $\# A_2^\circ(\omega;\alpha,n,(\theta,t)) \le c_n$ as well as $\# A_3^\circ(\omega;\alpha,n,(\theta,t)) \le c_n$ with $c_n = \lceil (v_n/\Delta_n^{\overline{v}}) + 1 \rceil$ for each $(\theta,t) \in [0,1]\times\mathbb{R}$, $\alpha > 0$ and $\omega \in J_n^{(2)}(\alpha) \cap Q_n$. Thus, (A.26) yields the assertion. □

A.3 Moments: Bounds and convergence results

In the remaining part of Appendix A we gather results on moments of functionals of processes which occur several times in Section 6.

Lemma A.13. *For $n \in \mathbb{N}$ let $\mu^{(n)}$ be a Poisson random measure with predictable compensator $\bar{\mu}^{(n)}(ds,dz) = \nu_s^{(n)}(dz)ds$ such that (6.8) is satisfied for all $n \in \mathbb{N}$ for $\Delta_n > 0$ and transition kernels $g^{(n)}$ from $([0,1], \mathbb{B}([0,1]))$ into $(\mathbb{R},\mathbb{B})$ with*

$$\Big(\lambda_1 - \text{ess sup}_{y\in[0,1]} \int (1 \wedge |z|^\beta) g^{(n)}(y,dz)\Big) \le K$$

for each $n \in \mathbb{N}$ and some $\beta \in [0,2]$, $K > 0$. Furthermore, let $c > 0$, $F \subset \{z \in \mathbb{R} \mid |z| > c\}$ and $N^{(n)} = \mathbb{1}_F \star \mu^{(n)}$. Then for $0 \le t_1 \le t_2 \le n\Delta_n$ the following equality in distribution holds

$$N_{t_2}^{(n)} - N_{t_1}^{(n)} =_d \text{Poiss}\big(\zeta_{t_2}^{(n)} - \zeta_{t_1}^{(n)}\big), \tag{A.27}$$

with

$$\zeta_t^{(n)} = \int_0^t \int_F \nu_s^{(n)}(dz)ds = n\Delta_n \int_0^{t/(n\Delta_n)} \int_F g^{(n)}(y,dz)dy, \tag{A.28}$$

for $t \in [0, n\Delta_n]$. Moreover, for $i \in \{1,\ldots,n\}$ the sets $A_n^{(i)} := \big\{N_{i\Delta_n}^{(n)} - N_{(i-1)\Delta_n}^{(n)} \le 1\big\}$ satisfy

$$\mathbb{P}\big((A_n^{(i)})^C\big) \le K\Delta_n^2 (c \wedge 1)^{-2\beta}.$$

Proof. (A.27) is a consequence of Theorem II.4.8 in Jacod and Shiryaev (2002). Furthermore, according to (A.27) we calculate as follows

$$\begin{aligned}\mathbb{P}\big((A_n^{(i)})^C\big) &= \exp\Big\{-\Big(\zeta_{i\Delta_n}^{(n)} - \zeta_{(i-1)\Delta_n}^{(n)}\Big)\Big\} \sum_{k=2}^{\infty} \frac{\big(\zeta_{i\Delta_n}^{(n)} - \zeta_{(i-1)\Delta_n}^{(n)}\big)^k}{k!} \\ &\le \Big(\zeta_{i\Delta_n}^{(n)} - \zeta_{(i-1)\Delta_n}^{(n)}\Big)^2 = \Big(n\Delta_n \int_{(i-1)/n}^{i/n} \int_F g^{(n)}(y,dz)dy\Big)^2,\end{aligned}$$

where the final equality in the above display is a consequence of (A.28). Now, using the assumption that the mapping $(y \mapsto \int (1 \wedge |z|^\beta) g^{(n)}(y,dz))$ is Lebesgue almost surely bounded on $[0,1]$ we obtain

$$\mathbb{P}((A_n^{(i)})^C) \le \Big(n\Delta_n (c\wedge 1)^{-\beta} \int_{(i-1)/n}^{i/n} \int (1 \wedge |z|^\beta) g^{(n)}(y,dz)dy\Big)^2 \le K\Delta_n^2 (c \wedge 1)^{-2\beta},$$

because $(c \wedge 1)^{-\beta}(1 \wedge |z|^\beta) \ge 1$ holds if $|z| > c$. □

Lemma A.14. *Let Assumption 6.12 be satisfied and let $\alpha > 0$. Then for $n \in \mathbb{N}$ large enough we have*

$$\mathbb{E}|a_i^n(\alpha)|^m \le \Big(\frac{K(\alpha)}{\sqrt{n\Delta_n}}\Big)^m \Delta_n,$$

for all $m \in \mathbb{N}$ and all $i = 1,\ldots,n$, where for $\alpha > 0$, $n \in \mathbb{N}$ and $i = 1,\ldots,n$

$$a_i^n(\alpha) = \frac{1}{\sqrt{n\Delta_n}} \rho_\alpha(\Delta_i^n \hat{X}''(\alpha)^n) \mathbb{1}_{\{|\Delta_i^n \hat{X}''(\alpha)^n| > \alpha/4\}},$$

with $\hat{X}''(\alpha)^n$ defined in (A.8) and where ρ_α is defined prior to (6.3).

Proof. For $n \in \mathbb{N}$, $i = 1,\ldots,n$ and $\alpha > 0$ define $N^{(\alpha,n)} = \mathbb{1}_{\{|z|>\alpha/4\}} \star \mu^{(n)}$ and $H_i^n(\alpha) = \{N_{i\Delta_n}^{(\alpha,n)} - N_{(i-1)\Delta_n}^{(\alpha,n)} \le 1\}$. Then Lemma A.13 shows that

$$\mathbb{P}\big(H_i^n(\alpha)^C\big) \le K(\alpha)\Delta_n^2.$$

Notice that on $H_i^n(\alpha)$ the quantity $\Delta_i^n \hat{X}''(\alpha)^n$ is either zero or equal to the only jump larger than $\alpha/4$ on the interval $((i-1)\Delta_n, i\Delta_n]$. Thus, we obtain

$$\begin{aligned}\mathbb{E}|a_i^n(\alpha)|^m &\leq \Big(\frac{K}{\sqrt{n\Delta_n}}\Big)^m \mathbb{P}\big(H_i^n(\alpha)^C\big) + \\ &\qquad + \Big(\frac{1}{\sqrt{n\Delta_n}}\Big)^m \mathbb{E}\Big\{ \int\limits_{(i-1)\Delta_n}^{i\Delta_n} \int |\rho_\alpha(z)|^m \mathbb{1}_{\{|z|>\alpha/4\}} \mathbb{1}_{H_i^n(\alpha)} \mu^{(n)}(\omega; du, dz)\Big\} \\ &\leq \Big(\frac{K}{\sqrt{n\Delta_n}}\Big)^m \Big(K(\alpha)\Delta_n^2 + \mathbb{E}\Big\{ \int\limits_{(i-1)\Delta_n}^{i\Delta_n} \int (1\wedge|z|^p)^m \mathbb{1}_{\{|z|>\alpha/4\}} \mu^{(n)}(\omega; du, dz)\Big\}\Big),\end{aligned}$$

where $K>0$ is chosen such that $|\rho(z)| \leq K(1\wedge|z|^p)$ for all $z\in\mathbb{R}$. Furthermore, due to $m\geq 1$ we have $(1\wedge|z|^p)^m \leq (1\wedge|z|^p)$ and consequently the definition of the predictable compensator of an optional $\mathcal{P}'$-σ-finite random measure (see Theorem II.1.8 in Jacod and Shiryaev (2002)) yields

$$\begin{aligned}\mathbb{E}|a_i^n(\alpha)|^m &\leq \Big(\frac{K(\alpha)}{\sqrt{n\Delta_n}}\Big)^m \Big\{\Delta_n^2 + n\Delta_n \int\limits_{(i-1)/n}^{i/n} \int (1\wedge|z|^p) \mathbb{1}_{\{|z|>\alpha/4\}} g^{(n)}(y,dz)dy\Big\} \\ &\leq \Big(\frac{K(\alpha)}{\sqrt{n\Delta_n}}\Big)^m \Delta_n,\end{aligned}$$

for $n\in\mathbb{N}$ large enough due to Assumption 6.12(a1) and $p>\beta$. □

Lemma A.15. *Let Assumption 6.12 be satisfied and let $\alpha>0$. Then for $n\in\mathbb{N}$ large enough we have*

$$\mathbb{E}|\bar{a}_i^n(\alpha)|^m \leq \Big(\frac{K}{\sqrt{n\Delta_n}}\Big)^m \Delta_n,$$

for all $m\in\mathbb{N}$ and all $i=1,\ldots,n$, where for $\alpha>0$, $n\in\mathbb{N}$ and $i=1,\ldots,n$

$$\bar{a}_i^n(\alpha) = \frac{1}{\sqrt{n\Delta_n}} \rho_\alpha^\circ(\Delta_i^n \tilde{X}''(8\alpha)^n) \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n| > v_n \vee \Delta_n^{\overline{v}}\}},$$

with $\tilde{X}''(8\alpha)^n$ defined in (A.8), $\overline{v}>0$ is the constant from Assumption 6.12(a(4)I) and where ρ_α° is defined prior to (6.3).

Proof. For $n\in\mathbb{N}$, $i=1,\ldots,n$ and $\alpha>0$ we define the processes $\bar{N}^{(\alpha,n)} = \mathbb{1}_{\{u_n<|z|\leq 2\alpha\}} \star \mu^{(n)}$, where $u_n = v_n^\ell$ with ℓ the constant in (A.6). If we define the sets $\bar{H}_i^n(\alpha) = \{\bar{N}_{i\Delta_n}^{(\alpha,n)} - \bar{N}_{(i-1)\Delta_n}^{(\alpha,n)} \leq 1\}$, Lemma A.13 yields

$$\mathbb{P}\big(\bar{H}_i^n(\alpha)^C\big) \leq K(\delta)\Delta_n^{2-2\ell\overline{w}(\beta+\delta)}, \tag{A.29}$$

for all $\delta>0$ and $n\in\mathbb{N}$ large enough. Consequently, using the fact that on $\bar{H}_i^n(\alpha)$ the increment $\Delta_i^n \tilde{X}''(8\alpha)^n$ is either zero or equal to the only jump of absolute size in $(u_n, 2\alpha]$ on the interval

$((i-1)\Delta_n, i\Delta_n]$ we obtain

$$\mathbb{E}|\bar{a}_i^n(\alpha)|^m \le \Big(\frac{K(\delta)}{\sqrt{n\Delta_n}}\Big)^m \mathbb{P}\big(\bar{H}_i^n(\alpha)^C\big) +$$

$$+ \Big(\frac{1}{\sqrt{n\Delta_n}}\Big)^m \mathbb{E} \int\limits_{(i-1)\Delta_n}^{i\Delta_n} \int |\rho_\alpha^\circ(z)|^m \mathbb{1}_{\{|z|>v_n \vee \Delta_n^{\overline{v}}\}} \mathbb{1}_{\bar{H}_i^n(\alpha)} \mu^{(n)}(\omega; du, dz)$$

$$\le \Big(\frac{K(\delta)}{\sqrt{n\Delta_n}}\Big)^m \Big\{ \Delta_n^{2-2\ell\overline{w}(\beta+\delta)} + \mathbb{E} \int\limits_{(i-1)\Delta_n}^{i\Delta_n} \int (1\wedge |z|^p)^m \mu^{(n)}(\omega; du, dz) \Big\},$$

for every $\delta > 0$, where $K(\delta) > 0$ is chosen such that (A.29) holds and $|\rho(z)| \le K(\delta)(1\wedge |z|^p)$ (see Assumption 6.12(a2)). Thus, due to $(1\wedge|z|^p)^m \le (1\wedge|z|^p)$ and the definition of the predictable compensator of an optional $\mathcal{P}'$-σ-finite random measure (see Theorem II.1.8 in Jacod and Shiryaev (2002)) we obtain for some small $\delta > 0$ and $n \in \mathbb{N}$ large enough

$$\mathbb{E}|\bar{a}_i^n(\alpha)|^m \le \Big(\frac{K(\delta)}{\sqrt{n\Delta_n}}\Big)^m \Big\{ \Delta_n^{2-2\ell\overline{w}(\beta+\delta)} + n\Delta_n \int\limits_{(i-1)/n}^{i/n} \int (1\wedge|z|^p) g^{(n)}(y, dz) dy \Big\}$$

$$\le \Big(\frac{K(\delta)}{\sqrt{n\Delta_n}}\Big)^m \Delta_n,$$

because of $p > \beta$ and Assumption 6.12(a1), as well as $\ell < 1/(2\beta\overline{w})$. □

The proof of the following lemma requires the notion of Orlicz norms. Recall from Section 2.2 in Van der Vaart and Wellner (1996) that for $\Lambda : \mathbb{R}_+ \to \mathbb{R}$ a non-decreasing, convex function with $\Lambda(0) = 0$ and a random variable Z the Orlicz norm is defined as

$$\|Z\|_\Lambda = \inf\big\{C > 0 \mid \mathbb{E}\Lambda\big(|Z|/C\big) \le 1\big\},$$

where we set $\inf \emptyset = \infty$. It is easy to check that if Λ equals the function $x \mapsto x^p$ for some $p \ge 1$, the corresponding Orlicz norm is the well-known L^p-norm

$$\|Z\|_p = \big(\mathbb{E}|Z|^p\big)^{1/p}.$$

Furthermore for $\Lambda_1(x) := e^x - 1$ a straight forward calculation gives

$$\|Z\|_p \le p!\|Z\|_{\Lambda_1}, \qquad \text{for all } p \in \mathbb{N}, \tag{A.30}$$

because $x^p \le p!(e^x - 1)$ for all $x \in \mathbb{R}_+$ by the series expansion of the exponential function.

Lemma A.16. *Let Assumption 6.12 be satisfied and for $n \in \mathbb{N}$ let $(Z_i^n)_{i=1,\dots,n}$ be independent random variables with mean zero such that there exist constants $C_1, C_2 > 0$ with*

$$\mathbb{E}|Z_i^n|^m \le m!\Big(\frac{C_1}{\sqrt{n\Delta_n}}\Big)^{m-2}\frac{C_2}{n}, \tag{A.31}$$

for every integer $m \ge 2$. Then we have

$$\mathbb{E}\Big\{\sup_{A\in\mathfrak{S}_n}\Big|\sum_{i\in A} Z_i^n\Big|\Big\} = o(1),$$

as $n \to \infty$ for $\mathfrak{S}_n = \{M \subset \{1,\dots,n\} \mid \#M \le c_n\}$ with $c_n = \lceil (v_n/\Delta_n^{\overline{r}}) + 1\rceil$.

Proof. The modified Bernstein inequality (Lemma 2.2.11 in Van der Vaart and Wellner (1996)) and (A.31) yield

$$\mathbb{P}\Big(\big|\sum_{i\in A} Z_i^n\big| > x\Big) \leq 2e^{-\frac{1}{2}\frac{x^2}{b_n+d_nx}}$$

for every $x \in \mathbb{R}_+$, $A \in \mathfrak{S}_n$ with $b_n = 2C_2c_n/n$ and $d_n = C_1/\sqrt{n\Delta_n}$, because each $A \in \mathfrak{S}_n$ consists of at most c_n elements. Therefore, by Lemma 2.2.10 in the previously mentioned reference, the fact that $\#\mathfrak{S}_n \leq (n+1)^{c_n}$ and (A.30) we obtain for a universal constant C and $n \geq 2$

$$\begin{aligned}\mathbb{E}\Big\{\sup_{A\in\mathfrak{S}_n}\Big|\sum_{i\in A} Z_i^n\Big|\Big\} &\leq \Big\|\sup_{A\in\mathfrak{S}_n}\Big|\sum_{i\in A} Z_i^n\Big|\Big\|_{\Lambda_1} \leq C\big(d_n\log(1+(n+1)^{c_n}) + \sqrt{b_n\log(1+(n+1)^{c_n})}\big)\\ &\leq K\frac{1}{\sqrt{n\Delta_n}}\big(v_n/\Delta_n^{\overline{r}}+2\big)\log(2n) + K\big(v_n/\Delta_n^{\overline{r}}+2\big)\sqrt{\log(2n)/n}\\ &\leq K\frac{\log(2n)}{\sqrt{n\Delta_n^{(1+2\overline{r}-2\overline{w})\vee 1}}} = K\frac{\Delta_n^{\delta/2}\log(2n)}{\sqrt{n\Delta_n^{((1+2\overline{r}-2\overline{w})\vee 1)+\delta}}} \to 0, \end{aligned} \tag{A.32}$$

with some $\delta > 0$ such that Assumption 6.12(b7) is satisfied. The final convergence in (A.32) holds, because by Assumption 6.12(b4) we have $\Delta_n = o(n^{-u})$ for some $0 < u < 1$. □

Lemma A.17. *Grant Assumption 6.12. Then we have $a_n(\alpha) = o(1)$ and $b_n(\alpha) = o(1)$ for all $\alpha > 0$ as $n \to \infty$, where*

$$a_n(\alpha) = \frac{1}{\sqrt{n\Delta_n}}\sum_{i=1}^n \mathbb{E}\Big\{\big|\Delta_i^n\hat{X}''(\alpha)^n\big|^p \mathbb{1}_{\{|\Delta_i^n\hat{X}''(\alpha)^n|\leq 2v_n\}}\Big\}$$

and

$$b_n(\alpha) = \frac{v_n}{2\sqrt{n\Delta_n}}\sum_{i=1}^n \mathbb{E}\big|\Delta_i^n\hat{X}''(\alpha)^n\big|^{p-1},$$

with $\hat{X}''(\alpha)^n$ defined in (A.8).

Proof. Obviously, $|z|^p\mathbb{1}_{\{|z|\leq 2v_n\}} \leq 2v_n|z|^{p-1}$ holds for all $z \in \mathbb{R}$. Consequently, $a_n(\alpha) \leq 4b_n(\alpha)$ and it suffices to verify $\lim_{n\to\infty} b_n(\alpha) = 0$. For $\gamma \in \mathbb{R}_+$ set

$$\widehat{\delta}_\alpha^{(n)}(\gamma) = \lambda_1 - \text{ess sup}_{y\in[0,1]}\Big(\int |z|^\gamma \mathbb{1}_{\{|z|>\alpha/4\}} g^{(n)}(y,dz)\Big).$$

Then Assumption 6.12(a1) and (a3) yield $\{\widehat{\delta}_\alpha^{(n)}(1) \vee \widehat{\delta}_\alpha^{(n)}(p-1)\} \leq K < \infty$ for all $\alpha > 0$, $n \in \mathbb{N}$

and we obtain the desired result with Lemma A.2 and Assumption 6.12(b4) as follows:

$$
\begin{aligned}
b_n(\alpha) &\leq \frac{Kv_n}{\sqrt{n\Delta_n}} \sum_{i=1}^{n} \Big\{ \int\limits_{(i-1)\Delta_n}^{i\Delta_n} \int |z|^{p-1} \mathbb{1}_{\{|z|>\alpha/4\}} \nu_s^{(n)}(dz) ds + \\
&\qquad\qquad + \mathbb{1}_{\{p-1\geq 1\}} \Big(\int\limits_{(i-1)\Delta_n}^{i\Delta_n} \int |z| \mathbb{1}_{\{|z|>\alpha/4\}} \nu_s^{(n)}(dz) ds \Big)^{p-1} \Big\} \\
&= \frac{Kv_n}{\sqrt{n\Delta_n}} \sum_{i=1}^{n} \Big\{ n\Delta_n \int\limits_{(i-1)/n}^{i/n} \int |z|^{p-1} \mathbb{1}_{\{|z|>\alpha/4\}} g^{(n)}(y, dz) dy + \\
&\qquad\qquad + \mathbb{1}_{\{p-1\geq 1\}} \Big(n\Delta_n \int\limits_{(i-1)/n}^{i/n} \int |z| \mathbb{1}_{\{|z|>\alpha/4\}} g^{(n)}(y, dz) dy \Big)^{p-1} \Big\} \\
&\leq \frac{Knv_n}{\sqrt{n\Delta_n}} \big\{ \Delta_n + \Delta_n^{(p-1)\vee 1} \big\} = O\Big(\sqrt{n\Delta_n^{1+2\overline{w}}} \Big) = o(1).
\end{aligned}
$$

□

Lemma A.18. *Grant Assumption 6.12. Then we have $c_n(\alpha) = o(1)$ and $d_n(\alpha) = o(1)$ for all $\alpha > 0$ as $n \to \infty$, where*

$$
c_n(\alpha) = \frac{1}{\sqrt{n\Delta_n}} \sum_{i=1}^{n} \mathbb{E}\Big\{ \big|\Delta_i^n \tilde{X}''(8\alpha)^n\big|^p \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n| \leq 2v_n\}} \Big\},
$$

and

$$
d_n(\alpha) = \frac{v_n}{2\sqrt{n\Delta_n}} \sum_{i=1}^{n} \mathbb{E}\big|\Delta_i^n \tilde{X}''(8\alpha)^n\big|^{p-1},
$$

with $\tilde{X}''(8\alpha)^n$ defined in (A.8).

Proof. Obviously, $|z|^p \mathbb{1}_{\{|z|\leq 2v_n\}} \leq 2v_n |z|^{p-1}$ holds for each $z \in \mathbb{R}$. Thus, $c_n(\alpha) \leq 4d_n(\alpha)$ and it is enough to verify $\lim_{n\to\infty} d_n(\alpha) = 0$. Let $\alpha > 0$ be fixed and define further for $\gamma \in \mathbb{R}_+$

$$
\widehat{\delta}_{n,\alpha}(\gamma) = \lambda_1 - \text{ess sup}_{y\in[0,1]} \Big(\int |z|^\gamma \mathbb{1}_{\{u_n<|z|\leq 2\alpha\}} g^{(n)}(y, dz) \Big).
$$

Note that due to Assumption 6.12(a1) and $p - 1 > \beta$ we have for each small $\delta > 0$:

$$
\widehat{\delta}_{n,\alpha}(1) \leq K(\delta) u_n^{-(\beta+\delta-1)_+} \quad \text{and} \quad \widehat{\delta}_{n,\alpha}(p-1) \leq K(\delta).
$$

Furthermore, Lemma A.2 gives

$$
\begin{aligned}
&\mathbb{E}\big|\Delta_i^n \tilde{X}''(8\alpha)^n\big|^{p-1} \le \\
&\le K \times \begin{cases} n\Delta_n \int\limits_{(i-1)/n}^{i/n} \int |z|^{p-1}\mathbb{1}_{\{u_n<|z|\le 2\alpha\}} g^{(n)}(y,dz)dy, & \text{if } p\le 2,\\ n\Delta_n \int\limits_{(i-1)/n}^{i/n} \int |z|^{p-1}\mathbb{1}_{\{u_n<|z|\le 2\alpha\}} g^{(n)}(y,dz)dy + \\ \qquad +\Big(n\Delta_n \int\limits_{(i-1)/n}^{i/n} \int |z|\mathbb{1}_{\{u_n<|z|\le 2\alpha\}} g^{(n)}(y,dz)dy\Big)^{p-1}, & \text{if } p>2, \end{cases}\\
&\le K \times \begin{cases} \Delta_n \widehat{\delta}_{n,\alpha}(p-1), & \text{if } p\le 2,\\ \Delta_n \widehat{\delta}_{n,\alpha}(p-1) + \Delta_n^{p-1}(\widehat{\delta}_{n,\alpha}(1))^{p-1}, & \text{if } p>2, \end{cases}\\
&= K(\delta) \times \begin{cases} \Delta_n, & \text{if } p\le 2,\\ \Delta_n + \Delta_n^{p-1} u_n^{-(p-1)(\beta+\delta-1)_+}, & \text{if } p>2, \end{cases}
\end{aligned}
$$

for each $1 \le i \le n$ and $\delta > 0$ small enough. Thus, when $p \le 2$, Assumption 6.12(b4) yields

$$
d_n(\alpha) \le K(\delta)\frac{1}{\sqrt{n\Delta_n}} n\Delta_n v_n = O\Big(\sqrt{n\Delta_n^{1+2\overline{w}}}\Big) = o(1).
$$

In the case $p > 2$ we obtain also from Assumption 6.12(b4) for δ small enough:

$$
\begin{aligned}
d_n(\alpha) &\le K(\delta)\left\{\frac{1}{\sqrt{n\Delta_n}} n\Delta_n v_n + \frac{1}{\sqrt{n\Delta_n}} n\Delta_n^{p-1} v_n u_n^{-(\beta+\delta-1)_+(p-1)}\right\}\\
&\le K(\delta)\left\{\sqrt{n\Delta_n^{1+2\overline{w}}} + \sqrt{n\Delta_n^{2(p-1)-2(p-1)(\beta+\delta-1)_+\ell\overline{w}-1+2\overline{w}}}\right\}\\
&\le K(\delta)\sqrt{n\Delta_n^{1+2\overline{w}}} = o(1),
\end{aligned}
$$

where the last inequality above is clear for $\beta < 1$, for $\beta = 1$ we have $2(p-1) - 2(p-1)(\beta+\delta-1)\ell\overline{w} - 1 > 1$ from $p > 2$, and in the case $\beta > 1$ we calculate using $\ell < \frac{1}{2\beta\overline{w}}$ and $p > 1+\beta$:

$$
\begin{aligned}
2(p-1) - 2(p-1)(\beta+\delta-1)\ell\overline{w} - 1 &= 2(p-1)(1-(\beta+\delta-1)\ell\overline{w}) - 1\\
&> 2(p-1)\Big(1 - \frac{\beta+\delta-1}{2\beta}\Big) - 1\\
&= (p-1)\Big(1 + \frac{1}{\beta} - \frac{\delta}{\beta}\Big) - 1\\
&> \beta\Big(1 + \frac{1}{\beta} - \frac{\delta}{\beta}\Big) - 1 = \beta - \delta > 1.
\end{aligned}
$$

□

Lemma A.19. *Grant Assumption 6.12. Then we have* $y_n^{(\alpha)} = o(1)$ *and* $z_n^{(\alpha)} = o(1)$ *for all* $\alpha > 0$ *as* $n \to \infty$*, where*

$$
\begin{aligned}
y_n^{(\alpha)} = \frac{1}{\sqrt{n\Delta_n}} \sum_{i=1}^n \mathbb{E}\Big\{&\big|\big|\Delta_i^n \tilde{X}'^n + \Delta_i^n \tilde{X}''(8\alpha)^n\big|^p - \big|\Delta_i^n \tilde{X}''(8\alpha)^n\big|^p\big| \times\\
&\times \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n| \le \Delta_n^{\overline{v}}\}} \mathbb{1}_{\{|\Delta_i^n \tilde{X}'^n + \Delta_i^n \tilde{X}''(8\alpha)^n| > v_n\}} \mathbb{1}_{\{|\Delta_i^n \tilde{X}'^n| \le v_n/2\}} \mathbb{1}_{Q_n}\Big\},
\end{aligned}
$$

and

$$z_n^{(\alpha)} = \frac{1}{\sqrt{n\Delta_n}} \sum_{i=1}^{n} \mathbb{E}\Big\{ \big|\Delta_i^n \tilde{X}''(8\alpha)^n\big|^p \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n| \leq \Delta_n^{\overline{v}}\}} \mathbb{1}_{Q_n} \Big\},$$

with $\overline{v} > 0$ *the constant from Assumption 6.12(a(4)I) and where the involved processes and the set* Q_n *are defined in (A.8) and (A.9), respectively.*

Proof. First we consider $y_n^{(\alpha)}$. The mean value theorem yields

$$y_n^{(\alpha)} \leq \frac{1}{\sqrt{n\Delta_n}} \sum_{i=1}^{n} \mathbb{E}\Big\{ \big|\Delta_i^n \tilde{X}'^n\big| p \xi_i^{p-1} \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n| \leq \Delta_n^{\overline{v}}\}} \times \\ \times \mathbb{1}_{\{|\Delta_i^n \tilde{X}'^n + \Delta_i^n \tilde{X}''(8\alpha)^n| > v_n\}} \mathbb{1}_{\{|\Delta_i^n \tilde{X}'^n| \leq v_n/2\}} \mathbb{1}_{Q_n} \Big\},$$

for some ξ_i between $|\Delta_i^n \tilde{X}''(8\alpha)^n|$ and $|\Delta_i^n \tilde{X}'^n + \Delta_i^n \tilde{X}''(8\alpha)^n|$. Next using the fact that due to the indicators $|\Delta_i^n \tilde{X}'^n| \leq |\Delta_i^n \tilde{X}''(8\alpha)^n|$ holds, we obtain

$$y_n^{(\alpha)} \leq \frac{K}{\sqrt{n\Delta_n}} \sum_{i=1}^{n} \mathbb{E}\Big\{ \big|\Delta_i^n \tilde{X}'^n\big| \big|\Delta_i^n \tilde{X}''(8\alpha)^n\big|^{p-1} \mathbb{1}_{\{|\Delta_i^n \tilde{X}''(8\alpha)^n| \leq \Delta_n^{\overline{v}}\}} \mathbb{1}_{\{|\Delta_i^n \tilde{X}'^n| \leq v_n/2\}} \mathbb{1}_{Q_n} \Big\}.$$

Note that on Q_n the sum $\Delta_i^n \tilde{X}''(8\alpha)^n$ consists of at most one jump. Therefore, we can calculate with the definition of the predictable compensator of the random measure associated with the jumps of $X^{(n)}$:

$$\begin{aligned} y_n^{(\alpha)} &\leq \frac{K v_n}{\sqrt{n\Delta_n}} \sum_{i=1}^{n} \mathbb{E}\Big\{ \big(|z|^{p-1} \mathbb{1}_{\{u_n < |z| \leq \Delta_n^{\overline{v}}\}} \mathbb{1}_{\{(i-1)\Delta_n < s \leq i\Delta_n\}} \big) \star \mu^{(n)} \Big\} \\ &= \frac{K v_n}{\sqrt{n\Delta_n}} \sum_{i=1}^{n} \int_{(i-1)\Delta_n}^{i\Delta_n} \int |z|^{p-1} \mathbb{1}_{\{u_n < |z| \leq \Delta_n^{\overline{v}}\}} \nu_s^{(n)}(dz) ds \\ &= \frac{K n\Delta_n v_n}{\sqrt{n\Delta_n}} \int_0^1 \int |z|^{p-1} \mathbb{1}_{\{u_n < |z| \leq \Delta_n^{\overline{v}}\}} g^{(n)}(y, dz) dy = o\Big(\sqrt{n\Delta_n^{1+2\overline{w}}} \Big) = o(1), \end{aligned}$$

by Assumption 6.12(a1) and (b4), because $p - 1 > \beta$.

Now we show the claim $z_n^{(\alpha)} = o(1)$. This can be seen again by the definition of the predictable compensator of the random measure associated with the jumps of $X^{(n)}$. By the fact that on Q_n $\Delta_i^n \tilde{X}''(8\alpha)^n$ is either 0 or equal to the only jump of absolute size in $(u_n, 2\alpha]$ on the interval $((i-1)\Delta_n, i\Delta_n]$, we have:

$$\begin{aligned} z_n^{(\alpha)} &\leq \frac{1}{\sqrt{n\Delta_n}} \sum_{i=1}^{n} \mathbb{E}\Big\{ \big(|z|^{p} \mathbb{1}_{\{u_n < |z| \leq \Delta_n^{\overline{v}}\}} \mathbb{1}_{\{(i-1)\Delta_n < s \leq i\Delta_n\}} \big) \star \mu^{(n)} \Big\} \\ &= \sqrt{n\Delta_n} \int_0^1 \int |z|^{p} \mathbb{1}_{\{u_n < |z| \leq \Delta_n^{\overline{v}}\}} g^{(n)}(y, dz) dy = O\Big(\sqrt{n\Delta_n^{1+2\overline{v}(p-\beta-\delta)}} \Big) = o(1), \end{aligned}$$

according to Assumption 6.12(b5), for some appropriate $\delta > 0$. □

Lemma A.20. *Let Assumption 6.12 be satisfied and let $f : \mathbb{R} \to \mathbb{R}$ be a Borel measurable function with $|f(z)| \leq K(1 \wedge |z|^p)$ for all $z \in \mathbb{R}$ and some $K > 0$. Then we have*

$$\sup_{i \in \{1,\dots,n\}} \mathbb{E}\big(\big|f(\Delta_i^n L^{(n)})\big|\big) = O(\Delta_n) \tag{A.33}$$

Proof. By the assumptions on f we obtain for $i = 1, \dots, n$ from Proposition 6.3

$$\begin{aligned}
&\mathbb{E}\big(\big|f(\Delta_i^n L^{(n)})\big|\big) \leq K\mathbb{E}\big(1 \wedge |\Delta_i^n L^{(n)}|^p\big) \\
&\leq Kn\Delta_n \int\limits_{(i-1)/n}^{i/n} \int (1 \wedge |z|^p) g^{(n)}(y, dz)dy + O\big(\Delta_n^2 v_n^{-2((\beta+\delta)\wedge 2)} + \Delta_n v_n^{p-((\beta+\delta)\wedge 2)}\big) \\
&\leq K\Big(n\Delta_n \int\limits_{(i-1)/n}^{i/n} \int (1 \wedge |z|^{(\beta+\delta)\wedge 2}) g_0(y, dz)dy + \sqrt{n\Delta_n} \int\limits_{(i-1)/n}^{i/n} \int (1 \wedge |z|^{(\beta+\delta)\wedge 2}) g_1(y, dz)dy \\
&\qquad + n\Delta_n a_n \int\limits_{(i-1)/n}^{i/n} \int (1 \wedge |z|^{(\beta+\delta)\wedge 2}) g_2(y, dz)dy\Big) + O\big(\Delta_n^2 v_n^{-2((\beta+\delta)\wedge 2)} + \Delta_n v_n^{p-((\beta+\delta)\wedge 2)}\big) \\
&= O(\Delta_n) + O\big((\Delta_n/n)^{1/2}\big) + o\big((\Delta_n/n)^{1/2}\big) + O\big(\Delta_n^2 v_n^{-2((\beta+\delta)\wedge 2)} + \Delta_n v_n^{p-((\beta+\delta)\wedge 2)}\big),
\end{aligned} \tag{A.34}$$

for each $\delta > 0$, because by Assumption 6.12(a1) we have $\int (1 \wedge |z|^{(\beta+\delta)\wedge 2}) g_i(y, dz) \leq K(\delta)$ for Lebesgue almost every $y \in [0, 1]$ for all $i \in \{0, 1, 2\}$ and some $K(\delta) > 0$. Furthermore, in the display above a_n denotes a sequence of non-negative real numbers with $a_n = o((n\Delta_n)^{-1/2})$ and $\mathcal{R}_n(y, B) \leq a_n g_2(y, B)$ for all $y \in [0, 1]$, $B \in \mathbb{B}$, $n \in \mathbb{N}$ according to Assumption 6.12. Now, (A.34) yields (A.33) because of three reasons: first $(\Delta_n/n)^{1/2} \leq \Delta_n$ for large $n \in \mathbb{N}$, moreover, $p > \beta$ so $v_n^{p-((\beta+\delta)\wedge 2)} \leq 1$ for large $n \in \mathbb{N}$ and if $\beta = 2$ we have $2((\beta + \delta) \wedge 2)\overline{w} = 4\overline{w} < 1$ due to $\overline{w} < (1/2\beta)$, while in the case $\beta < 2$ we obtain $2((\beta + \delta) \wedge 2)\overline{w} = 2(\beta + \delta)\overline{w} < 1$ for $\delta > 0$ small enough using $\overline{w} < (1/2\beta)$ again. □

Lemma A.21. *Grant Assumption 6.12 and let Q_n be the set defined in (A.9). Then we have for sufficiently large $n \in \mathbb{N}$*

$$\mathbb{E}\big(\big|\rho(\Delta_i^n X^{(n)})\big| \mathbb{1}_{\{|\Delta_i^n X^{(n)}| > v_n\}} \mathbb{1}_{Q_n}\big) \leq K\Delta_n \tag{A.35}$$

$$\mathbb{E}\big(\big|\rho(\Delta_i^n X^{(n)})\big|\big|\rho(\Delta_j^n X^{(n)})\big| \mathbb{1}_{\{|\Delta_i^n X^{(n)}| > v_n\}} \mathbb{1}_{\{|\Delta_j^n X^{(n)}| > v_n\}} \mathbb{1}_{Q_n}\big) \leq K\Delta_n^2 \tag{A.36}$$

for $i, j = 1, \dots, n$ with $i \neq j$, where the constant $K > 0$ is independent of n, i and j.

Proof. Recall the decomposition $X^{(n)} = \tilde{X}'^n + \tilde{X}''^n$ and the sets

$$A_i^n = \{|\Delta_i^n \tilde{X}'^n| \leq v_n/2\} \cap \{\Delta_i^n N^n \leq 1\}$$

in (A.8). According to (A.9) we then have

$$Q_n = \bigcap_{i=1}^n A_i^n$$

and in order to show (A.35) we obtain

$$\begin{aligned}\mathbb{E}\big(\big|\rho(\Delta_i^n X^{(n)})\big|\mathbb{1}_{\{|\Delta_i^n X^{(n)}|>v_n\}}\mathbb{1}_{Q_n}\big) &\leq \mathbb{E}\big(\big|\rho(\Delta_i^n \tilde{X}'^n + \Delta_i^n \tilde{X}''^n)\big|\mathbb{1}_{\{|\Delta_i^n \tilde{X}''^n|>v_n/2\}}\mathbb{1}_{Q_n}\big)\\ &\leq \mathbb{E}\big(\big(\big|\rho(\Delta_i^n \tilde{X}''^n)\big| + \big|\rho'(\xi_i^n)\Delta_i^n \tilde{X}'^n\big|\big)\mathbb{1}_{\{|\Delta_i^n \tilde{X}''^n|>v_n/2\}}\mathbb{1}_{Q_n}\big)\\ &\leq K\mathbb{E}\big((1\wedge\big|\Delta_i^n \tilde{X}''^n\big|^p)\mathbb{1}_{\{|\Delta_i^n \tilde{X}''^n|>u_n\}}\mathbb{1}_{Q_n} + v_n\big|\Delta_i^n \tilde{X}''^n\big|^{p-1}\mathbb{1}_{\{|\Delta_i^n \tilde{X}''^n|>u_n\}}\mathbb{1}_{Q_n}\big),\end{aligned} \tag{A.37}$$

for some ξ_i^n between $\Delta_i^n \tilde{X}''^n$ and $\Delta_i^n \tilde{X}'^n + \Delta_i^n \tilde{X}''^n$ using the mean value theorem and the fact that $|\Delta_i^n \tilde{X}'^n| \leq v_n/2$ on Q_n. Notice furthermore that due to $|\Delta_i^n \tilde{X}'^n| \leq v_n/2$ on Q_n the condition $|\Delta_i^n X^{(n)}| > v_n$ implies $|\Delta_i^n \tilde{X}''^n| > v_n/2$ and consequently $|\Delta_i^n \tilde{X}''^n| > |\Delta_i^n \tilde{X}'^n|$. The final inequality in (A.37) follows with the assumptions on ρ and the definition of $u_n = (v_n)^\ell$ with $\ell > 1$ in (A.7), such that $u_n < v_n/2$ holds for large $n \in \mathbb{N}$. On Q_n $\Delta_i^n \tilde{X}''^n$ is either zero or equal to the only jump in $((i-1)\Delta_n, i\Delta_n]$ of absolute size larger than u_n. Thus, the definition of the predictable compensator of an optional $\mathcal{P}'$-σ-finite random measure (Theorem II.1.8 in Jacod and Shiryaev (2002)) gives

$$\begin{aligned}\mathbb{E}\big((1\wedge\big|\Delta_i^n \tilde{X}''^n\big|^p)\mathbb{1}_{\{|\Delta_i^n \tilde{X}''^n|>u_n\}}\mathbb{1}_{Q_n}\big) &= \mathbb{E}\big(\big((1\wedge|z|^p)\mathbb{1}_{\{|z|>u_n\}}\mathbb{1}_{Q_n}\mathbb{1}_{\{(i-1)\Delta_n<s\leq i\Delta_n\}}\big)\star\mu^{(n)}\big)\\ &\leq \mathbb{E}\big((1\wedge|z|^p)\mathbb{1}_{\{(i-1)\Delta_n<s\leq i\Delta_n\}}\star\mu^{(n)}\big)\\ &= n\Delta_n\int_{(i-1)/n}^{i/n}\int(1\wedge|z|^p)g^{(n)}(y,dz)dy \leq K\Delta_n,\end{aligned} \tag{A.38}$$

where the last inequality above is a consequence of Assumption 6.12(a1) and $p > \beta$. With the same reasoning we obtain for the second summand in (A.37)

$$\begin{aligned}\mathbb{E}\big(\big|\Delta_i^n \tilde{X}''^n\big|^{p-1}\mathbb{1}_{\{|\Delta_i^n \tilde{X}''^n|>u_n\}}\mathbb{1}_{Q_n}\big) &= \mathbb{E}\big(\big(|z|^{p-1}\mathbb{1}_{\{|z|>u_n\}}\mathbb{1}_{Q_n}\mathbb{1}_{\{(i-1)\Delta_n<s\leq i\Delta_n\}}\big)\star\mu^{(n)}\big)\\ &\leq \mathbb{E}\big(|z|^{p-1}\mathbb{1}_{\{(i-1)\Delta_n<s\leq i\Delta_n\}}\star\mu^{(n)}\big)\\ &= n\Delta_n\int_{(i-1)/n}^{i/n}\int|z|^{p-1}g^{(n)}(y,dz)dy \leq K\Delta_n,\end{aligned} \tag{A.39}$$

using Assumption 6.12(a3) for the last estimate above. (A.37), (A.38) and (A.39) yield (A.35). In order to prove (A.36) we use the mean value theorem, the definition of Q_n and the assumptions on ρ to obtain for $i \neq j$ similar to (A.37)

$$\begin{aligned}&\mathbb{E}\big(\big|\rho(\Delta_i^n X^{(n)})\big|\big|\rho(\Delta_j^n X^{(n)})\big|\mathbb{1}_{\{|\Delta_i^n X^{(n)}|>v_n\}}\mathbb{1}_{\{|\Delta_j^n X^{(n)}|>v_n\}}\mathbb{1}_{Q_n}\big)\\ &\quad\leq \mathbb{E}\big(\big(\big|\rho(\Delta_i^n \tilde{X}''^n)\big| + \big|\rho'(\xi_i^n)\Delta_i^n \tilde{X}'^n\big|\big)\big(\big|\rho(\Delta_j^n \tilde{X}''^n)\big| + \big|\rho'(\xi_j^n)\Delta_j^n \tilde{X}'^n\big|\big)\times\\ &\qquad\qquad\times\mathbb{1}_{\{|\Delta_i^n \tilde{X}''^n|>v_n/2\}}\mathbb{1}_{\{|\Delta_j^n \tilde{X}''^n|>v_n/2\}}\mathbb{1}_{Q_n}\big)\\ &\quad\leq K\mathbb{E}\big((1\wedge\big|\Delta_i^n \tilde{X}''^n\big|^p)(1\wedge\big|\Delta_j^n \tilde{X}''^n\big|^p)\mathbb{1}_{\{|\Delta_i^n \tilde{X}''^n|>u_n\}}\mathbb{1}_{\{|\Delta_j^n \tilde{X}''^n|>u_n\}}\mathbb{1}_{Q_n}\big)\\ &\qquad+ Kv_n\mathbb{E}\big((1\wedge\big|\Delta_i^n \tilde{X}''^n\big|^p)\big|\Delta_j^n \tilde{X}''^n\big|^{p-1}\mathbb{1}_{\{|\Delta_i^n \tilde{X}''^n|>u_n\}}\mathbb{1}_{\{|\Delta_j^n \tilde{X}''^n|>u_n\}}\mathbb{1}_{Q_n}\big)\\ &\qquad+ Kv_n\mathbb{E}\big((1\wedge\big|\Delta_j^n \tilde{X}''^n\big|^p)\big|\Delta_i^n \tilde{X}''^n\big|^{p-1}\mathbb{1}_{\{|\Delta_i^n \tilde{X}''^n|>u_n\}}\mathbb{1}_{\{|\Delta_j^n \tilde{X}''^n|>u_n\}}\mathbb{1}_{Q_n}\big)\\ &\qquad+ Kv_n^2\mathbb{E}\big(\big|\Delta_i^n \tilde{X}''^n\big|^{p-1}\big|\Delta_j^n \tilde{X}''^n\big|^{p-1}\mathbb{1}_{\{|\Delta_i^n \tilde{X}''^n|>u_n\}}\mathbb{1}_{\{|\Delta_j^n \tilde{X}''^n|>u_n\}}\mathbb{1}_{Q_n}\big),\end{aligned} \tag{A.40}$$

for some ξ_i^n between $\Delta_i^n \tilde{X}''^n$ and $\Delta_i^n \tilde{X}'^n + \Delta_i^n \tilde{X}''^n$ and some ξ_j^n between $\Delta_j^n \tilde{X}''^n$ and $\Delta_j^n \tilde{X}'^n + \Delta_j^n \tilde{X}''^n$. $\Delta_i^n \tilde{X}''^n$ is on Q_n either zero or equal to the only jump of absolute size larger than u_n on the interval $((i-1)\Delta_n, i\Delta_n]$. The same is true for $\Delta_j^n \tilde{X}''^n$ on $((j-1)\Delta_n, j\Delta_n]$. Therefore, the definition of the predictable compensator of an optional $\mathcal{P}'$-σ-finite random measure yields for the first resulting summand in (A.40)

$$
\begin{aligned}
&\mathbb{E}\big(\big(1 \wedge \big|\Delta_i^n \tilde{X}''^n\big|^p\big)\big(1 \wedge \big|\Delta_j^n \tilde{X}''^n\big|^p\big) \mathbb{1}_{\{|\Delta_i^n \tilde{X}''^n| > u_n\}} \mathbb{1}_{\{|\Delta_j^n \tilde{X}''^n| > u_n\}} \mathbb{1}_{Q_n}\big) \\
&\quad = \mathbb{E}\Big(\big((1 \wedge |z|^p) \mathbb{1}_{\{|z| > u_n\}} \mathbb{1}_{\{(i-1)\Delta_n < s \le i\Delta_n\}} \star \mu^{(n)}\big) \times \\
&\qquad\qquad\qquad \times \big((1 \wedge |z|^p) \mathbb{1}_{\{|z| > u_n\}} \mathbb{1}_{\{(j-1)\Delta_n < s \le j\Delta_n\}} \star \mu^{(n)}\big) \mathbb{1}_{Q_n}\Big) \\
&\quad \le \mathbb{E}\Big(\big((1 \wedge |z|^p) \mathbb{1}_{\{(i-1)\Delta_n < s \le i\Delta_n\}} \star \mu^{(n)}\big)\big((1 \wedge |z|^p) \mathbb{1}_{\{(j-1)\Delta_n < s \le j\Delta_n\}} \star \mu^{(n)}\big)\Big) \\
&\quad = \mathbb{E}\big((1 \wedge |z|^p) \mathbb{1}_{\{(i-1)\Delta_n < s \le i\Delta_n\}} \star \mu^{(n)}\big) \mathbb{E}\big((1 \wedge |z|^p) \mathbb{1}_{\{(j-1)\Delta_n < s \le j\Delta_n\}} \star \mu^{(n)}\big) \le K \Delta_n^2.
\end{aligned} \tag{A.41}
$$

The final equality above follows using the fact that $\mu^{(n)}$ is a Poisson random measure and thus both involved factors are independent (see Theorem II.4.8 in Jacod and Shiryaev (2002)). The last estimate in (A.41) is a consequence of (A.38). The remaining summands in (A.40) can be treated similarly by exploiting the properties of a Poisson random measure as well as (A.38) and (A.39). □

B Results on the limiting process from Theorem 6.1

B.1 Useful properties of the Gaussian limit and its covariance semimetric

In the following we collect several lemmas which are useful to obtain bounds for the expectation of sup-functionals of the process $\mathbb{G}_f$. In particular, we apply them in the proof of Proposition 6.6 in order to show asymptotical uniform d-equicontinuity in probability of a sequence of processes $\mathbb{G}_{f_n}$ for some suitable semimetric d.

Lemma B.1. *Grant Assumption 6.12, let $f : \mathbb{R} \to \mathbb{R}$ be Borel measurable with $|f(z)| \le K(1 \wedge |z|^p)$ for all $z \in \mathbb{R}$ and some $K > 0$. Furthermore, let $\mathbb{G}_f$ be the tight centered Gaussian process in $\ell^\infty([0,1] \times \mathbb{R})$ defined in Theorem 6.1. Then for $(\theta_1, t_1), (\theta_2, t_2) \in [0,1] \times \mathbb{R}$ the L^8-norm satisfies*

$$\|\mathbb{G}_f(\theta_1, t_1) - \mathbb{G}_f(\theta_2, t_2)\|_8 = 105^{\frac{1}{8}} d_f((\theta_1, t_1); (\theta_2, t_2)),$$

with d_f the semimetric defined in (6.2).

Proof. For $(\theta_1, t_1), (\theta_2, t_2) \in [0,1] \times \mathbb{R}$ with $\theta_1 \le \theta_2$ we have

$$\begin{aligned}
&\mathbb{E}\big(\mathbb{G}_f(\theta_1, t_1) - \mathbb{G}_f(\theta_2, t_2)\big)^2 \\
&\quad = H_f((\theta_1, t_1); (\theta_1, t_1)) - 2H_f((\theta_1, t_1); (\theta_2, t_2)) + H_f((\theta_2, t_2); (\theta_2, t_2)) \\
&\quad = \int_0^{\theta_1} \int_{-\infty}^{t_1} f^2(z) g_0(y, dz) dy - 2 \int_0^{\theta_1} \int_{-\infty}^{t_1 \wedge t_2} f^2(z) g_0(y, dz) dy + \int_0^{\theta_2} \int_{-\infty}^{t_2} f^2(z) g_0(y, dz) dy \\
&\quad = d_f^2((\theta_1, t_1); (\theta_2, t_2)),
\end{aligned}$$

where the last equation follows by distinguishing the cases $t_1 \le t_2$ and $t_2 \le t_1$. Therefore, using the properties of the normal distribution we obtain for the L^8-norm

$$\|\mathbb{G}_f(\theta_1, t_1) - \mathbb{G}_f(\theta_2, t_2)\|_8 = 105^{\frac{1}{8}} d_f((\theta_1, t_1); (\theta_2, t_2)),$$

for arbitrary $(\theta_1, t_1), (\theta_2, t_2) \in [0,1] \times \mathbb{R}$. □

Lemma B.2. *Grant Assumption 6.12 and let $f : \mathbb{R} \to \mathbb{R}$ be Borel measurable with $|f(z)| \le K(1 \wedge |z|^p)$ for all $z \in \mathbb{R}$ and some $K > 0$. Then the tight centered Gaussian process $\mathbb{G}_f$ defined in Theorem 6.1 is separable with respect to the semimetric d_f in the sense of Theorem 2.2.4 in Van der Vaart and Wellner (1996). More precisely, for every $\delta > 0$*

$$\sup_{d_f((\theta_1,t_1);(\theta_2,t_2))<\delta} \big|\mathbb{G}_f(\theta_1, t_1) - \mathbb{G}_f(\theta_2, t_2)\big| = \sup_{\substack{d_f((\theta_1,t_1);(\theta_2,t_2))<\delta \\ (\theta_1,t_1),(\theta_2,t_2)\in([0,1]\times\mathbb{R})\cap\mathbb{Q}^2}} \big|\mathbb{G}_f(\theta_1, t_1) - \mathbb{G}_f(\theta_2, t_2)\big|$$

holds almost surely.

Proof. By the assumptions on the involved quantities for each $t \in \mathbb{R}$ the function $[0,1] \ni \theta \mapsto \int_0^\theta \int_{-\infty}^t f^2(z) g_0(y, dz) dy$ is continuous and for each $\theta \in [0,1]$ the function $\mathbb{R} \ni t \mapsto \int_0^\theta \int_{-\infty}^t f^2(z) g_0(y, dz) dy$ is right-continuous. As a consequence, we can find for every $\varepsilon > 0$ and $(\theta_1, t_1) \in [0,1] \times \mathbb{R}$ a $(\theta_2, t_2) \in ([0,1] \times \mathbb{R}) \cap \mathbb{Q}^2$ with $d_f((\theta_1, t_1); (\theta_2, t_2)) < \varepsilon$. Thus, for every $\delta > 0$

$$\sup_{d_f((\theta_1,t_1);(\theta_2,t_2))<\delta} \big|\mathbb{G}_f(\theta_1, t_1) - \mathbb{G}_f(\theta_2, t_2)\big| = \sup_{\substack{d_f((\theta_1,t_1);(\theta_2,t_2))<\delta \\ (\theta_1,t_1),(\theta_2,t_2)\in([0,1]\times\mathbb{R})\cap\mathbb{Q}^2}} \big|\mathbb{G}_f(\theta_1, t_1) - \mathbb{G}_f(\theta_2, t_2)\big|$$

holds almost surely, because the sample paths of $\mathbb{G}_f$ are almost surely uniformly d_f-continuous. □

Lemma B.3. *Grant Assumption 6.12 and let $f : \mathbb{R} \to \mathbb{R}$ be Borel measurable with $|f(z)| \le C(1 \wedge |z|^p)$ for all $z \in \mathbb{R}$ and some $C > 0$. Then for d_f the semimetric defined in (6.2) the semimetric space $([0,1] \times \mathbb{R}, d_f)$ is totally bounded and there exists a $K > 0$ which depends only on C such that for every $\varepsilon > 0$*

$$D(\varepsilon, d_f) \le K/\varepsilon^4,$$

where $D(\varepsilon, d_f)$ denotes the packing number of $[0,1] \times \mathbb{R}$ with respect to d_f at distance $\varepsilon > 0$.

Proof. By the well-known relation $D(\varepsilon, d_f) \leq N(\varepsilon/2, d_f)$ of the packing number and the covering number $N(\varepsilon/2, d_f)$ of $([0,1]\times\mathbb{R}, d_f)$ at radius $\varepsilon/2$ it suffices to show that there exists a $K > 0$ with $N(\varepsilon/2, d_f) \leq K/\varepsilon^4$ for every $\varepsilon > 0$.

The measure $\mathbb{B} \ni A \mapsto C^2 \int_0^1 \int_A (1 \wedge |z|^{2p}) g_0(y, dz) dy$ is finite. Therefore, for each $\varepsilon > 0$ we can find a finite partition $\{t_0 = -\infty < t_1 < \ldots < t_m < t_{m+1} = \infty\}$ of $\overline{\mathbb{R}}$ with $m \leq K/\varepsilon^2$ for some $K > 0$ which depends only on C such that $C^2 \int_0^1 \int_{t_j}^{t_{j+1}} (1 \wedge |z|^{2p}) g_0(y, dz) dy < \varepsilon^2/64$ for each $j = 0, \ldots, m$. For the same reason there is also a finite partition $\{\theta_0 = 0 < \theta_1 < \ldots < \theta_\ell < \theta_{\ell+1} = 1\}$ of $[0,1]$ with $\ell \leq K/\varepsilon^2$ and $C^2 \int_{\theta_i}^{\theta_{i+1}} \int (1 \wedge |z|^{2p}) g_0(y, dz) dy < \varepsilon^2/64$ for all $i = 0, \ldots, \ell$. Furthermore, consider the collection $M := \{(\theta_i, t_j) \mid i = 1, \ldots, \ell; j = 1, \ldots, m\}$ which consists of at most K/ε^4 points. Then for an arbitrary $(\theta, t) \in [0,1]\times\mathbb{R}$ let $i_0 \in \operatorname{argmin}\{|\theta_i - \theta| \mid i = 1, \ldots, \ell\}$, $j_0 \in \operatorname{argmin}\{|t_j - t| \mid j = 1, \ldots, m\}$ and choose $i_1 \in \{0, \ldots, \ell+1\}$, $j_1 \in \{0, \ldots, m+1\}$ such that $\theta \in [\theta_{i_0} \wedge \theta_{i_1}, \theta_{i_0} \vee \theta_{i_1}]$ as well as $t \in [t_{j_0} \wedge t_{j_1}, t_{j_0} \vee t_{j_1}]$ to obtain

$$
\begin{aligned}
d_f((\theta,t);(\theta_{i_0}, t_{j_0})) \leq 2 \max\Bigg\{ &\Big(C^2 \int_0^1 \int_{t_{j_0}\wedge t_{j_1}}^{t_{j_0}\vee t_{j_1}} (1 \wedge |z|^{2p}) g_0(y, dz) dy \Big)^{\frac{1}{2}}, \\
&\Big(C^2 \int_{\theta_{i_0}\wedge\theta_{i_1}}^{\theta_{i_0}\vee\theta_{i_1}} \int (1 \wedge |z|^{2p}) g_0(y, dz) dy \Big)^{\frac{1}{2}} \Bigg\} \\
\leq \varepsilon/4 < \varepsilon/2.
\end{aligned}
$$

Thus, we have $N(\varepsilon/2, d_f) \leq K/\varepsilon^4$, because by the inequality above the d_f-balls with radius $\varepsilon/2$ around the points of M cover $[0,1]\times\mathbb{R}$. $\square$

B.2 An auxiliary result on the supremum of the Gaussian limit

A further application of the lemmas in Section B.1 is the proposition below which is necessary to prove Theorem 6.8.

Proposition B.4. *Grant Assumption 6.12 and for $\alpha > 0$ let ρ_α° be the function defined prior to (6.3). Then we have*

$$
\lim_{\alpha\to 0} \mathbb{E}\Big(\sup_{(\theta,t)\in[0,1]\times\mathbb{R}} |\mathbb{G}_{\rho_\alpha^\circ}(\theta, t)| \Big) = 0.
$$

Proof. We want to use Corollary 2.2.5 in Van der Vaart and Wellner (1996) for the convex, non-decreasing, non-zero function $\varphi(x) = x^8$ which clearly satisfies $\varphi(0) = 0$ and $\limsup_{x,y\to\infty} \varphi(x)\varphi(y)/\,\varphi(cxy) < \infty$ for some constant $c > 0$. Due to Lemma B.2 the process $\mathbb{G}_{\rho_\alpha^\circ}$ is separable in the sense of this corollary. Furthermore, Lemma B.1 shows

$$
\|\mathbb{G}_{\rho_\alpha^\circ}(\theta_1, t_1) - \mathbb{G}_{\rho_\alpha^\circ}(\theta_2, t_2)\|_8 = 105^{\frac{1}{8}} d_{\rho_\alpha^\circ}((\theta_1, t_1); (\theta_2, t_2)),
$$

for all $(\theta_1, t_1), (\theta_2, t_2) \in [0,1]\times\mathbb{R}$, where

$$
d_{\rho_\alpha^\circ}((\theta_1, t_1); (\theta_2, t_2)) = \Big\{ \int_0^{\theta_1} \int_{t_1\wedge t_2}^{t_1\vee t_2} (\rho_\alpha^\circ(z))^2 g_0(y, dz) dy + \int_{\theta_1}^{\theta_2} \int_{-\infty}^{t_2} (\rho_\alpha^\circ(z))^2 g_0(y, dz) dy \Big\}^{1/2}, \quad (\theta_1 \leq \theta_2)
$$

is the semimetric for which the sample paths of $\mathbb{G}_{\rho_\alpha^\circ}$ are almost surely uniformly continuous. Thus, according to Corollary 2.2.5 in Van der Vaart and Wellner (1996) there exists a constant $K > 0$ which does not depend on ρ or α such that

$$\Big\| \sup_{(\theta_1,t_1),(\theta_2,t_2)\in[0,1]\times\mathbb{R}} |\mathbb{G}_{\rho_\alpha^\circ}(\theta_1,t_1) - \mathbb{G}_{\rho_\alpha^\circ}(\theta_2,t_2)| \Big\|_8 \le K \int_0^{\bar{d}(\alpha)} D(\varepsilon, d_{\rho_\alpha^\circ})^{\frac{1}{8}} d\varepsilon, \tag{B.1}$$

where $D(\varepsilon, d_{\rho_\alpha^\circ})$ denotes the packing number of $[0,1] \times \mathbb{R}$ at distance ε with respect to the semimetric $d_{\rho_\alpha^\circ}$ and where

$$\begin{aligned} \bar{d}(\alpha) &= \operatorname{diam}([0,1] \times \mathbb{R}; d_{\rho_\alpha^\circ}) = \sup_{(\theta_1,t_1),(\theta_2,t_2)\in[0,1]\times\mathbb{R}} d_{\rho_\alpha^\circ}((\theta_1,t_1);(\theta_2,t_2)) \\ &\le \Big\{ 2 \int_0^1 \int (\rho_\alpha^\circ(z))^2 g_0(y,dz)dy \Big\}^{\frac{1}{2}} \overset{\alpha\to 0}{\longrightarrow} 0 \end{aligned} \tag{B.2}$$

is the diameter of $[0,1]\times\mathbb{R}$ with respect to $d_{\rho_\alpha^\circ}$. The convergence in the display above holds due to Lebesgue's dominated convergence theorem by the assumptions on ρ_α° and on g_0. Moreover, Lemma B.3 gives a constant $K > 0$ which is independent of α such that $D(\varepsilon, d_{\rho_\alpha^\circ}) \le K/\varepsilon^4$. Thus, with (B.1) and (B.2) we obtain the desired result:

$$\begin{aligned} \mathbb{E}\Big(\sup_{(\theta,t)\in[0,1]\times\mathbb{R}} |\mathbb{G}_{\rho_\alpha^\circ}(\theta,t)| \Big) &\le \|\mathbb{G}_{\rho_\alpha^\circ}(0,0)\|_2 + \Big\| \sup_{(\theta_1,t_1),(\theta_2,t_2)\in[0,1]\times\mathbb{R}} |\mathbb{G}_{\rho_\alpha^\circ}(\theta_1,t_1) - \mathbb{G}_{\rho_\alpha^\circ}(\theta_2,t_2)| \Big\|_8 \\ &\le K \int_0^{\bar{d}(\alpha)} \varepsilon^{-\frac{1}{2}} d\varepsilon = K\bar{d}(\alpha)^{\frac{1}{2}} \overset{\alpha\to 0}{\longrightarrow} 0. \end{aligned}$$

□

C Auxiliary Results

The following lemma shows that two bootstrapped random elements with values in some metric space $(\mathbb{D}, d)$ which differ only by a term of order $o_{\mathbb{P}}(1)$ converge simultaneously weakly conditional on the data in probability.

Lemma C.1. *Let $\hat{G}_n = \hat{G}_n(X_1, \dots, X_n, \xi_1, \dots, \xi_n)$ and $\hat{H}_n = \hat{H}_n(X_1, \dots, X_n, \xi_1, \dots, \xi_n)$ be two sequences of bootstrapped elements with values in some metric space $(\mathbb{D}, d)$ such that $d(\hat{G}_n, \hat{H}_n) \overset{\mathbb{P}^*}{\to} 0$. Then for a tight Borel measurable process G in $\mathbb{D}$, we have $\hat{G}_n \rightsquigarrow_\xi G$ if and only if $\hat{H}_n \rightsquigarrow_\xi G$.*

Proof. See Bücher (2011), Lemma A.1. □

The next auxiliary result is useful in order to show consistency of the test procedures in this paper. In the assertion of this proposition $(\xi^{(b)})_{b=1,\dots,B}$ for some $B \in \mathbb{N}$ denote independent sequences $\xi^{(b)} = (\xi_i^{(b)})_{i\in\mathbb{N}}$ of random variables satisfying Assumption 3.6. Furthermore, $\hat{G}_{\rho,\xi^{(b)}}^{(n)}$ denotes the process defined in (6.7) calculated with respect to the b-th multiplier sequence $\xi^{(b)}$.

Proposition C.2. *Let $B \in \mathbb{N}$. If $\mathbf{H}_1^{(loc)}$ is true and each of the independent multiplier sequences $(\xi^{(b)})_{b=1,\dots,B}$ satisfies Assumption 3.6, then we have*

$$\left(\mathbb{T}_\rho^{(n)}, \hat{\mathbb{T}}_{\rho,\xi^{(1)}}^{(n)}, \dots, \hat{\mathbb{T}}_{\rho,\xi^{(B)}}^{(n)}\right) \rightsquigarrow \left(\mathbb{T}_\rho + \mathbb{T}_{\rho,g_1}, \mathbb{T}_{\rho,(1)}, \dots, \mathbb{T}_{\rho,(B)}\right)$$

in $(\ell^\infty([0,1] \times \mathbb{R}))^{B+1}$ and

$$\left(\mathbb{H}_\rho^{(n)}, \hat{\mathbb{H}}_{\rho,\xi^{(1)}}^{(n)}, \dots, \hat{\mathbb{H}}_{\rho,\xi^{(B)}}^{(n)}\right) \rightsquigarrow \left(\mathbb{H}_\rho + D_\rho^{(g_1)}, \mathbb{H}_{\rho,(1)}, \dots, \mathbb{H}_{\rho,(B)}\right)$$

in $(\ell^\infty(C \times \mathbb{R}))^{B+1}$, where $\rightsquigarrow$ denotes (unconditional) weak convergence (with respect to the (joint) probability measure $\mathbb{P}$), furthermore $\mathbb{T}_{\rho,(1)}, \dots, \mathbb{T}_{\rho,(B)}$ are independent copies of the Gaussian process $\mathbb{T}_\rho$ in Theorem 3.1 and $\mathbb{H}_{\rho,(1)}, \dots, \mathbb{H}_{\rho,(B)}$ are independent copies of the stochastic process $\mathbb{H}_\rho$ defined in Theorem 4.4.

Proof. The claim follows by exactly the same reasoning as in the proof of Proposition A.2 in Bücher et al. (2017). □