



HAL
open science

ESTIMATING FAST MEAN-REVERTING JUMPS IN ELECTRICITY MARKET MODELS

Thomas Deschatre, Marc Hoffmann

► **To cite this version:**

Thomas Deschatre, Marc Hoffmann. ESTIMATING FAST MEAN-REVERTING JUMPS IN ELECTRICITY MARKET MODELS. 2018. hal-01966337

HAL Id: hal-01966337

<https://hal.science/hal-01966337>

Preprint submitted on 28 Dec 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ESTIMATING FAST MEAN-REVERTING JUMPS IN ELECTRICITY MARKET MODELS

THOMAS DESCHATRE, OLIVIER FÉRON AND MARC HOFFMANN

ABSTRACT. Based on empirical evidence of fast mean-reverting spikes, we model electricity price processes $X + Z^\beta$ as the sum of a continuous Itô semimartingale X and a mean-reverting compound Poisson process $Z_t^\beta = \int_0^t \int_{\mathbb{R}} x e^{-\beta(t-s)} \underline{p}(ds, dt)$ where $\underline{p}(ds, dt)$ is Poisson random measure with intensity $\lambda ds \otimes dt$. In a first part, we investigate the estimation of (λ, β) from discrete observations and establish asymptotic efficiency in various asymptotic settings. In a second part, we discuss the use of our inference results for correcting the value of forward contracts on electricity markets in presence of spikes. We implement our method on real data in the French, German and Australian market over 2015 and 2016 and show in particular the effect of spike modelling on the valuation of certain strip options. In particular, we show that some out-of-the-money options have a significant value if we incorporate spikes in our modelling, while having a value close to 0 otherwise.

Mathematics Subject Classification (2010): 62M86, 60J75, 60G35, 60F05.

Keywords: Financial statistics, Discrete observations, Electricity market modelling, Derivatives pricing.

1. INTRODUCTION

1.1. Motivation. A striking empirical feature of electricity spot prices is the presence of spikes, that can be described by a jump in the price process immediately followed by a fast mean reversion (see Figure 1 showing the behaviour of electricity spot prices in different markets over one year of historical data). These spikes are due to the non-storability of electricity, an abrupt change in the demand or the offer (due to weather conditions, outages and so on) having a direct impact on prices. For risk management purposes, the modelling of these extreme events is essential. And, due to the non-storability of electricity, the modelling of forward contracts (used as hedging products) are also needed. If $(S_t)_{t \geq 0}$ denotes the electricity spot price, the forward price $f(t, T)$ at time t delivering 1 megawatt hour (MWh) at time T can be defined as

$$(1) \quad f(t, T) = \mathbb{E}[S_T | \mathcal{F}_t], \quad t \geq 0$$

where \mathcal{F}_t is the available information up to time t and the expectation is taken under a risk-neutral probability. In this context, one usually faces two major issues: first and prior to data analysis, a stochastic model that captures the main characteristics of spot prices, and especially the presence of fast mean-reverting spikes has to be set, however simple enough to give tractable formulas for the forward prices $f(t, T)$. Second, the chosen model must be calibrated with efficient statistical procedures to show its adequacy to the data, and to properly quantify risk measures. The main difficulty is the estimation of the characteristics of the spikes. To this end, Cartea and Figueroa [8] extend the popular and tractable approach of Lucia *et al.* [16] by introducing jumps in the price process, resulting in the model

$$\log S_t = \rho(t) + Y_t, \quad dY_t = -\beta Y_t dt + \sigma(t) dW_t + \log J dN_t, \quad t \geq 0,$$

where $\rho(t)$ and $\sigma(t)$ are deterministic components, (W_t) is a Wiener process, (N_t) is a Poisson process and J is the jump size drawn proportional to a log-normal distribution. A similar model is proposed in Geman and Roncoroni [12], adding up a threshold parameter that determines the sign of the jumps.

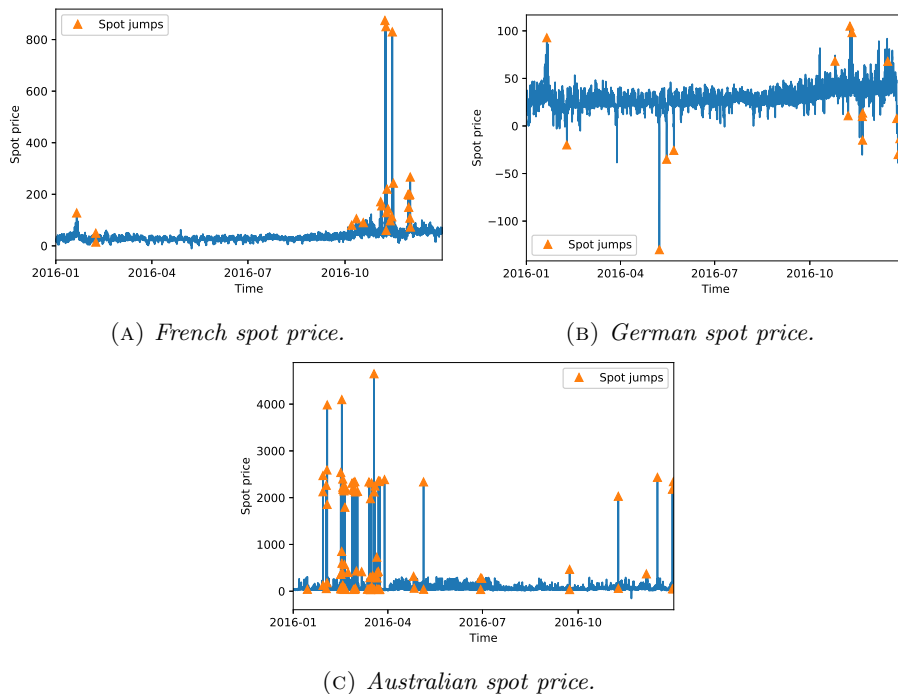


FIGURE 1. Series of the spot price during the year 2016 for France, Germany and Australia. The frequency of the data is 1 hour for France and Germany and 30 minutes for Australia. Spot jumps are estimated using a threshold of $5\hat{\sigma}\Delta_n^{-0.01}$ where $\hat{\sigma}$ is the multi-power variation of order 20.

In these approaches, the mean reverting coefficient $\beta > 0$ is the same for the continuous component and for the spike component. However, statistical evidence shows that the mean reversion of the spike component is much stronger than the one of the Brownian component, see for instance Benth *et al.* [6]. The estimated β then underestimates the mean reversion of the spike component and overestimates the one of the continuous component. A similar model slightly more realistic is also proposed by Geman and Roncoroni [12] but this one does not provide explicit formulas for deriving $f(t, T)$. Yet another approach is undertaken in Benth *et al.* [5, 6] with multi-factor models:

$$S_t = \sum_{i=1}^m w_i Y_t^i, \quad dY_t^i = -\beta_i Y_t^i dt + dL_t^i, \quad t \geq 0, \quad i = 1, \dots, m,$$

for some weights w_i , and where $(L_t^i)_{t \geq 0}$ are independent time-inhomogeneous subordinators ensuring that $(S_t)_{t \geq 0}$ remains nonnegative. Benth and collaborators establish in [5, 6] that $m = 2$ is sufficient for modelling purposes, each factor $(Y_t^i)_{t \geq 0}$ having its own mean reverting parameter, allowing for a fast mean reversion and a slower one. However, the use of subordinators implies

that the volatility of the process seems to be underestimated. Finally, multi-factor models with a Brownian component and a spike component are studied in Meyer and Tankov [19], Schmidt [23] and Gonzales *et al.* [13]. Meyer and Tankov estimate the mean-reverting parameters using spectral methods and the jumps are detected by filtering. In Schmidt [23], the parameters of the model are estimated using maximum likelihood with the EM algorithm, implying an approximation of the process with its Euler scheme. Gonzales and co-authors develop a Bayesian framework and recover the parameters of the model by MCMC. In a more general context than electricity price modelling, Moreno *et al.* [20] use a method of moments to estimate the parameters of a jump diffusion model when the log-price is the sum of an arithmetic Brownian motion and a mean reverting compound Poisson process.

In this paper, we mostly focus on the estimation of the characteristics of the spike process in the case of a wide and unifying range of models, including the aforementioned ones and allowing us to get analytical formulas of forward prices. More precisely, the goal of the paper is at least threefold:

- i) Introduce a general electricity spot price model consistent with historical data that encompasses the previous approaches and overcomes their limitations.
- ii) Develop within this model efficient and robust statistical procedures that estimate the spikes' characteristics.
- iii) Derive an explicit correction formula for the value $f(t, T)$ of forward contracts revealing the effect of spikes along maturities.

Therefore, our proposed approach can be easily used in practice, solving both estimation and pricing issues.

1.2. Main results. We consider an extended framework that encompasses [19], [23] and [13]. In particular, our approach does not require that the continuous part of the price process is an Ornstein-Uhlenbeck, a necessary condition in the aforementioned models.

A semimartingale model with fast mean-reverting jumps. On a rich enough filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ that will accommodate all the considered random quantities, we model the electricity spot price $X_t = S_t$ or $X_t = \log S_t$ by

$$(2) \quad X_t = X_t^c + Z_t^\beta, \quad t \geq 0,$$

where $(X_t^c)_{t \geq 0}$ is a continuous Itô semimartingale and $(Z_t^\beta)_{t \geq 0}$ is the so-called *spike process*, governed by a mean-reverting factor $\beta > 0$. More specifically, we assume that

$$(3) \quad X_t^c = X_0^c + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s, \quad t \geq 0$$

where $(\sigma_t)_{t \geq 0}$ and $(\mu_t)_{t \geq 0}$ are two adapted càdlàg processes, $(W_t)_{t \geq 0}$ a (\mathcal{F}_t) -standard Brownian motion and

$$(4) \quad Z_t^\beta = \int_0^t \int_{\mathbb{R}} x e^{-\beta(t-s)} \underline{p}(ds, dx), \quad t \geq 0,$$

with \underline{p} a random Poisson measure on $[0, \infty) \times \mathbb{R}$ independent of $(W_t)_{t \geq 0}$, with intensity

$$\underline{q} = \lambda dt \otimes \nu(dx),$$

for some $\lambda > 0$ and a probability measure $\nu(dx)$ on \mathbb{R} . We thus model the electricity spot price as a classical continuous Itô semimartingale $(X_t^c)_{t \geq 0}$ allowing for the usual financial fluctuations and usual models (factor models, mean-reverting models and so on) to which we add a perturbation

$(Z_t^\beta)_{t \geq 0}$ of “spikes” or “jumps”, triggered by exogenous physical hazard, at intensity λ and magnitude $\nu(dx)$, but with a relaxation period $1/\beta$ comparable to λ that accounts for the absorption of such events by the market toward resulting in stable prices at large scales. The term *comparable* is a bit vague at this stage, and will be assessed precisely in Section 2.1, enabling us to speak of *fast mean-reversion*. In this setting, model (2)-(3)-(4) is well posed and can reproduce, at least *visually*, the general shape of electricity spot markets, compare historical data from Figure 1 and sample paths simulations given in Figure 2 and detailed in the simulation Section 3.2.

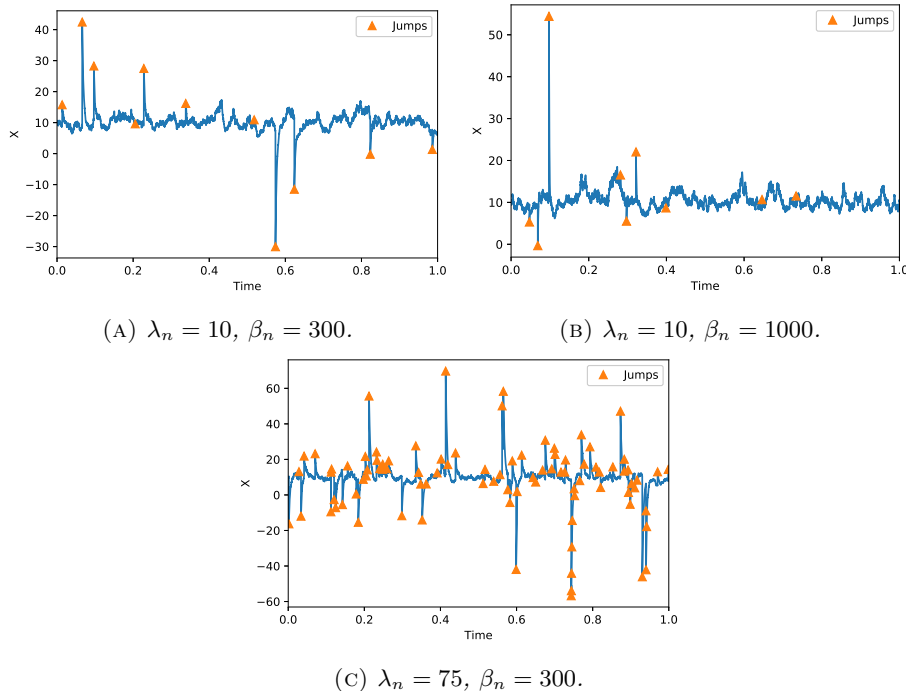


FIGURE 2. Simulation of the process X in the case of a model having continuous part defined in (12) and with jump sizes having law $0.4(-\mathcal{E}(40)) + 0.6\mathcal{E}(30)$ for different values of λ_n and β_n .

Statistical setting. We assume that we observe the process $(X_t)_{t \geq 0}$ given by (2)-(3)-(4) over the time interval $[0, T]$ on a regular grid

$$0 = t_{0,n} < t_{1,n} < \dots, t_{n,n} = T, \quad t_{i,n} = i\Delta_n, \quad \text{for } 0 \leq i \leq n,$$

with mesh Δ_n . Thus we have n (or rather $n + 1$) observations

$$(5) \quad X^n = (X_0, X_{\Delta_n}, \dots, X_{n\Delta_n} = X_T).$$

In the following, for a given process Y , we use the classical notations $\Delta_i^n Y = Y_{t_{i,n}} - Y_{t_{i-1,n}}$ and $\Delta Y_s = Y_s - Y_{s-}$.

Asymptotics are taken as $n \rightarrow \infty$. We assume that T is constant, and we take $T = 1$ with no loss of generality. Equivalently, $\Delta_n = 1/n \rightarrow 0$ as $n \rightarrow \infty$. This asymptotic setting is usually referred to as the “high-frequency” framework (for instance the classical textbook [3] by Aït-Sahalia and

Jacod), but this terminology is a bit misleading: our framework certainly belongs to statistical finance, but it has no link to high-frequency finance or microstructure modelling of any sort. In practice, we apply our methodology to three markets: the French EPEX, the German EPEX and the Australian electricity spot in Queensland, see Section 3 below. We use data between 2015, Jan. 01 and 2016, Dec. 31. with hourly data (even less in the case of Australian data), so that $n = 17064$ is considered to be large. Equivalently, one hour is considered to be small in front of 2 years. In our setting, the important fact about the assumption that T is fixed is that we leave out any stationarity or ergodicity of the underlying process. We thus make an implicit statistical robustness assumption, which we believe is of importance when considering recent and changing energy markets over such time horizons.

The parameters of interest are $\lambda, \beta > 0$ that govern our correction formulas (see the application to forward contracts prices $f(t, T)$ below). In particular we leave out the issue of identifying the continuous semimartingale part $(X_t^c)_{t \geq 0}$ *i.e.* the drift $(\mu_t)_{t \geq 0}$ and the volatility process $(\sigma_t)_{t \geq 0}$ as well as the jump distribution $\nu(dx)$.

The mean-reversion factor over the observation increment $[t_{i-1,n}, t_{i,n}]$ is of size $\beta \Delta_n$, and by requiring $\beta \Delta_n$ to be large compared to the order of magnitude $\sqrt{\Delta_n}$ of $\Delta_t^n X^c$, we may hope to recover β asymptotically. We thus introduce the asymptotic setting $\beta = \beta_n$ with the requirement

$$(6) \quad \beta = \beta_n \rightarrow \infty \text{ while } \beta_n \sqrt{\Delta_n} \rightarrow \infty$$

Condition (6) becomes

$$(7) \quad \beta_n \sqrt{\Delta_n} \lambda_n \rightarrow \infty$$

if we let $\lambda = \lambda_n \rightarrow \infty$, another necessary condition that enough jumps are available over the observation period $[0, T]$. A second crucial assumption is

$$(8) \quad \beta_n \Delta_n \lesssim 1,$$

since otherwise, the spikes caused by the jumps of \underline{p} are absorbed by the Brownian fluctuations of X^c due to the fast relaxation period $1/\beta_n$ and therefore cannot be detected by X^n .

Statistical results. Heavily relying on classical techniques in high-frequency finance (for instance [3, Theorem 10.26, p.374]), we estimate in a first step the times and sizes of the jumps which are random quantities, taking into account the interplay between β_n , λ_n and Δ_n dictated by the asymptotic regime (6)-(7)-(8), see Proposition 4. In a second step, we construct an estimator of β_n based on an estimator \widehat{s}_n of the right-derivative or *instantaneous slope* of $t \mapsto Z_t^\beta$ right after a jump is detected. The estimator \widehat{s}_n is based on averaging of instantaneous slope proxies of the form $\Delta X_{T_q}(1 - e^{-\beta_n \Delta_n})$ that govern the relaxation effect after a jump of size ΔX_{T_q} has occurred at time T_q and it enables us to consider

$$\widehat{\beta}_n = -\frac{1}{\Delta_n} \log(1 - \widehat{s}_n)$$

as our estimator of β_n . Since β_n itself varies with n and grows to infinity, the notion of convergence has to be considered carefully. Under suitable assumptions, we prove in Theorem 6 that the relative error

$$(9) \quad \mathcal{E}_n = \frac{\widehat{\beta}_n - \beta_n}{\beta_n} \rightarrow 0$$

in probability as $n \rightarrow \infty$. The error \mathcal{E}_n has two components: a first term of order $1/(\beta_n \sqrt{\lambda_n \Delta_n})$ due to Brownian oscillations, and a second term of order $\min\{\lambda_n/\beta_n, 1/\sqrt{\beta_n}\} + \sqrt{\lambda_n}/\beta_n$ that accounts for the effect of jumps that are still present in the price process despite the relaxation effect. When $\beta_n \sqrt{\Delta_n \lambda_n} \rightarrow \infty$ and $\lambda_n/\beta_n \lesssim 1$, we have \mathcal{E}_n converges to 0. If we assume further $\sqrt{\beta_n}/\lambda_n \rightarrow 0$, we obtain a central limit theorem for \mathcal{E}_n with a Gaussian limit and an explicit rate of convergence that depends on the interplay between λ_n, β_n and Δ_n . We have an analogous result (although less demanding) for the estimation of the jump intensity λ_n detailed in Proposition 5.

Application to pricing forward contracts. We show in Theorem 7 that in the model for the spot price (with $X = S$) defined by (2)-(3)-(4), the price $f(t, T)$ of a forward contract is given by

$$f(t, T) = f^c(t, T) + f^\beta(t, T),$$

where

$$f^c(t, T) = \mathbb{E}[X_T^c | \mathcal{F}_t]$$

and

$$f^\beta(t, T) = e^{-\beta(T-t)} Z_t^\beta + \frac{\lambda}{\beta} \int_{\mathbb{R}} x \nu(dx) \left(1 - e^{-\beta(T-t)}\right).$$

The term $f^c(t, T)$ corresponds to the price of the forward contract in a continuous case framework. The computation of this value has been extensively studied for different continuous models and it is known analytically for the most common models, see for instance [5, 6] among others. The term $f^\beta(t, T)$ is a correction that follows from our approach. It is of order λ/β and is usually small for the applications we have in mind, see the practical implementation Sections 3 and 4. On balance, the presence of spikes does not significantly impact the price of forward contracts to within these order of magnitudes. This is consistent with our data, for which spikes are not observed on forward prices. By neglecting the term $f^\beta(t, T)$, we can calibrate the process X_t^c to the observed forward prices and the process Z_t^β to the observed spot prices with the estimation procedure proposed in this paper. The results are similar for the modelling of the logarithm of the spot price (*i.e* when $X = \log(S)$) by (2)-(3)-(4), see Theorem 8. We implement the prices of the forward contracts and Call options from a model calibrated to historical data on electricity prices in Section 4 and we show the impact of the spike modelling on the valuation of a strip of options with payoff of the form $\sum_{i=1}^I (S_{t_{i,n}} - K)^+$ for different times $t_{i,n}$. As expected, the value of this option increases if we add significant large spikes. In particular, we show that some out-of-the-money options have a significant value if we incorporate spikes in our modelling, while having a value close to 0 otherwise.

1.3. Organisation of the paper. Section 2 develops a rigorous mathematical framework for the stochastic model (2)-(3)-(4) and gives the explicit construction of the estimators described in Section 1.2 above together with their asymptotic properties in Propositions 4, 5 and Theorem 6. Section 3 establishes the numerical feasibility and consistency of our statistical estimation results on simulated and real data, based over two years (2015 and 2016) of electricity spot prices in three different markets (French, German and Australian). Section 4 is devoted to the application of our model and statistical results to forward contracts. We establish in Theorem 7 a correction formula to get analytical forward prices and study the valuation of a strip of European Call options. The proofs are given in Section 5.

2. STATISTICAL RESULTS

2.1. Model assumptions. We consider the process $(X_t)_{t \geq 0}$ defined by (2)-(3)-(4) in Section 1.2. Following closely the standard notation of Aït-Sahalia and Jacod [3], if ν is a positive measure and

f a ν -integrable function, we write $f(x) \star \nu = \int_{\mathbb{R}} f(x) \nu(dx)$. Remember also that we work over a finite time horizon $T = 1$.

Assumption 1. *We have $\mathbb{E}[(X_t^c)^2] < \infty$ for every $t \geq 0$. Moreover, $t \mapsto \sigma_t$ is continuous on $[0, 1]$ and for some deterministic $\underline{\sigma}, \bar{\sigma}, c_0 > 0$, we have $0 < \underline{\sigma}^2 \leq \inf_t \sigma_t^2 \leq \sup_t \sigma_t^2 \leq \bar{\sigma}^2$, $\sup_t |\mu_t| \leq c_0$, $\nu(\{0\}) = 0$ and $|x|^2 \star \nu < \infty$.*

Since our asymptotic results will be given in distribution (see Theorem 6 below), the conditions on the drift $(\mu_t)_{t \geq 0}$ and $(\sigma_t)_{t \geq 0}$ can substantially be weakened (in order to accommodate for instance diffusion coefficients of the form $\sigma_t = X_t^c h(X_t^c)$ with a bounded h or even locally integrable) by standard localisation procedures, see for instance [14, Section 4.4.1]. We observe (5) and asymptotics are taken as $n \rightarrow \infty$ or equivalently $\Delta_n = n^{-1} \rightarrow 0$. We also allow $\lambda = \lambda_n$ and $\beta = \beta_n$ to either grow to ∞ with n or remain bounded, compare Equations (6)-(7)-(8) and the accompanying discussion in Section 1.2 above.

Assumption 2. *We have*

$$\lambda_n \lesssim \beta_n, \quad \beta_n \Delta_n \lesssim 1, \quad \text{and} \quad \lambda_n \Delta_n \rightarrow 0.$$

The condition $\lambda_n \lesssim \beta_n$ ensures the stability of X_t as $n \rightarrow \infty$ since $\text{Var}(X_t) = \text{Var}(X_t^c) + \text{Var}(Z_t^{\beta_n}) = \text{Var}(X_t^c) + |x|^2 \star \nu \frac{\lambda_n t}{2\beta_n} (1 - e^{-2\beta_n}) \rightarrow \infty$ if $\sup_n \lambda_n / \beta_n = \infty$. The condition $\beta_n \Delta_n \lesssim 1$ is necessary to identify spikes (or jumps): otherwise, a spike that occurs in the interval $((i-1)\Delta_n, i\Delta_n]$ will be absorbed by the relaxation effect before we observe $X_{i\Delta_n}$. Finally, the condition $\lambda_n \Delta_n \rightarrow 0$ controls the no accumulation of jumps within the rate of observation.

In order to estimate the times and sizes of the jumps, we need the following:

Assumption 3. *We have either (I) or (II), where*

(I) *For some $\varpi \in (0, 1/2)$:*

(i) $\lambda_n^2 \Delta_n \rightarrow 0$, (ii) $\lambda_n \mathbf{1}_{\{|x| > \Delta_n^{1/2-\varpi}/(\beta_n \Delta_n)\}} \star \nu \rightarrow 0$, and (iii) $\lambda_n \mathbf{1}_{\{|x| < \Delta_n^{1/2-\varpi}\}} \star \nu \rightarrow 0$.

(II) *For some $\varpi \in (0, 1/2)$ and a sequence of integers $k_n \geq 1$:*

(i) $\lambda_n^2 \Delta_n k_n^2 \rightarrow 0$, (ii) $\lambda_n \mathbf{1}_{\{|x| > e^{\beta_n \Delta_n k_n} \Delta_n^{1/2-\varpi}\}} \star \nu \rightarrow 0$, (iii) $\lambda_n \mathbf{1}_{\{|x| < \Delta_n^{1/2-\varpi}\}} \star \nu \rightarrow 0$

and (iv) $\lambda_n e^{-(\beta_n \Delta_n^{1-\varpi})^2} \rightarrow 0$.

Assumption 3 (I) implies $\beta_n \Delta_n = o(1)$. Condition (i) ensures that the number of jumps in a interval of size Δ_n is essentially 1. In the case where β_n is bounded, (ii) is implied by (i) and we have the usual conditions for the detection of jumps (see Mancini [17]). Condition (ii) controls the size of the mean-reversion. Condition (iii) controls the size of the small jumps that cannot converge too fast to 0. If the jumps are bounded below by some constant as Mancini [17] the condition is automatically satisfied.

Assumption 3 (II) (iv) implies that $\beta_n \Delta_n^{1-\varpi} \rightarrow \infty$ and in particular $\beta_n \Delta_n^{1/2} \rightarrow \infty$, implying that the mean reversion of order $\beta_n \Delta_n$ is stronger than the order of magnitude $\Delta_n^{1/2}$ of Brownian increments. It also allows for the case $\beta_n \Delta_n \approx 1$. In the setting of Assumption 3 (II), the mean reversion is more difficult to distinguish from the jumps and in the case $\beta_n \Delta_n \approx 1$, the jumps and the drift have the same size and are not distinguishable from their size solely. Condition (i) states that there is at most one jump in an interval of size k_n which is large enough for the spike to vanish. Assumption 3 (II) also implies that $\lambda_n^2 / \beta_n \rightarrow 0$. Assumption 3 (II) allows for high values of β_n but the number of jumps needs then to be smaller than in case Assumption 3 (I) compared to β_n .

2.2. Estimation of the jumps times and λ_n . We first construct estimators of the sequence of the jumps and of their intensity λ_n . Assumption 3 is in force. Let

$$N_t = \sum_{s \leq t} \mathbf{1}_{\{\Delta X_s \neq 0\}}, \quad t \geq 0,$$

denote the number of jumps of $(X_t)_{t \geq 0}$ up to time t and let $T_1 < T_2 < \dots < T_q < \dots$ denote the random times at which jumps occur. By construction, the sizes of jumps $(\Delta X_{T_q})_{q \geq 1}$ form a sequence of independent and identically distributed random variables independent of $(N_t)_{t \geq 0}$. Let $i_{n,q}$ be the random integer such that

$$(i(n, q) - 1) \Delta_n < T_q \leq i(n, q) \Delta_n.$$

Define the increasing sequence

$$\mathcal{I}_n(1) < \dots < \mathcal{I}_n(\widehat{\lambda}_n)$$

of indices $i \in \{1, \dots, n\}$ defined by the realisation of the following successive events:

$$\left\{ \frac{|\Delta_i^n X|}{\sqrt{\Delta_n}} > v_n \right\} \text{ under Assumption 3 (I)}$$

and

$$\left\{ \frac{|\Delta_i^n X|}{\sqrt{\Delta_n}} > v_n, \Delta_i^n X \Delta_{i+1}^n X < 0 \right\} \text{ under Assumption 3 (II).}$$

with $v_n \asymp \Delta_n^{-\varpi}$, following the notations of [3, Ch. 10].

Under Assumption 3 (II), we need the supplementary condition $\Delta_i^n X \Delta_{i+1}^n X < 0$ for the following reason: whenever a jump occurs, the mean reverting is dominant in the next observation interval and has a direction opposite to the sign of the jump. Furthermore, it enables us to discard the increments caused by the mean reversion that are large enough to be detected as jumps. Indeed, if we detect a false jump due to the mean reversion effect, the next increments will follow the same dynamics and it will share the same sign with first increment. The property that no jump lies within the next observation interval is ensured by the existence of k_n . Let

$$\Omega_n = \left\{ \widehat{\lambda}_n = N_1, \forall q \in \{1, \dots, N_1\} : T_q \in (\mathcal{I}_n(q) \Delta_n - \Delta_n, \mathcal{I}_n(q) \Delta_n) \right\}.$$

Proposition 4. *Work under Assumptions 1, 2 and 3. We have $\mathbb{P}(\Omega_n) \rightarrow 1$.*

The proof of Proposition 4 relies on a result of Aït-Sahalia and Jacod [3, Theorem 10.26, p.374]. However, the presence of a drift term $-\beta_n \int_0^t Z_s^{\beta_n} ds$ that depends on n together with the fact that $\lambda_n \rightarrow \infty$ makes the extension not trivial. Proposition 4 also provides us with an estimator $\widehat{\lambda}_n$ of λ_n . In the case $\lambda_n \rightarrow \infty$, we have the following asymptotic property:

Proposition 5. *Work under Assumption 1, 2 and 3 and assume that $\lambda_n \rightarrow \infty$. We have*

$$(10) \quad \sqrt{\lambda_n} \frac{\widehat{\lambda}_n - \lambda_n}{\lambda_n} \rightarrow \mathcal{N}(0, 1)$$

in distribution as $\lambda_n \rightarrow \infty$.

This result is straightforward and asymptotically optimal: consider indeed the seemingly richer experiment where one continuously observes a Poisson process $(\mathcal{P}_t)_{0 \leq t \leq 1}$ with intensity $\lambda > 0$. The variable \mathcal{P}_1 is a sufficient statistic and the Cramer-Rao bound tells us that any unbiased estimator $\widehat{\lambda}$ necessarily satisfies $\mathbb{E}[(\widehat{\lambda} - \lambda)^2] \geq I(\lambda)^{-1}$, where $I(\lambda) = 1 + \lambda^{-1}$ is the Fisher information associated to the observation of \mathcal{P}_1 , *i.e.* a Poisson random variable with parameter λ . Equivalently $\mathbb{E}[(\frac{\widehat{\lambda} - \lambda}{\lambda})^2] \geq (\lambda^{-1} + \lambda^{-2}) \sim \lambda^{-1}$ as $\lambda \rightarrow \infty$ which is consistent with the convergence (10).

A natural estimator of the jump sizes is $\Delta_{\mathcal{I}_n(q)}^n X$ for $q \in \{1, \dots, \widehat{\lambda}_n\}$, see [3, Theorem 10.21, p.370]. In our case, $\Delta_{\mathcal{I}_n(q)}^n X$ is equal to $\Delta X_{T_q} e^{-\beta_n(T_q - \mathcal{I}_n(q)\Delta_n)}$ plus a negligible term. If $\beta_n \Delta_n \rightarrow 0$, $\Delta_{\mathcal{I}_n(q)}^n X$ is then equivalent to ΔX_{T_q} but if $\beta_n \Delta_n \asymp 1$, the bias $\Delta X_{T_q} (1 - e^{-\beta_n(T_q - \mathcal{I}_n(q)\Delta_n)})$ remains and it is not possible to identify the size of the jump. However, if $\lambda_n \rightarrow \infty$, one can infer some statistical properties of the jumps size. Indeed, $(T_q - \mathcal{I}_n(q)\Delta_n)_{1 \leq q \leq \widehat{\lambda}_n}$ has a known distribution and this error can be averaged. We can easily prove the following result:

$$\frac{m\beta_n\Delta_n}{(1 - e^{-m\beta_n\Delta_n})\widehat{\lambda}_n} \sum_{q=1}^{\widehat{\lambda}_n} (\Delta_{\mathcal{I}_n(q)}^n X)^m \rightarrow x^m \star \nu$$

in probability for every integer $m > 0$ such that $x^m \star \nu < \infty$. The proof for the case $m = 1$ appears in the proof of Theorem 6. Combining this result with our estimator of β_n provided below enables us to have an estimator for the moments of ν .

2.3. Estimation of β_n . We are ready to construct an estimator of β_n . Define the sign function as $\text{sgn}(x) = 1$ if $x \geq 0$ and -1 otherwise. On the event $\{\widehat{\lambda}_n > 0\}$, define $\widehat{\beta}_n$ via

$$(11) \quad \exp(-\Delta_n \widehat{\beta}_n) = \max \left\{ 1 + \frac{\sum_{q=1}^{\widehat{\lambda}_n} \text{sgn}(\Delta_{\mathcal{I}_n(q)}^n X) (\Delta_{\mathcal{I}_n(q)+1}^n X + 2\Delta_n \sum_{j=1}^{q-1} \Delta_{\mathcal{I}_n(j)}^n X)}{\sum_{q=1}^{\widehat{\lambda}_n} |\Delta_{\mathcal{I}_n(q)}^n X|}, \Delta_n \right\}$$

and set $\widehat{\beta}_n = 0$ otherwise. Our main result describes precisely the behaviour of $\widehat{\beta}_n$ under the different asymptotic regimes of interest.

Theorem 6. *Work under Assumptions 1, 2 and 3. Let $\beta_n \sqrt{\lambda_n \Delta_n} \rightarrow \infty$. On the set $\{\widehat{\lambda}_n > 0\}$, we have*

$$\frac{|\widehat{\beta}_n - \beta_n|}{\beta_n} \lesssim \lambda_n \Delta_n + \frac{1}{\beta_n \sqrt{\lambda_n \Delta_n}} + \min \left\{ \frac{1}{\sqrt{\beta_n}}, \frac{\lambda_n}{\beta_n} \right\}$$

in probability. More precisely,

i) On $\{\widehat{\lambda}_n > 0\}$, we have

$$\frac{\widehat{\beta}_n - \beta_n}{\beta_n} = \mathcal{M}_n + \mathcal{V}_n \mathcal{J}_n^T,$$

where

$$\mathcal{M}_n = e^{\beta_n \Delta_n} \frac{\lambda_n}{\beta_n} \frac{(x \star \nu) (\text{sgn}(x) \star \nu)}{|x| \star \nu} \left(\frac{e^{\beta_n \Delta_n} - 1}{\beta_n \Delta_n} - 1 \right),$$

$\mathcal{V}_n = (\mathcal{V}_n^{(i)})_{1 \leq i \leq 4} \in \mathbb{R}^4$ is such that

$$\left\{ \begin{array}{l} \mathcal{V}_n^{(1)} = e^{\beta_n \Delta_n} \frac{\sqrt{\lambda_n}}{\sqrt{3\beta_n} |x| \star \nu} ((\text{sgn}(x) \star \nu)^2 (|x|^2 \star \nu) + (x \star \nu)^2 - 2(\text{sgn}(x) \star \nu)(|x|^2 \star \nu))^{1/2}, \\ \mathcal{V}_n^{(2)} = e^{\beta_n \Delta_n} \min \left\{ \left(\frac{|x|^2 \star \nu}{(|x| \star \nu)^2} \frac{1}{2\beta_n} \frac{(1 - e^{-2\beta_n \Delta_n})}{2\beta_n \Delta_n} \right)^{1/2}, \frac{\lambda_n}{\beta_n} \right\}, \\ \mathcal{V}_n^{(3)} = e^{\beta_n \Delta_n} \left(\frac{\beta_n \Delta_n}{1 - e^{-\beta_n \Delta_n}} \right) \frac{\sqrt{\int_0^1 \sigma_s^2 ds}}{|x| \star \nu \sqrt{\lambda_n \beta_n \Delta_n}}, \\ \mathcal{V}_n^{(4)} = e^{\beta_n \Delta_n} \frac{\sqrt{\int_0^1 \sigma_s^2 ds} \sqrt{\Delta_n}}{|x| \star \nu \sqrt{\lambda_n}}, \end{array} \right.$$

and $\mathcal{J}_n = (\mathcal{J}_n^{(i)})_{1 \leq i \leq 4} \in \mathbb{R}^4$ is bounded in probability as $n \rightarrow \infty$.

ii) If $\lambda_n \rightarrow \infty$, then

$$(\mathcal{J}_n^{(3)}, \mathcal{J}_n^{(4)}) \rightarrow \mathcal{N}(0, \text{Id}_{\mathbb{R}^2})$$

in distribution as $n \rightarrow \infty$.

iii) If $\lambda_n \rightarrow \infty$, $|x|^3 \star \nu < \infty$ and $(\text{sgn}(x) \star \nu)^2 |x|^2 \star \nu + (x \star \nu)^2 - 2\text{sgn}(x) \star \nu |x|^2 \star \nu \neq 0$, we have

$$(\mathcal{J}_n^{(1)}, \mathcal{J}_n^{(3)}, \mathcal{J}_n^{(4)}) \rightarrow \mathcal{N}(0, \text{Id}_{\mathbb{R}^3})$$

in distribution as $n \rightarrow \infty$.

iv) If $\beta_n/\lambda_n^2 \rightarrow 0$, the conditions of iii) and $|x|^4 \star \nu < \infty$ hold together, we finally obtain

$$\mathcal{J}_n \rightarrow \mathcal{N}(0, \text{Id}_{\mathbb{R}^4})$$

in distribution as $n \rightarrow \infty$.

Some remarks on the different error terms: **1)** the term of order $1/(\beta_n \sqrt{\lambda_n \Delta_n})$ accounts for the presence of a Brownian motion in the term $(X_t^c)_{t \geq 0}$. When λ_n is bounded, we need $\beta_n \sqrt{\Delta_n} \rightarrow \infty$ or equivalently $\sqrt{\Delta_n} = o(\beta_n \Delta_n)$: the size of the slope of $(Z_t^\beta)_{t \geq 0}$ after a jump needs to dominate the Brownian motion part which is of order $\sqrt{\Delta_n}$. In the case where $\lambda_n \rightarrow \infty$, we can average the error due to the Brownian martingale part and then diminish the order of the error. In that case, we do not need the restriction $\sqrt{\Delta_n} = o(\beta_n \Delta_n)$ anymore but rather $\sqrt{\Delta_n/\lambda_n} = o(\beta_n \Delta_n)$. **2)** The error terms of order $\min\{\frac{1}{\sqrt{\beta_n}}, \frac{\lambda_n}{\beta_n}\}$, $\frac{\sqrt{\lambda_n}}{\beta_n}$ and $\lambda_n \Delta_n$ account for the jumps that occur before the observation increment used to estimate the slope of the process. **3)** The term $2\Delta_n \sum_{j=1}^{q-1} \Delta_{\mathcal{I}_n(j)}^n X$ introduced in the definition of $\hat{\beta}_n$ in (11) is a bias correction that enables us to obtain a consistent estimator in the case $\lambda_n/\beta_n \approx 1$.

3. PRACTICAL IMPLEMENTATION

3.1. Choice of the threshold v_n . The method to detect the jumps is based on the classical use of the threshold which is proportional to $\Delta_n^{-\varpi}$. As for the choice of the threshold and ϖ , no exact method is provided in the literature. However, the threshold is recommended to be chosen of the form $v_n = C\hat{\sigma}\Delta_n^{-\varpi}$ in [2, Section 5.3] and [3, Section 6.2.2, p. 187] where C is a constant and $\hat{\sigma}$ is an estimator of the integrated volatility $\int_0^1 \sigma_s^2 ds$. A popular rule-of-thumb consists in picking ϖ close to 0. Moreover, [2] suggests to choose C between 3 and 5.

It remains to find an estimator of the integrated volatility. A natural choice is the multipower variation estimator, see [4] and [25] for more details. The order of the multipower variation estimator is set to 20 in the practical applications in this paper, which is high compared to the orders typically chosen in the literature. This choice is justified by the strong mean reversion of the spikes. As for the jumps, spikes have large increments that need to be compensated in the multipower variation estimator and can be present during two or three time steps. We compensate these large increments with a higher order of the multipower variation. Some simulations on simple models show that an order of 20 looks reasonable.

3.2. Numerical illustration. In this section, we study the performances of our estimation procedures on simulated data of the process defined by (2)-(3)-(4). We tested a wide range of values for (λ_n, β_n) in order to illustrate the results of Theorem 6. To be consistent with real data, the process is simulated with a step time $\Delta_n = 10^{-4}$, which is the order of magnitude corresponding to one year of hourly data observations. We pick

$$(12) \quad dX_t^c = X_t^c \left(\left(\frac{1}{2} 2^2 - 100 \log(X_t^c) \right) dt + 2dW_t \right)$$

corresponding to the logarithm of an Ornstein-Uhlenbeck process with the mean reverting parameter equal to 100 and volatility parameter equal to 2. The sizes of the jumps follow the law $0.4(-\mathcal{E}(15)) + 0.6\mathcal{E}(10)$, with $\mathcal{E}(\rho)$ denoting the exponential distribution with parameter $\rho > 0$. Figure 2 illustrates a sample paths of the process for different parameters λ_n and β_n . We realize 10000 simulations. We use jump detection under Assumption 3(I) as well as under Assumption 3(II), the latter corresponding to filtering the previous jumps by keeping only increments that have successive opposite signs.

Table 1, 2 and 3 show the different results of estimated $(\hat{\lambda}_n, \hat{\beta}_n)$ using a threshold equal to $C\hat{\sigma}\Delta_n^{-0.01}$, for $C = 3, 4, 5$. As expected, for large β_n , the estimation on filtered jumps gives satisfactory results of $\hat{\beta}_n$, whatever the value of λ_n satisfying the different assumptions, and seems to slightly underestimate λ_n , whereas the estimation under Assumption 3(I) gives bad results on $\hat{\beta}_n$ and overestimates λ_n . For small values of β_n , the estimation under Assumption 3(I) is more efficient for both λ_n and β_n . The results for $\beta_n = 20$ highlights, as expected, the need in Assumption 3(II) to have a small number of jumps ($\lambda_n = 10$) to get satisfactory results. Because of the expected β_n on real data, we will focus, in the next section, on the estimation procedure on filtered jumps (*i.e.* under Assumption 3(II)). This choice is also justified by the fact that underestimating the numbers of spikes seems to have a lower impact on the quality of $\hat{\beta}_n$ than overestimating them. One also observes that the choice of C has an impact on the estimation of λ_n : the threshold to select jumps increases with C and then the estimated λ_n which is the number of jumps decreases. On the contrary, it has a low impact on the estimation of β_n , except in the case $C = 3$ for small λ_n where the number of spikes is overestimated, leading to a bad estimator for β_n . Finally, the impact of the order of the multipower estimator of the volatility is illustrated in Figure 3. The impact is negligible from an order of 20, where the estimators mean of λ_n and β_n remain constant and give a good estimation. Before 20, it has only an impact on the estimation of λ_n : the spike effect has not disappeared in the estimator and $\hat{\sigma}$ is overestimated, inducing less detected spikes and an underestimation for λ_n . This supports the choice of a high order for the multipower estimator.

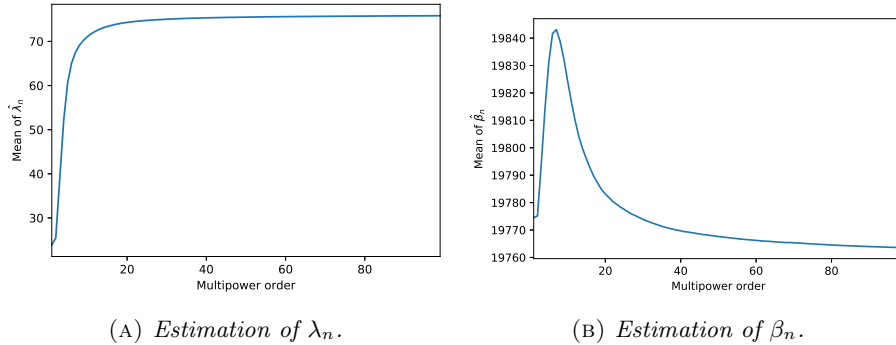


FIGURE 3. Estimation of $(\hat{\lambda}_n, \hat{\beta}_n)$ for different values of the order of the multipower estimator of the volatility $\hat{\sigma}$ under Assumption 3(II), with continuous part defined by (12), $\lambda_n = 75$, $\beta_n = 20000$ and jump size distribution $0.4(-\mathcal{E}(15)) + 0.6\mathcal{E}(10)$. The threshold v_n is chosen equal to $5\hat{\sigma}\Delta_n^{-0.01}$. The mean of the estimators is computed with 10^4 Monte-Carlo simulations and $\Delta_n = 10^{-4}$.

(λ_n, β_n)	Assumption 3 (I)				Assumption 3 (II)			
	$\hat{\lambda}_n$		$\hat{\beta}_n$		$\hat{\lambda}_n$		$\hat{\beta}_n$	
	Mean	C. I.	Mean	C. I.	Mean	C. I.	Mean	C. I.
(10, 2)	27.0	[18, 36]	4.4	[-16.6, 25.8]	13.3	[8, 20]	86.5	[13.6, 179.6]
(10, 20)	21.3	[11, 35]	-31.6	[-152.2, 23.7]	8.9	[5, 14]	37.1	[18.1, 66.8]
(10, 200)	95.1	[70, 117]	-592.6	[-1839.0, 28.7]	9.4	[5, 15]	204.5	[186.5, 225.7]
(10, 2000)	94.1	[54, 137]	-4430.1	[-5860.2, -2271.0]	9.7	[5, 15]	2003.0	[1978.9, 2026.2]
(10, 20000)	45.7	[31, 62]	-727.9	[-1262.9, -77.7]	18.0	[12, 25]	19152.0	[17840.8, 19845.5]
(75, 2)	79.6	[66, 94]	2.2	[-2.3, 6.6]	42.1	[32, 53]	42.5	[-3.5, 232.1]
(75, 20)	73.0	[59, 88]	19.7	[12.4, 27.8]	49.5	[39, 60]	50.6	[24.4, 158.9]
(75, 200)	106.7	[82, 132]	147.0	[94.5, 190.7]	63.3	[52, 75]	226.2	[198.8, 289.4]
(75, 2000)	351.3	[290, 411]	-2483.0	[-3471.0, -1566.8]	69.4	[56, 83]	2021.5	[1959.7, 2116.0]
(75, 20000)	229.2	[189, 271]	-1158.3	[-1291.5, -1016.8]	80.4	[66, 96]	19717.3	[18745.8, 20257.2]

TABLE 1. Performance of $(\hat{\lambda}_n, \hat{\beta}_n)$ for different values of (λ_n, β_n) under Assumption 3 (I) and (II), with continuous part defined by (12) and jump size distribution $0.4(-\mathcal{E}(15)) + 0.6\mathcal{E}(10)$. The threshold v_n is chosen equal to $3\hat{\sigma}\Delta_n^{-0.01}$ with $\hat{\sigma}$ is the multi-power variation estimator of order 20. The means and quantile intervals 5%-95% are computed with 10^4 Monte-Carlo simulations and $\Delta_n = 10^{-4}$.

(λ_n, β_n)	Assumption 3 (I)				Assumption 3 (II)			
	$\hat{\lambda}_n$		$\hat{\beta}_n$		$\hat{\lambda}_n$		$\hat{\beta}_n$	
	Mean	C. I.	Mean	C. I.	Mean	C. I.	Mean	C. I.
(10, 2)	10.6	[6, 16]	2.0	[-9.7, 13.7]	5.8	[2, 10]	29.7	[4.6, 46.4]
(10, 20)	10.2	[5, 16]	18.4	[2.4, 32.4]	7.6	[4, 12]	29.0	[16.8, 41.3]
(10, 200)	54.8	[39, 69]	-256.6	[-996.6, 114.1]	9.3	[5, 14]	203.8	[186.5, 224.9]
(10, 2000)	81.5	[47, 118]	-4216.3	[-5908.7, -1902.0]	9.6	[5, 15]	2002.8	[1978.9, 2025.7]
(10, 20000)	30.9	[16, 48]	-1151.6	[-1528.6, -742.1]	10.8	[5, 17]	19778.4	[19185.2, 20168.1]
(75, 2)	74.2	[61, 89]	2.0	[-2.4, 6.2]	40.2	[30, 51]	40.5	[-3.8, 221.3]
(75, 20)	72.3	[59, 87]	19.8	[12.8, 27.8]	49.2	[39, 60]	50.4	[24.4, 158.9]
(75, 200)	77.3	[61, 97]	178.8	[139.1, 212.2]	62.1	[51, 74]	225.4	[198.2, 287.8]
(75, 2000)	281.7	[231, 330]	-1888.3	[-2843.5, -1004.8]	67.4	[55, 81]	2020.8	[1959.3, 2114.6]
(75, 20000)	216.7	[177, 257]	-1200.1	[-1307.8, -1085.3]	76.6	[62, 91]	19760.4	[18778.1, 20302.9]

TABLE 2. Performance of $(\hat{\lambda}_n, \hat{\beta}_n)$ for different values of (λ_n, β_n) under Assumption 3 (I) and (II), with continuous part defined by (12) and jump size distribution $0.4(-\mathcal{E}(15)) + 0.6\mathcal{E}(10)$. The threshold v_n is chosen equal to $4\hat{\sigma}\Delta_n^{-0.01}$ with $\hat{\sigma}$ is the multi-power variation estimator of order 20. The means and quantile intervals 5%-95% are computed with 10^4 Monte-Carlo simulations and $\Delta_n = 10^{-4}$.

3.3. Practical implementation on real data. Electricity spot historical data do exhibit spikes with strong mean reversion, see Figure 1. We expect to obtain relatively high values for β_n , a necessary condition in order to apply our estimation procedure, especially under Assumption 3 (II). The goal is then to estimate the parameters λ_n and β_n of the process Z^{β_n} on a time series of spot prices, assuming that the spot price is the sum of a continuous semimartingale and a spike process. We dispose of the following data:

(λ_n, β_n)	Assumption 3 (I)				Assumption 3 (II)			
	$\hat{\lambda}_n$		$\hat{\beta}_n$		$\hat{\lambda}_n$		$\hat{\beta}_n$	
	Mean	C. I.	Mean	C. I.	Mean	C. I.	Mean	C. I.
(10, 2)	9.95	[5, 15]	2.01	[-9.21, 12.85]	5.45	[2, 9]	19.5	[4.26, 39.4]
(10, 20)	9.9	[5, 15]	19.84	[7.6, 32.4]	7.5	[3, 12]	28.8	[16.9, 40.2]
(10, 200)	33.4	[17, 47]	-76.7	[-540, 158]	9.3	[5, 14]	204	[188, 224]
(10, 2000)	72.5	[42, 105]	-3968	[-5934, -1633]	9.6	[5, 15]	2002	[1979, 2023]
(10, 20000)	29.2	[15, 46]	-1207	[-1508, -871]	10.2	[5, 16]	19861	[19337, 20207]
(75, 2)	73.7	[60, 88]	2	[-2.46, 6.17]	39.9	[30, 51]	39.2	[-4, 204]
(75, 20)	71.9	[59, 86]	19.8	[12.9, 27.6]	49.9	[39, 60]	49.9	[24.5, 155.6]
(75, 200)	68.7	[55, 83]	191	[157, 219]	61	[50, 73]	225	[198, 291]
(75, 2000)	234	[192, 273]	-1403	[-2334, -556]	65.6	[54, 78]	2019	[1958, 2109]
(75, 20000)	207	[170, 247]	-1216	[-1308, -1113]	74.2	[60, 89]	19785	[18790, 20310]

TABLE 3. Performance of $(\hat{\lambda}_n, \hat{\beta}_n)$ for different values of (λ_n, β_n) under Assumption 3 (I) and (II), with continuous part defined by (12) and jump size distribution $0.4(-\mathcal{E}(15)) + 0.6\mathcal{E}(10)$. The threshold v_n is chosen equal to $5\hat{\sigma}\Delta_n^{-0.01}$ with $\hat{\sigma}$ is the multi-power variation estimator of order 20. The means and quantile intervals 5%-95% are computed with 10^4 Monte-Carlo simulations and $\Delta_n = 10^{-4}$.

- (1) French electricity EPEX spot prices between the first of January of 2015 (included) and the first of January 2017 (not included) with data each hour ¹,
- (2) German electricity EPEX spot prices between the first of January of 2015 (included) and the first of January 2017 (not included) with data each hour ¹,
- (3) Australian electricity spot prices in Queensland between the first of January of 2015 (included) and the first of January 2017 (not included) with data each 30 minutes ².

We estimate those parameters using a threshold $v_n = C\hat{\sigma}\Delta_n^{-0.01}$, with $\hat{\sigma}$ the multi-power variation of order 20 and C a constant set to 3, 4 or 5. Results are presented in Table 4. Figure 1 gives the time series of these three sets of data with jumps time estimated in the case $C = 5$.

As expected, the estimated $\hat{\lambda}_n$ is sensitive to the value of C , the number of detected jumps decreasing with C . The estimated $\hat{\beta}_n$ is much less sensitive, although we can observe a slight increase of values with C in the French and German markets. We will see in section 4.3 the sensitivity of these estimators to the value of a strip of Call options.

Market	$C = 3$	$C = 4$	$C = 5$
French	(100, 19170)	(51, 20259)	(35, 21042.533)
German	(145, 9848)	(62, 13438)	(34, 14531)
Australian	(337, 22897)	(227, 22883)	(177, 22884)

TABLE 4. Estimation of (λ_n, β_n) for different markets using a threshold of the form $v_n = C\hat{\sigma}\Delta_n^{-0.01}$ where $\hat{\sigma}$ is the multi-power variation estimator of order 20 and C takes different values.

¹Source: <https://www.epexspot.com/>

²Source: <https://www.aemo.com.au/>

4. DERIVATIVE PRICING IN THE ELECTRICITY MARKET

4.1. Forward price formula. Due to the non-storability of electricity, the spot price is not an asset. The modelling of (and then an analytical formula for) the forward prices (i.e. the real assets and hedging products) is essential for risk management purposes. The question of the choice of the risk-neutral probability is addressed in the next section. Here, we consider that the electricity spot price $(S_t)_{t \geq 0}$ (respectively the logarithm of the spot price $(S_t)_{t \geq 0}$) is modelled by $S_t = X_t^c + Z_t^\beta$ (respectively $\log(S_t) = X_t^c + Z_t^\beta$), according to (2)-(3)-(4) under a risk-neutral probability \mathbb{Q} .

The forward price $f(t, T)$ quoted at time t , delivering 1MWh at time T , is defined by:

$$f(t, T) = \mathbb{E}^{\mathbb{Q}}[S_T | \mathcal{F}_t].$$

The available contracts in the electricity markets are of the form $f(t, T, \theta)$: a contract that delivers 1MWh continuously from T to $T + \theta$. The delivery period θ can be one week, one month, one year and so on. For example, the contract called "one-week-ahead" (1WAH) will deliver 1MWh continuously from the first hour of next Monday to the last hour of the following Sunday; the contract called "one-month-ahead" (1MAH) will deliver 1MWh continuously between the first and the last hour of next month. By classical no arbitrage arguments [5] the price of such a product is defined by:

$$f(t, T, \theta) = \frac{1}{\theta} \int_T^{T+\theta} f(t, u) du.$$

Theorems 7 and 8 give analytic formulae for the forward price $f(t, T)$ respectively for the modelling of the spot price and the modelling of the logarithm of the spot price by (2)-(3)-(4).

Theorem 7. *Suppose that the spot price is modelled by $S_t = X_t^c + Z_t^\beta$, $t \geq 0$, according to (2)-(3)-(4) under a risk-neutral probability \mathbb{Q} .*

- (1) *We have an explicit representation of $f(t, T) = \mathbb{E}^{\mathbb{Q}}[S_T | \mathcal{F}_t]$ given by*

$$f(t, T) = f^c(t, T) + f^\beta(t, T),$$

with

$$f^c(t, T) = \mathbb{E}^{\mathbb{Q}}[X_T^c | \mathcal{F}_t]$$

and

$$f^\beta(t, T) = e^{-\beta(T-t)} Z_t^\beta + \frac{\lambda x \star \nu}{\beta} (1 - e^{-\beta(T-t)}).$$

- (2) *We also have $f(t, T, \theta) = f^c(t, T, \theta) + f^\beta(t, T, \theta)$, with $f^c(t, T, \theta) = \frac{1}{\theta} \int_T^{T+\theta} f^c(t, u) du$ and*

$$f^\beta(t, T, \theta) = e^{-\beta(T-t)} \left(\frac{1 - e^{-\beta\theta}}{\beta\theta} \right) Z_t^\beta + \frac{\lambda x \star \nu}{\beta} \left[1 - e^{-\beta(T-t)} \left(\frac{1 - e^{-\beta\theta}}{\beta\theta} \right) \right].$$

The proof is straightforward. Theorem 7 has three major consequences.

- (1) The proposed model (2)-(3)-(4) allows us to get analytical formulas for the forward prices $f(t, T)$ and $f(t, T, \theta)$, provided that we have analytical formulas for the continuous part $f^c(t, T)$. This then covers a wide range of models.
- (2) We can easily write the model (2)-(3)-(4) in its equivalent form on forward prices by $f(t, T) = f^c(t, T) + f^\beta(t, T)$ with

$$(13) \quad f^\beta(t, T) = \int_0^t \int_{\mathbb{R}} x e^{-\beta(T-t)} \underline{p}(ds, dx).$$

It is then easy to consider the proposed model as an extension of any classical (continuous) model written on the forward prices, allowing to represent spikes in the spot price dynamics.

- (3) If λ/β is small and β is large, the impact of f^β on the forward prices is negligible and the additive spike process has only an impact on the spot prices. This is consistent with the observations in the electricity markets, the forward prices showing no spikes.

These consequences are of significant importance. Especially in the case where λ/β is small and β is large, this means that the spike process can be treated independently, both in parameter estimation and in simulation. Indeed, consider any existing (continuous) model describing (or simulating) $f^c(t, T)$ and calibrated on forward prices, the proposed model then consists in adding (simulations of) the spike process calibrated on spot prices following the estimation procedure previously described.

Theorem 8. *Suppose that the logarithm of the spot price is modelled by $\log(S_t) = X_t^c + Z_t^\beta$, $t \geq 0$, according to (2)-(3)-(4) under a risk-neutral probability \mathbb{Q} . Let assume that $\int_{\mathbb{R}} e^{ux} \nu(dx) < \infty$ for all $u \in [0, 1]$. We have an explicit representation of $f(t, T) = \mathbb{E}^{\mathbb{Q}}[S_T | \mathcal{F}_t]$ given by*

$$f(t, T) = f^c(t, T) f^\beta(t, T),$$

with

$$f^c(t, T) = \mathbb{E}^{\mathbb{Q}}[e^{X_T^c} | \mathcal{F}_t]$$

and

$$f^\beta(t, T) = e^{e^{-\beta(T-t)} Z_t^\beta} e^{\frac{\lambda}{\beta} \int_0^1 (\int_{\mathbb{R}} e^{ux} \nu(dx) - \int_{\mathbb{R}} e^{ue^{-\beta(T-t)x}} \nu(dx)) du}.$$

The proof is straightforward and comments for the results of Theorem 7 can be transposed to the ones of Theorem 8. In particular, if λ/β is small and β is large, the term $f^\beta(t, T)$ is close to 1 and $f(t, T)$ to $f^c(t, T)$.

4.2. Specific model and change of measure. We address the problem of choosing the risk neutral probability which we illustrate with a specific and simple model: in the rest of this section, we consider the model defined by

$$f(t, T) = \int_0^t \mu_s ds + f^c(t, T) + f^\beta(t, T),$$

with $f^\beta(t, T)$ defined by (13) and where the continuous part $f^c(t, T)$ follows the dynamics

$$df^c(t, T) = f^c(t, T) (\sigma_l dW_t^l + \sigma_s e^{-\alpha(T-t)} dW_t^s)$$

with $(W_t^l, W_t^s)_{t \geq 0}$ a two-dimensional Brownian motion under the historical probability \mathbb{P} with correlation ρ . This dynamics corresponds to a classical two factors model for forward prices of electricity [15] or gas [26]. The forward price is driven by a short term factor with volatility $\sigma_s e^{-\alpha(T-t)}$ and a long term factor with volatility σ_l . The short term volatility $\sigma_s e^{-\alpha(T-t)}$ captures the Samuelson effect: the volatility increases when $T - t$ decreases. The spot price is then equal to $S_t = \int_0^t \mu_s ds + X_t^c + Z_t^\beta$ with

$$X_t^c = f^c(0, t) \exp \left(-\frac{1}{2} \left[\sigma_l^2 t + \sigma_s^2 \frac{1 - e^{-2\alpha t}}{2\alpha} + 2\rho\sigma_l\sigma_s \frac{1 - e^{-\alpha t}}{\alpha} \right] + \sigma_l dW_t^l + \sigma_s \int_0^t e^{-\alpha(t-u)} dW_u^s \right),$$

and the model then falls within the class of models (2)-(3)-(4).

We have seen that the forward products are not impacted by the spikes if λ/β is small and β is large. However, it can have an important impact on options on the spot, for instance on a strip of Call options, with payoff of the form $\sum_{i=1}^I (S_{t_{i,n}} - K)^+$ for prescribed dates $t_{i,n}$. If we consider

an option with payoff having a single component $(S_t - K)^+$, the jump process will have a weak impact: the probability to have a jump at time t is equal to 0 and even if there is a jump before, it disappears very quickly. However, the jump process may have a significant impact on the value of options with payoff $\sum_{i=1}^I (S_{t_{i,n}} - K)^+$ because the probability of having spikes on $[0, 1]$ is non zero. (Note that only upward spikes will have an impact on the price of these options.)

Unlike spot prices, forward contracts are tradable assets. In the following, we assume absence of arbitrage opportunity. According to the fundamental theorem of asset pricing, there exists a probability measure \mathbb{Q} equivalent to the historical measure \mathbb{P} such that $f(t, T)$ is a local martingale under \mathbb{Q} . Because of the presence of jumps, the market is incomplete and \mathbb{Q} is not unique. According to [22, Theorem 2], there exists a predictable process $(\gamma_t)_{t \geq 0}$ and a predictable process $(Y(\cdot, t, x))_{t \geq 0, x \in \mathbb{R}}$ such that:

- 1) $\mu_t + \gamma_t c_t + \int_0^t \int_{\mathbb{R}} x \lambda Y(t, x) e^{-\beta(T-t)} \nu(dx) = 0$ ($\mathbb{P} \otimes dt$ almost-surely),
- 2) $\int_0^1 \gamma_s^2 c_s ds < \infty$ almost surely,
- 3) $\int_0^1 \int_{\mathbb{R}} |x|^2 \wedge |x| Y(t, x) e^{-\beta(T-t)} \lambda \nu(dx) < \infty$ ($\mathbb{P} \otimes dt$ almost -surely).

with c_t equal to $f^c(t, T) (\sigma_l^2 + \sigma_s^2 e^{-2\alpha(T-t)} + 2\rho\sigma_l\sigma_s e^{-\alpha(T-t)})^{1/2}$ in our case. Under the equivalent measure, $f(t, T)$ is an Itô semi-martingale with drift 0, volatility c_t and jump measure $p^* = Yp$ following

$$df(t, T) = df^c(t, T) + df^\beta(t, T)$$

with

$$df^c(t, T) = f^c(t, T) \left(\sigma_l dW_t^{l,*} + \sigma_s e^{-\alpha(T-t)} dW_t^{s,*} \right),$$

$$df^\beta(t, T) = \int_{\mathbb{R}} x e^{-\beta(T-t)} \left(\underline{p}^*(dt, dx) - \lambda Y(t, x) \nu(dx) dt \right)$$

for two Brownian motions $(W^{s,*}, W^{l,*})$ under the new measure. The volatility does not change unlike the intensity and the law of jump sizes of the Poisson process.

In order to choose the change of martingale measure, one usually consider an optimisation criterion. One of the most common used criterion is the local risk-minimisation introduced by Föllmer and Schweizer (see [24] for details). The variance of the cost of the strategy is minimised locally, infinitesimally at each time. This strategy corresponds to choose the minimal martingale measure defined in [11]. Under certain assumptions, this measure is a true probability measure and the asset is a local martingale under this measure. Furthermore, the intensity changes and depends on the drift μ , which is also true for most common criteria. Since we work on a finite time framework, the drift is not identifiable and it is not possible to estimate it.

In the following we choose the historical approach of Merton consisting in picking a change of probability that does not affect the intensity and the jump sizes of the Poisson measure [18]. The equivalent probability measure is defined by

$$\frac{d\mathbb{Q}^M}{d\mathbb{P}} = \exp \left(\int_0^1 \theta_u d \frac{\sigma_l W_u^l + \sigma_s e^{-\beta(T-u)} W_u^s}{(\sigma_l^2 + \sigma_s^2 e^{-2\alpha(T-u)} + 2\rho\sigma_l\sigma_s e^{-\alpha(T-u)})^{1/2}} - \frac{1}{2} \int_0^1 \theta_u^2 du \right)$$

with $\theta_u = \frac{-(\mu_u + e^{-\beta(T-u)} \int_{\mathbb{R}} x \nu(dx))}{f^c(u, T) (\sigma_l^2 + \sigma_s^2 e^{-2\alpha(T-u)} + 2\rho\sigma_l\sigma_s e^{-\alpha(T-u)})^{1/2}}$. The Nikonikov condition is satisfied so it defines in turn a genuine probability measure. Under \mathbb{Q}^M , the price of the forward contract $f(t, T)$ follows the dynamics $df(t, T) = df^c(t, T) + df^\beta(t, T)$ with

$$df^c(t, T) = f^c(t, T) \left(\sigma_l dW_t^{l, \mathbb{Q}^M} + \sigma_s e^{-\alpha(T-t)} dW_t^{s, \mathbb{Q}^M} \right)$$

and

$$df^\beta(t, T) = \int_{\mathbb{R}} x e^{-\beta(T-t)} (\underline{p}(dt, dx) - \lambda \nu(dx) dt),$$

where W^{l, \mathbb{Q}^M} and W^{s, \mathbb{Q}^M} are two \mathbb{Q}^M -Brownian motions. Merton chooses this probability considering that the risk associated to the jumps is diversifiable. As noticed in Tankov in [9, Section 10.1], using this strategy leaves one exposed to the risk of the jumps. It only corrects the average effect of jumps (provided that the jump component of the electricity price is independent of the other assets, which is the case here: we understand the electricity spikes caused by physical exogenous events; it can in particular be caused by the production capacity and the demand which are not assets (see the structural model of Aid et al. [1] for instance). Finally, the price of an option with payoff $H(S_T) = H(f(T, T))$ at time t is given by $\mathbb{E}^{\mathbb{Q}^M}[H(S_T) | \mathcal{F}_t]$.

4.3. Application to Call option pricing in the French market. In the following, we focus on the French market and we work on the model of Section 4.2. We dispose of the hourly spot and daily forward prices in 2015 and 2016.

Parameters of f^β . We use the parameters found in Table 4 to calibrate Z^β to the spot prices. We model the size of the jumps by its empirical distribution, each jump being estimated with $\Delta_{T_n(q)}^n X$, knowing that a bias (mentioned in the end of section 2.2) is present.

Parameters of f^c . We consider the following forward products in the French market: 1 to 4 Week-ahead, 1 to 3 Month-ahead, 1 to 4 Quarter-ahead and 1 and 2 Year-ahead products. As $\hat{\lambda}_n/\hat{\beta}_n$ is small and $\hat{\beta}_n$ is large, we can neglect the jump part on the forward prices and consider that the forward products have only a continuous part. We use the method of Féron and Daboussi [10] to calibrate the parameters of f^c to the observed forward prices. We find for the different parameters $\alpha = 12.56 y^{-1}$, $\sigma_s = 1.03 y^{-\frac{1}{2}}$, $\sigma_l = 0.25 y^{-\frac{1}{2}}$ and $\rho = -0.11$.

Forward products. In Figure 4, we display a simulation of the spot price, the 1WAH and the 1MAH with and in absence of spikes. The parameters of the spike component are the one of Table 4 with $C = 5$. We observe that the difference between the trajectory of the forward products with and without spikes is very small but significant for the spot price.

Strip of Call options. We consider an option of payoff $\sum_{i=1}^I (S_{t_{i,n}} - K)^+$, with $I = 8760$ corresponding to possible exercises each hour of one year. This choice of payoff is motivated by the valuation of a power plant: the produced electricity is sold on the market at S and the production cost is K . A high strike corresponds to peaking power stations activated when other plants can not satisfy the total load. The price of such an option is equal to $\mathbb{E}^{\mathbb{Q}^M}[\sum_{i=1}^I (S_{t_{i,n}} - K)^+]^3$. We give in Table 5 confidence intervals at level 95% for the option price with the different strikes 100, 200 and 300, computed using Monte Carlo method with 10000 simulations. We consider the case where there is no spikes and the cases with spikes using the different threshold of the form $v_n = C\hat{\sigma}\Delta_n^{-0.01}$ with $C = 3$, $C = 4$ and $C = 5$. Considering spikes leads to higher value for the strip options. Furthermore, options valued at zero have now non negligible values. We notice that the choice of the threshold have an impact on the price of the option. Indeed, a higher C leads to less jumps and a smaller λ , see Table 4, and then to a lower price as most of the jumps are positive. However, the impact is low since we keep the larger jumps when C is increased which are the ones impacting the price for high strikes.

³For simplicity, we consider a risk-free rate equal to zero

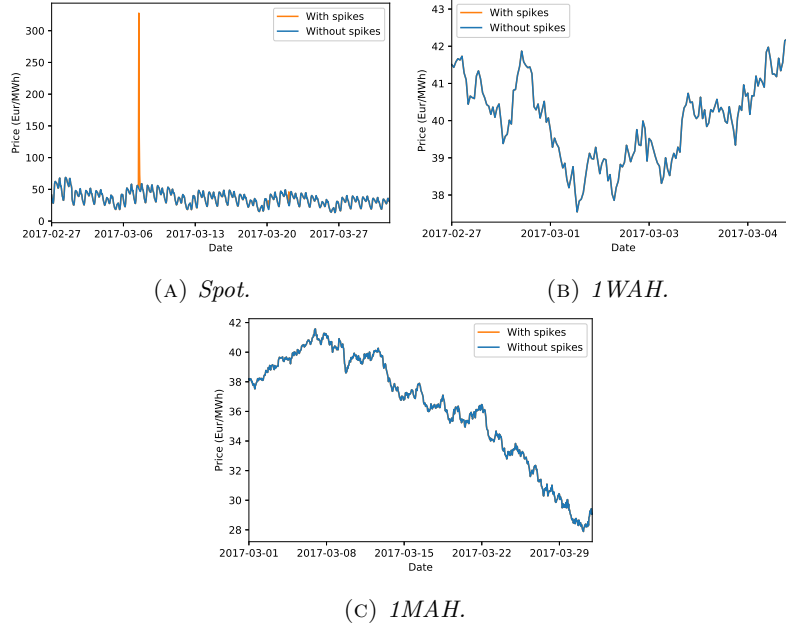


FIGURE 4. Simulation of different products in a two factor model with and without spikes between the 27th of February 2017 and the 31st of March 2017. We illustrate the spot, the 1WAH starting the 27th of February 2017 and the 1MAH starting the 01st of March 2017.

Model/Strike	100	200	300
Without spike	[1716.22, 1806.83]	[0.0089, 0.063]	[0, 0]
Spikes, $C = 3$	[2482.57, 2576.17]	[434.21, 450.26]	[264.14, 274.00]
Spikes, $C = 4$	[2442.66, 2536.26]	[412.98, 428.44]	[251.72, 262.29]
Spikes, $C = 5$	[2417.24, 2510.79]	[397.74, 412.76]	[242.04, 252.41]

TABLE 5. Confidence intervals at level 95% for the price of strip options computed using Monte Carlo method with 10000 simulations for different strikes and different models.

5. PROOFS

In the following proofs, we set the drift $(\mu_t)_{t \geq 0}$ vanishes identically. Generalizing to the non-null drift case is done using the usual argument based on Girsanov theorem. Also, we assume for simplicity that $(\sigma_t)_{t \geq 0}$ is a deterministic function, in order to simplify in particular the proofs for central limit theorems.

5.1. **Proof of Proposition 4.** The proof follows the path of [3, Theorem 10.26, p.374] and is also close to Mancini [17] in spirit. We will denote by ξ_ν a random variable distributed according to ν . Let

$$\mathcal{A}_n = \{i \in \{1, \dots, n\}, i \neq i(n, q) \ \forall q \geq 1\}$$

be the set of indices i such that the interval $((i-1)\Delta_n, i\Delta_n]$ contains no jump.

Proof of Proposition 4 under Assumption 3 (I). We first need to show

$$(14) \quad \mathbb{P}\left(\max_{i \in \mathcal{A}_n} \frac{|\Delta_i^n X|}{\sqrt{\Delta_n}} > v_n\right) \rightarrow 0,$$

$$(15) \quad \mathbb{P}\left(\min_{i \in \mathcal{A}_n^c} \frac{|\Delta_i^n X|}{\sqrt{\Delta_n}} < v_n\right) \rightarrow 0$$

and

$$(16) \quad \mathbb{P}\left(\max_{1 \leq i \leq n} \Delta_i^n N \geq 2\right) \rightarrow 0.$$

We have

$$\frac{\Delta_i^n X}{\sqrt{\Delta_n}} = \frac{-\beta_n \int_{t_{i-1,n}}^{t_{i,n}} Z_s^\beta ds}{\sqrt{\Delta_n}} + \frac{\Delta_i^n X'}{\sqrt{\Delta_n}},$$

with

$$X'_t = \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s + \int_0^t \int_{\mathbb{R}} x \underline{p}(ds, dx).$$

By [3, Equation (10.71), p.374] we have $\mathbb{P}\left(\sup_{i \in \mathcal{A}_n} \frac{|\Delta_i^n X'}{\sqrt{\Delta_n}} > v_n\right) \rightarrow 0$, Therefore, in order to prove (14), we need to show that

$$(17) \quad \mathbb{P}\left(\sup_{i \in \mathcal{A}_n} \frac{|\beta_n \int_{t_{i-1,n}}^{t_{i,n}} Z_s^\beta ds|}{\sqrt{\Delta_n}} > v_n\right) \rightarrow 0.$$

Since $|\beta_n \int_{t_{i-1,n}}^{t_{i,n}} Z_s^\beta ds| = (1 - e^{-\beta_n \Delta_n}) |Z_{t_{i-1,n}}^\beta|$ for $i \in \mathcal{A}_n$, we have

$$\begin{aligned} \mathbb{P}\left(\max_{i \in \mathcal{A}_n} \frac{|\beta_n \int_{t_{i-1,n}}^{t_{i,n}} Z_s^\beta ds|}{\sqrt{\Delta_n}} > v_n\right) &\leq \mathbb{P}\left(\beta_n \sqrt{\Delta_n} \sup_{t \in [0,1]} |Z_t^\beta| > v_n\right) \\ &\leq \mathbb{P}\left(\sup_{t \in [0,1]} \int_0^t \int_{\mathbb{R}} |x| e^{-\beta_n(t-s)} \underline{p}(ds, dx) > \frac{v_n}{\beta_n \sqrt{\Delta_n}}\right) \\ &\leq 2\lambda_n \mathbb{P}\left(|\xi_\nu| > \frac{v_n}{\beta_n \sqrt{\Delta_n}}\right) \end{aligned}$$

by Markov's inequality and [7, Equation (10)] on the expectation of the crossings of shot noise processes. This last term converges to 0 by Assumption 3 (ii) and (17) follows which completes the proof of (14).

We next turn to (16). The left hand side of (16) is equal to $\mathbb{P}(\cup_{i=1}^n \Delta_i^n N \geq 2) \lesssim \lambda_n^2 \Delta_n$ which converges to 0 if $\lambda_n \sqrt{\Delta_n} \rightarrow 0$. With no loss of generality we may (and will) work on the set $\{\max_{1 \leq i \leq n} \Delta_i^n N \leq 1\}$. In the interval $((i(n,q) - 1) \Delta_n, i(n,q) \Delta_n]$, there is only one jump and we have

$$\Delta_{i(n,q)}^n Z^\beta = -(1 - e^{-\beta_n \Delta_n}) Z_{t_{i(n,q)-1,n}} + e^{-\beta_n(i(n,q)\Delta_n - T_q)} \Delta X_{T_q}$$

for all $q \geq 1$, therefore

$$\begin{aligned} \frac{|\Delta_{i(n,q)}^n X|}{\sqrt{\Delta_n}} &\geq \frac{e^{-\beta_n \Delta_n} |\Delta X_{T_q}|}{\sqrt{\Delta_n}} - \frac{|\Delta_{i(n,q)}^n X^c|}{\sqrt{\Delta_n}} - \frac{|(1 - e^{-\beta_n \Delta_n}) Z_{t_{i(n,q)-1,n}}^\beta|}{\sqrt{\Delta_n}} \\ &\geq \min_{1 \leq q \leq N_1} \frac{e^{-\beta_n \Delta_n} |\Delta X_{T_q}|}{\sqrt{\Delta_n}} - \max_{i \in \mathcal{A}_n^c} \frac{|\Delta_i^n X^c|}{\sqrt{\Delta_n}} - \max_{1 \leq i \leq n} \frac{\beta_n \Delta_n |Z_{t_{i,n}}^\beta|}{\sqrt{\Delta_n}}. \end{aligned}$$

It follows that $\mathbb{P}(\min_{i \in \mathcal{A}_n^c} \frac{|\Delta_i^n X|}{\sqrt{\Delta_n}} \leq v_n)$ is dominated by the sum of the three following terms:

$$(18) \quad \mathbb{P}\left(\min_{1 \leq q \leq N_1} |\Delta X_{T_q}| e^{-\beta_n \Delta_n} \leq 3v_n \sqrt{\Delta_n}\right),$$

$$(19) \quad \mathbb{P}\left(\min_{1 \leq q \leq N_1} |\Delta X_{T_q}| e^{-\beta_n \Delta_n} \leq 3 \max_{i \in \mathcal{A}_n^c} |\Delta_i^n X^c|\right)$$

and

$$(20) \quad \mathbb{P}\left(\min_{1 \leq q \leq N_1} |\Delta X_{T_q}| e^{-\beta_n \Delta_n} \leq 3 \max_{1 \leq i \leq n} \beta_n \Delta_n |Z_{t_{i,n}}^\beta|\right).$$

The term (18) equals

$$\mathbb{E}\left[1 - \left(\mathbb{P}(|\xi_\nu| > 3v_n \sqrt{\Delta_n} e^{\beta_n \Delta_n})\right)^{N_1}\right] = 1 - \exp\left(-\lambda_n \mathbb{P}(|\xi_\nu| \leq 3v_n \sqrt{\Delta_n} e^{\beta_n \Delta_n})\right)$$

and converges to 0 under the assumption $\lambda_n \mathbb{P}(|\xi_\nu| \leq \Delta_n^{\frac{1}{2} - \varpi}) \rightarrow 0$. The term (19) is dominated by

$$(21) \quad \mathbb{P}\left(\min_{1 \leq q \leq N_1} |\Delta X_{T_q}| e^{-\beta_n \Delta_n} \leq 3v_n \sqrt{\Delta_n}\right) + \mathbb{P}\left(v_n \sqrt{\Delta_n} \leq \max_{i \in \mathcal{A}_n^c} |\Delta_i^n X^c|\right).$$

The left hand side of (21) is equal to (18) and converges to 0. According to [17, Corollary 3.3], for $i \in \{1, \dots, n\}$,

$$\mathbb{P}(\Delta_i^n X^c > v_n \sqrt{\Delta_n}) \leq 2e^{-v_n^2/2\bar{\sigma}^2}.$$

The right hand side of (21) is dominated by

$$\mathbb{E}\left[\sum_{q=1}^{N_1} \mathbb{P}(|\Delta_i^n X^c| \geq v_n \sqrt{\Delta_n})\right] \leq \mathbb{E}[N_1] 2e^{-v_n^2/2\bar{\sigma}^2} = 2\lambda_n e^{-v_n^2/2\bar{\sigma}^2} \rightarrow 0.$$

The term (20) is dominated by

$$(22) \quad \mathbb{P}\left(\min_{1 \leq q \leq N_1} |\Delta X_{T_q}| e^{-\beta_n \Delta_n} \leq 3v_n \sqrt{\Delta_n}\right) + \mathbb{P}\left(v_n \sqrt{\Delta_n} \leq \max_{1 \leq i \leq n} \beta_n \Delta_n |Z_{t_{i,n}}^\beta|\right).$$

The left hand side of (21) is equal to (18) and converges to 0. The right hand side of (21) also converges to 0 with the same argument as for (17).

Proof of Proposition 4 under Assumption 3 (II). Since

$$\mathbb{P}\left(\max_{0 \leq i \leq n-k_n} |N_{(i+k_n)\Delta_n} - N_{i\Delta_n}| \geq 2\right) \lesssim \lambda_n^2 \Delta_n k_n^2 \rightarrow 0,$$

therefore, we only need to prove the result on the set $\mathcal{B}_n = \left\{ \max_{0 \leq i \leq n-k_n} |N_{(i+k_n)\Delta_n} - N_{i\Delta_n}| \leq 1 \right\}$.

We need to show:

$$(23) \quad \mathbb{P}\left(\exists i \in \mathcal{A}_n, \frac{|\Delta_i^n X|}{\sqrt{\Delta_n}} > v_n \text{ and } \Delta_i^n X \Delta_{i+1}^n X < 0 \cap \mathcal{B}_n\right) \rightarrow 0$$

and

$$(24) \quad \mathbb{P}\left(\exists i \in \mathcal{A}_n^c, \frac{|\Delta_i^n X|}{\sqrt{\Delta_n}} < v_n \text{ or } \Delta_i^n X \Delta_{i+1}^n X > 0 \cap \mathcal{B}_n\right) \rightarrow 0.$$

We bound (23) above by the sum of

$$\mathbb{P}\left(\exists i \in \mathcal{A}_n, \frac{|\Delta_i^n X^c|}{\sqrt{\Delta_n}} > \frac{v_n}{2} \text{ and } \Delta_i^n X \Delta_{i+1}^n X < 0 \cap \mathcal{B}_n\right)$$

and

$$\mathbb{P}\left(\exists i \in \mathcal{A}_n, \frac{|\Delta_i^n Z^\beta|}{\sqrt{\Delta_n}} > \frac{v_n}{2} \text{ and } \Delta_i^n X \Delta_{i+1}^n X < 0 \cap \mathcal{B}_n\right).$$

The first term is bounded by

$$\mathbb{P}\left(\max_{1 \leq i \leq n} |\Delta_i^n X^c| > \frac{v_n \sqrt{\Delta_n}}{2}\right) \leq 2ne^{-v_n^2/8\sigma^2}$$

and converges to 0. For the second term, we consider two cases: a jump occurs before $i\Delta_n$ or no such jump occurs, which leads us to further consider

$$(25) \quad \mathbb{P}\left(\exists i \in \mathcal{A}_n, \frac{|\Delta_i^n Z^\beta|}{\sqrt{\Delta_n}} > \frac{v_n}{2}, \Delta_i^n X \Delta_{i+1}^n X < 0 \text{ and } \exists q \in \{1, \min\{k_n-2, i-1\}\}, i-q \in \mathcal{A}_n^c \cap \mathcal{B}_n\right)$$

and

$$(26) \quad \mathbb{P}\left(\exists i \in \mathcal{A}_n, \frac{|\Delta_i^n Z^\beta|}{\sqrt{\Delta_n}} > \frac{v_n}{2}, \Delta_i^n X \Delta_{i+1}^n X < 0 \text{ and } \forall q \in \{1, \min\{k_n-2, i-1\}\}, i-q \in \mathcal{A}_n\right) \cap \mathcal{B}_n.$$

For (25) since we work on \mathcal{B}_n , we have $i+1 \in \mathcal{A}_n$ hence $\Delta_{i+1}^n Z^\beta = -(1 - e^{-\beta_n \Delta_n}) Z_{t_{i,n}}^\beta = e^{-\beta_n \Delta_n} \Delta_i^n Z^\beta$ and (25) is dominated by the probability of the event

$$\left\{ \exists i \in \mathcal{A}_n, -|\Delta_i^n X^c| |\Delta_{i+1}^n X^c| - |\Delta_i^n X^c| |\Delta_i^n Z^\beta| e^{-\beta_n \Delta_n} - |\Delta_{i+1}^n X^c| |\Delta_i^n Z^\beta| + \frac{v_n \sqrt{\Delta_n}}{2} |\Delta_i^n Z^\beta| e^{-\beta_n \Delta_n} < 0 \right. \\ \left. \text{and } \frac{|\Delta_i^n Z^\beta|}{\sqrt{\Delta_n}} > \frac{v_n}{2} \right\}$$

equal to

$$\mathbb{P}\left(\exists i \in \mathcal{A}_n, \frac{v_n \sqrt{\Delta_n}}{2} |\Delta_i^n Z^\beta| e^{-\beta_n \Delta_n} < 2 \max_{1 \leq j \leq n} |\Delta_j^n X^c| |\Delta_i^n Z^\beta| + \left(\max_{1 \leq j \leq n} |\Delta_j^n X^c|\right)^2 \text{ and } \frac{|\Delta_i^n Z^\beta|}{\sqrt{\Delta_n}} > \frac{v_n}{2}\right)$$

and dominated by

$$\mathbb{P}\left(2 \max_{1 \leq j \leq n} |\Delta_j^n X^c| > \frac{v_n \sqrt{\Delta_n}}{4}\right) + \mathbb{P}\left(\left(\max_{1 \leq j \leq n} |\Delta_j^n X^c|\right)^2 > \frac{1}{2} \left(\frac{v_n \sqrt{\Delta_n}}{2}\right)^2\right) \rightarrow 0.$$

Concerning (26), we have, if $k_n \leq i+1$, we have that $|\Delta_i^n Z^\beta|$ is equal to

$$(1 - e^{-\beta_n \Delta_n}) e^{-\beta_n \Delta_n (k_n - 2)} |Z_{(i-k_n+1)\Delta_n}^\beta| \leq \beta_n \Delta_n \sup_{t \in [0,1]} \int_0^t \int_{\mathbb{R}} |x| e^{-\beta_n (t-s)} \underline{p}(ds, dx) e^{-\beta_n \Delta_n (k_n - 2)}.$$

The inequality remains true if $i+1 < k_n$ as $\Delta_i^n Z^\beta$ is equal to 0 in this case. Thus, (26) is dominated by

$$\mathbb{P}\left(\sup_{t \in [0,1]} \int_0^t \int_{\mathbb{R}} |x| e^{-\beta_n (t-s)} \underline{p}(ds, dx) > e^{\beta_n \Delta_n (k_n - 2)} \frac{v_n}{\beta_n \sqrt{\Delta_n}}\right) \leq \lambda_n \mathbb{P}\left(|\xi_\nu| > e^{\beta_n \Delta_n (k_n - 1)} \frac{v_n}{\beta_n \sqrt{\Delta_n}}\right) \rightarrow 0$$

using the same argument as for (17) and the convergence (23) follows.

We now turn to (24). It suffices to show that both terms

$$\mathbb{P}\left(\exists i \in \mathcal{A}_n^c, \frac{|\Delta_i^n X|}{\sqrt{\Delta_n}} \geq v_n \cap \mathcal{B}_n\right), \text{ and } \mathbb{P}\left(\exists i \in \mathcal{A}_n^c, \Delta_i^n X \Delta_{i+1}^n X \geq 0 \cap \mathcal{B}_n\right)$$

converge to 0. The proof for the first term is similar to the one of (15), the only difference being $\Delta_{i(n,q)}^n Z$ is equal to $(1 - e^{-\beta_n \Delta_n}) e^{-\beta_n \Delta_n (k_n - 2)} Z_{t_{i(n,q)-k_n+1,n}} + e^{-\beta_n (t_{i(n,q),n} - T_q)} \Delta X_{T_q}$ if $k_n \leq$

$i+1$ and $e^{-\beta_n(t_{i(n,q),n}-T_q)}\Delta X_{T_q}$ otherwise and that the term $\mathbb{P}(v_n\sqrt{\Delta_n} \leq 3 \max_{1 \leq i \leq n} \beta_n \Delta_n |Z_{t_{i,n}}^\beta|)$ needs to be replaced by

$$\mathbb{P}(v_n\sqrt{\Delta_n} \leq 3 \max_{1 \leq i \leq n} \beta_n \Delta_n e^{-\beta_n \Delta_n k_n} |Z_{t_{i,n}}^\beta|) \leq \lambda_n \mathbb{P}(|\xi_\nu| > e^{\beta_n \Delta_n (k_n-2)} \frac{v_n \sqrt{\Delta_n}}{\beta_n \Delta_n}) \rightarrow 0.$$

For the second term, it is sufficient to prove that $\Delta_{i(n,q)}^n X$ has the same sign as ΔX_{T_q} and that $\Delta_{i(n,q)+1}^n X$ has the opposite sign with probability one. We are thus led to show that

$$(27) \quad \mathbb{P}\left(\min_{1 \leq q \leq N_1} \Delta_{i(n,q)}^n X \Delta X_{T_q} < 0\right) \rightarrow 0$$

and

$$(28) \quad \mathbb{P}\left(\max_{1 \leq q \leq N_1} \Delta_{i(n,q)+1}^n X \Delta X_{T_q} > 0\right) \rightarrow 0.$$

We have that $\Delta_{i(n,q)}^n X \Delta X_{T_q}$ dominates

$$\begin{aligned} & - \max_{1 \leq q \leq N_1} |\Delta_{i(n,q)}^n X^c| |\Delta X_{T_q}| - (1 - e^{-\beta_n \Delta_n}) e^{-\beta_n \Delta_n k_n} \sup_{t \in [0,1]} \int_0^t \int_{\mathbb{R}} |x| e^{-\beta_n(t-s)} \underline{p}(ds, dx) |\Delta X_{T_q}| \\ & + e^{-\beta_n \Delta_n} |\Delta X_{T_q}| \min_{1 \leq q \leq N_1} |\Delta X_{T_q}|. \end{aligned}$$

We thus have that (27) is dominated by the probability of the event

$$\begin{aligned} & \left\{ e^{-\beta_n \Delta_n} \inf_{1 \leq q \leq N_1} |\Delta X_{T_q}| < \sup_{1 \leq q \leq N_1} |\Delta_{i(n,q)}^n X^c| \right. \\ & \quad \left. + (1 - e^{-\beta_n \Delta_n}) e^{-\beta_n \Delta_n k_n} \sup_{t \in [0,1]} \int_0^t \int_{\mathbb{R}} |x| e^{-\beta_n(t-s)} \underline{p}(ds, dx) \right\} \end{aligned}$$

that converges to 0 if we use a similar proof than the one of (15). The proof of (28) is similar since no jump occurs in the interval $[i(n,q), i(n,q)+1]$ and

$$\Delta_{i(n,q)+1}^n Z^\beta = - (1 - e^{-\beta_n \Delta_n}) Z_{i(n,q)\Delta_n}^\beta.$$

The term (28) is then dominated by the probability of the event

$$\begin{aligned} & \left\{ e^{-\beta_n \Delta_n} (1 - e^{-\beta_n \Delta_n}) \min_{1 \leq q \leq N_1} |\Delta X_{T_q}| < \max_{1 \leq q \leq N_1} |\Delta_{i(n,q)}^n X^c| \right. \\ & \quad \left. + (1 - e^{-\beta_n \Delta_n}) e^{-\beta_n \Delta_n (k_n+1)} \sup_{t \in [0,1]} \int_0^t \int_{\mathbb{R}} |x| e^{-\beta_n(t-s)} \underline{p}(ds, dx) \right\}. \end{aligned}$$

The convergence to 0 is obtained in the same way, except for the extra control of the terms

$$\mathbb{P}\left(\max_{1 \leq q \leq N_1} |\Delta_{i(n,q)}^n X^c| > v_n \sqrt{\Delta_n} (1 - e^{-\beta_n \Delta_n})\right) \leq \lambda_n e^{-v_n^2 (1 - e^{-\beta_n \Delta_n})^2 / 2\sigma^2}$$

and

$$\mathbb{P}\left(\sup_{t \in [0,1]} \int_0^t \int_{\mathbb{R}} |x| e^{-\beta_n(t-s)} \underline{p}(ds, dx) > e^{\beta_n \Delta_n (k_n+1)} v_n \sqrt{\Delta_n}\right) \leq \lambda_n \mathbb{P}(|\xi_\nu| > e^{\beta_n \Delta_n k_n} \Delta_n^{\frac{1}{2}-\varpi})$$

that both converge to 0. The proof of Proposition 4 is complete.

5.2. Proof of Theorem 6.

Preparation for the proof. In order to prove Theorem 6, we start by giving an oracle estimator of β_n when the jump times and their sizes are known.

Proposition 9. *Work under Assumption 1 and 2. Let $\beta_n \sqrt{\lambda_n \Delta_n} \rightarrow \infty$. Define $\widehat{\beta}_n^{\text{oracle}}$ via*

$$\exp(-\Delta_n \widehat{\beta}_n^{\text{oracle}}) = \max \left\{ 1 + \frac{\sum_{q \in \mathcal{E}_n} \text{sgn}(\Delta X_{T_q}) (\Delta_{i(n,q)+1}^n X + 2\Delta_n \sum_{j=1}^{q-1} \Delta X_{T_j})}{\sum_{q \in \mathcal{E}_n} \text{sgn}(\Delta X_{T_q}) \Delta_{i(n,q)}^n X} \mathbf{1}_{\{N_1 > 0\}}, \Delta_n \right\},$$

with $\mathcal{E}_n = \{q \in \{1, \dots, N_1\}, i(n, q) + 1 \in \mathcal{A}_n \text{ and } i(n, q) < i(n, q + 1)\}$.

i) *The following expansion holds on the set $\{N_1 > 0\}$:*

$$\frac{\widehat{\beta}_n^{\text{oracle}} - \beta_n}{\beta_n} = \mathcal{M}_n^{\text{oracle}} + \mathcal{V}_n \mathcal{J}_n^T,$$

with

$$\mathcal{M}_n^{\text{oracle}} = e^{\beta_n \Delta_n} \frac{\lambda_n}{\beta_n} \frac{(x \star \nu) (\text{sgn}(x) \star \nu)}{|x| \star \nu} \left(\frac{e^{\beta_n \Delta_n} - 1}{\beta_n \Delta_n} - \frac{\beta_n \Delta_n}{1 - e^{-\beta_n \Delta_n}} \right),$$

and $\mathcal{V}_n = (\mathcal{V}_n^{(i)})_{1 \leq i \leq 4} \in \mathbb{R}^4$ defined by

$$\begin{cases} \mathcal{V}_n^{(1)} &= e^{\beta_n \Delta_n} \frac{\sqrt{\lambda_n}}{\sqrt{3} \beta_n |x| \star \nu} \left((\text{sgn}(x) \star \nu)^2 (|x|^2 \star \nu) + (x \star \nu)^2 - 2(\text{sgn}(x) \star \nu)(|x|^2 \star \nu) \right)^{1/2}, \\ \mathcal{V}_n^{(2)} &= e^{\beta_n \Delta_n} \min \left\{ \left(\frac{|x|^2 \star \nu}{(|x| \star \nu)^2} \frac{1}{2\beta_n} \frac{(1 - e^{-2\beta_n \Delta_n})}{2\beta_n \Delta_n} \right)^{1/2}, \frac{\lambda_n}{\beta_n} \right\}, \\ \mathcal{V}_n^{(3)} &= e^{\beta_n \Delta_n} \left(\frac{\beta_n \Delta_n}{1 - e^{-\beta_n \Delta_n}} \right) \frac{\sqrt{\int_0^1 \sigma_s^2 ds}}{|x| \star \nu \sqrt{\lambda_n \beta_n \Delta_n}}, \\ \mathcal{V}_n^{(4)} &= e^{\beta_n \Delta_n} \frac{\sqrt{\int_0^1 \sigma_s^2 ds \sqrt{\Delta_n}}}{|x| \star \nu \sqrt{\lambda_n}}, \end{cases}$$

and $\mathcal{J}_n = (\mathcal{J}_n^{(i)})_{1 \leq i \leq 4} \in \mathbb{R}^4$ is bounded in probability as $n \rightarrow \infty$.

ii) *If $\lambda_n \rightarrow \infty$, then*

$$(\mathcal{J}_n^{(3)}, \mathcal{J}_n^{(4)}) \rightarrow \mathcal{N}(0, \text{Id}_{\mathbb{R}^2})$$

in distribution as $n \rightarrow \infty$.

iii) *If $\lambda_n \rightarrow \infty$, $|x|^3 \star \nu < \infty$ and $(\text{sgn}(x) \star \nu)^2 |x|^2 \star \nu + (x \star \nu)^2 - 2\text{sgn}(x) \star \nu |x|^2 \star \nu \neq 0$, we have*

$$(\mathcal{J}_n^{(1)}, \mathcal{J}_n^{(3)}, \mathcal{J}_n^{(4)}) \rightarrow \mathcal{N}(0, \text{Id}_{\mathbb{R}^3})$$

in distribution as $n \rightarrow \infty$.

iv) *If $\beta_n / \lambda_n^2 \rightarrow 0$, the conditions of iii) and $|x|^4 \star \nu < \infty$ hold together, we finally obtain*

$$\mathcal{J}_n \rightarrow \mathcal{N}(0, \text{Id}_{\mathbb{R}^4})$$

in distribution as $n \rightarrow \infty$.

Proposition 9 is the core of Theorem 6.

Proof of Proposition 9. Step 1). We first need three approximation results.

Lemma 10. *We have*

$$\left| \sum_{q \notin \mathcal{E}_n} \text{sgn}(\Delta X_{T_q}) \Delta_{i(n,q)}^n X \right| \lesssim \lambda_n^2 \Delta_n$$

in probability.

Proof. We plan to use the decomposition

$$\sum_{q \notin \mathcal{E}_n} \operatorname{sgn}(\Delta X_{T_q}) \Delta_{i(n,q)}^n X = \sum_{q=1}^{N_1} \operatorname{sgn}(\Delta X_{T_q}) \Delta_{i(n,q)}^n X \mathbf{1}_{\{T_{q+1}-T_q < t_{i(n,q),n}-T_q+\Delta_n\}} = I + II,$$

with

$$I = \sum_{q=1}^{N_1} \operatorname{sgn}(\Delta X_{T_q}) \Delta_{i(n,q)}^n X^c \mathbf{1}_{\{T_{q+1}-T_q < t_{i(n,q),n}-T_q+\Delta_n\}} \quad \text{and}$$

$$II = \sum_{q=1}^{N_1} \operatorname{sgn}(\Delta X_{T_q}) \Delta_{i(n,q)}^n Z^\beta \mathbf{1}_{\{T_{q+1}-T_q < t_{i(n,q),n}-T_q+\Delta_n\}}.$$

The term I is centred and as $|I| \leq \sum_{q=1}^{N_1} |\Delta_{i(n,q)}^n X^c| \mathbf{1}_{\{T_{q+1}-T_q < t_{i(n,q),n}-T_q+\Delta_n\}}$, we have

$$\begin{aligned} \mathbb{E}[I^2] &\leq \mathbb{E}\left[\left(\sum_{q=1}^{N_1} |\Delta_{i(n,q)}^n X^c| \mathbf{1}_{\{T_{q+1}-T_q < t_{i(n,q),n}-T_q+\Delta_n\}}\right)^2\right] \\ &\leq \mathbb{E}\left[\left(\sum_{i=1}^n |\Delta_i^n X^c| \Delta_i^n N \mathbf{1}_{\{\Delta_i^n N \geq 2 \text{ or } \Delta_{i+1}^n N \geq 1\}}\right)^2\right] \\ &= \sum_{i=1}^n \mathbb{E}[(\Delta_i^n X^c)^2] \mathbb{E}[(\Delta_i^n N)^2 \mathbf{1}_{\{\Delta_i^n N \geq 2 \text{ or } \Delta_{i+1}^n N \geq 1\}}] \end{aligned}$$

with the convention $\Delta_{n+1}^n N = 0$. Since $\mathbb{E}[(\Delta_i^n X^c)^2] \leq \bar{\sigma}^2 \Delta_n$ and $\mathbb{E}[(\Delta_i^n N)^2 (\mathbf{1}_{\{\Delta_i^n N \geq 2\}} + \mathbf{1}_{\{\Delta_{i+1}^n N \geq 1\}})] \lesssim \lambda_n^2 \Delta_n^2$ we obtain

$$\mathbb{E}[I^2] \lesssim (\lambda_n \Delta_n)^2.$$

In turn I is of order $\lambda_n \Delta_n$ hence negligible. For the term II , we have

$$|II| \leq \sum_{q=1}^{N_1} |\Delta_{i(n,q)}^n Z^\beta| \mathbf{1}_{\{T_{q+1}-T_q < t_{i(n,q),n}-T_q+\Delta_n\}},$$

and

$$\begin{aligned} |\Delta_{i(n,q)}^n Z^\beta| &= \left| - (1 - e^{-\beta_n \Delta_n}) Z_{t_{i(n,q)-1,n}}^\beta + \sum_{j=1}^{\Delta_{i(n,q)}^n N} \Delta_{i(n,q)}^j X e^{-\beta_n (t_{i(n,q),n} - T_{i(n,q)}^j)} \right| \\ &\leq \beta_n \Delta_n |Z_{t_{i(n,q)-1,n}}^\beta| + \sum_{j=1}^{\Delta_{i(n,q)}^n N} |\Delta X_{i(n,q)}^j|, \end{aligned}$$

where ΔX_i^j denotes the j th jump in the interval $((i-1)\Delta_n, i\Delta_n]$ that occurs at time T_i^j . First, we have that the term $\sum_{q=1}^{N_1} |Z_{t_{i(n,q)-1,n}}^\beta| \mathbf{1}_{\{T_{q+1}-T_q < t_{i(n,q),n}-T_q+\Delta_n\}}$ is dominated by

$$\sum_{i=1}^n \int_0^{t_{i-1,n}} |x| e^{-\beta_n (t_{i-1,n}-t)} \underline{p}(dt, dx) \Delta_i^n N (\mathbf{1}_{\{\Delta_i^n N \geq 2\}} + \mathbf{1}_{\{\Delta_{i+1}^n N \geq 1\}}).$$

Because of the independence of $\Delta_i^n N$ and $\Delta_{i+1}^n N$ conditional on $\mathcal{F}_{t_{i-1,n}}$, we derive that its expectation is less than

$$\sum_{i=1}^n \mathbb{E}\left[\int_0^{t_{i-1,n}} |x| e^{-\beta_n (t_{i-1,n}-t)} \underline{p}(dt, dx)\right] \mathbb{E}[\Delta_i^n N (\mathbf{1}_{\{\Delta_i^n N \geq 2\}} + \mathbf{1}_{\{\Delta_{i+1}^n N \geq 1\}})].$$

Second, since $\mathbb{E}[\int_0^{t_{i-1,n}} |x| e^{-\beta_n(t_{i-1,n}-t)} \underline{p}(dt, dx)] \lesssim \lambda_n/\beta_n$ and also $\mathbb{E}[\Delta_i^n N (\mathbf{1}_{\{\Delta_i^n N \geq 2\}} + \mathbf{1}_{\{\Delta_{i+1}^n N \geq 1\}})] \lesssim \lambda_n^2 \Delta_n^2$ we derive

$$(29) \quad \sum_{q=1}^{N_1} |Z_{t_{i(n,q)-1,n}}^\beta| \mathbf{1}_{\{T_{q+1}-T_q < t_{i(n,q),n}-T_q+\Delta_n\}} \lesssim \frac{\lambda_n^3 \Delta_n}{\beta_n}$$

in probability. In the same way, it is not difficult to see that

$$\sum_{q=1}^{N_1} \sum_{j=1}^{\Delta_{i(n,q)}^n N} |\Delta X_{i(n,q)}^j| \mathbf{1}_{\{T_{q+1}-T_q < 2\Delta_n\}} \leq \sum_{i=1}^n \Delta_i^n N \sum_{j=1}^{\Delta_i^n N} |\Delta X_i^j| (\mathbf{1}_{\{\Delta_i^n N \geq 2\}} + \mathbf{1}_{\{\Delta_{i+1}^n N \geq 1\}})$$

is of order $\lambda_n^2 \Delta_n$. The result of the lemma follows. \square

Lemma 11. *We have*

$$\left| \sum_{q=1}^{N_1} \operatorname{sgn}(\Delta X_{T_q}) \Delta_{i(n,q)}^n X - \int_0^1 \int_{\mathbb{R}} |x| e^{-\beta_n((\lfloor t \Delta_n^{-1} \rfloor + 1)\Delta_n - t)} \underline{p}(dt, dx) \right| \lesssim \lambda_n^2 \Delta_n + \sqrt{\lambda_n \Delta_n}$$

in probability.

Proof. We plan to use the decomposition $\sum_{q=1}^{N_1} \operatorname{sgn}(\Delta X_{T_q}) \Delta_{i(n,q)}^n X = I + II$, with

$$I = \sum_{q=1}^{N_1} \operatorname{sgn}(\Delta X_{T_q}) \Delta_{i(n,q)}^n X^c, \quad \text{and} \quad II = \sum_{q=1}^{N_1} \operatorname{sgn}(\Delta X_{T_q}) \Delta_{i(n,q)}^n Z^\beta.$$

With the notation of Lemma 10, we write

$$I = \sum_{i=1}^n \sum_{j=1}^{\Delta_i^n N} \operatorname{sgn}(\Delta X_i^j) \Delta_i^n X^c \mathbf{1}_{\{\Delta_i^n N \geq 1\}}$$

and in the same way as for the proof of Lemma 10, using the independence between X^c and N , it is not difficult to see that I is centred with variance of order $\lambda_n \Delta_n$. For the second term, we write $II = III + IV$, with

$$\begin{aligned} III &= \sum_{q=1}^{N_1} \operatorname{sgn}(\Delta X_{T_q}) \sum_{j=1}^{\Delta_{i(n,q)}^n N} \Delta X_{i(n,q)}^j e^{-\beta_n((\lfloor T_q \Delta_n^{-1} \rfloor + 1)\Delta_n - T_{i(n,q)}^j)}, \\ IV &= -(1 - e^{-\beta_n \Delta_n}) \sum_{q=1}^{N_1} \operatorname{sgn}(\Delta X_{T_q}) Z_{t_{i(n,q)-1,n}}^\beta. \end{aligned}$$

In the same way as in Lemma 10, the term III is equal to

$$\sum_{q=1}^{N_1} \operatorname{sgn}(\Delta X_{T_q}) \left(\Delta X_{T_q} e^{-\beta_n((\lfloor T_q \Delta_n^{-1} \rfloor + 1)\Delta_n - T_q)} + \sum_{\substack{j=1, \\ T_{i(n,q)}^j \neq T_q}}^{\Delta_{i(n,q)}^n N} \Delta X_q^j e^{-\beta_n((\lfloor T_q \Delta_n^{-1} \rfloor + 1)\Delta_n - T_{i(n,q)}^j)} \mathbf{1}_{\{\Delta_{i(n,q)}^n N \geq 2\}} \right),$$

which is nothing but $\int_0^1 \int_{\mathbb{R}} |x| e^{-\beta_n((\lfloor t\Delta_n^{-1} \rfloor + 1)\Delta_n - t)} \underline{p}(dt, dx)$ plus a remainder term of order $\lambda_n^2 \Delta_n$ in probability. Finally

$$\begin{aligned} |IV| &\leq \beta_n \Delta_n \sum_{q=1}^{N_1} \int_0^{t_{i(n,q)-1,n}} e^{-\beta_n(t_{i(n,q)-1,n} - t)} |x| \underline{p}(dt, dx) \\ &\leq \beta_n \Delta_n e^{\beta_n \Delta_n} \sum_{q=1}^{N_1} \int_0^{T_q^-} e^{-\beta_n(T_q^- - t)} |x| \underline{p}(dt, dx) \\ &= \beta_n \Delta_n e^{\beta_n \Delta_n} \int_0^1 \int_0^t |y| e^{-\beta_n(t-s)} \underline{p}(ds, dy) \underline{p}(dt, dx), \end{aligned}$$

and this term has expectation of order $\beta_n \Delta_n \lambda_n^2 / \beta_n \lesssim \lambda_n^2 \Delta_n$. \square

Lemma 12. *We have*

$$\left| \frac{\sum_{q \in \mathcal{E}_n} \operatorname{sgn}(\Delta X_{T_q}) \Delta_{i(n,q)}^n X}{\int_0^1 \int_{\mathbb{R}} |x| e^{-\beta_n((\lfloor t\Delta_n^{-1} \rfloor + 1)\Delta_n - t)} \underline{p}(dt, dx)} - 1 \right| \lesssim \lambda_n \Delta_n + \sqrt{\Delta_n / \lambda_n}$$

in probability.

Proof. If $\sup_n \lambda_n < \infty$, then $(\int_0^1 \int_{\mathbb{R}} |x| e^{-\beta_n((\lfloor t\Delta_n^{-1} \rfloor + 1)\Delta_n - t)} \underline{p}(dt, dx))^{-1}$ is bounded in probability. Otherwise, using $\int_0^1 e^{-\beta_n((\lfloor t\Delta_n^{-1} \rfloor + 1)\Delta_n - t)} dt = \frac{1 - e^{-\beta_n \Delta_n}}{\beta_n \Delta_n}$, we have

$$\frac{\int_0^1 \int_{\mathbb{R}} |x| e^{-\beta_n((\lfloor t\Delta_n^{-1} \rfloor + 1)\Delta_n - t)} \underline{p}(dt, dx)}{\lambda_n |x| \star \nu(\frac{1 - e^{-\beta_n \Delta_n}}{\beta_n \Delta_n})} \rightarrow 1$$

as $n \rightarrow \infty$ and the result follows by applying Lemma 10 and Lemma 11. \square

Step 2). We are ready to prove Proposition 9. Define the oracle slope

$$\widehat{s}_n^{\text{oracle}} = - \frac{\sum_{q \in \mathcal{E}_n} \operatorname{sgn}(\Delta X_{T_q}) (\Delta_{i(n,q)+1}^n X + 2\Delta_n \sum_{j=1}^{q-1} \Delta X_{T_j})}{\sum_{q \in \mathcal{E}_n} \operatorname{sgn}(\Delta X_{T_q}) \Delta_{i(n,q)}^n X} \mathbf{1}_{\{N_1 > 0\}}.$$

Using the canonical decomposition $X = X^c + Z^\beta$ and the fact that for $i \in \mathcal{A}_n$, we have

$$-\Delta_i^n Z^\beta = (1 - e^{-\beta_n \Delta_n}) Z_{t_{i-1,n},n}^\beta = (1 - e^{-\beta_n \Delta_n}) (\Delta_{i-1}^n Z^\beta + Z_{t_{i-2,n}}^\beta),$$

and we write

$$\widehat{s}_n^{\text{oracle}} = I + II + III,$$

with

$$\begin{aligned} I &= - \frac{\sum_{q \in \mathcal{E}_n} \operatorname{sgn}(\Delta X_{T_q}) \Delta_{i(n,q)+1}^n X^c}{\sum_{q \in \mathcal{E}_n} \operatorname{sgn}(\Delta X_{T_q}) \Delta_{i(n,q)}^n X} \mathbf{1}_{\{N_1 > 0\}}, \\ II &= (1 - e^{-\beta_n \Delta_n}) \frac{\sum_{q \in \mathcal{E}_n} \operatorname{sgn}(\Delta X_{T_q}) \Delta_{i(n,q)}^n Z^\beta}{\sum_{q \in \mathcal{E}_n} \operatorname{sgn}(\Delta X_{T_q}) \Delta_{i(n,q)}^n X} \mathbf{1}_{\{N_1 > 0\}}, \\ III &= (1 - e^{-\beta_n \Delta_n}) \frac{\sum_{q \in \mathcal{E}_n} \operatorname{sgn}(\Delta X_{T_q}) (Z_{t_{i(n,q)-1,n}}^\beta - \frac{2\Delta_n}{1 - e^{-\beta_n \Delta_n}} \sum_{j=1}^{q-1} \Delta X_{T_j})}{\sum_{q \in \mathcal{E}_n} \operatorname{sgn}(\Delta X_{T_q}) \Delta_{i(n,q)}^n X} \mathbf{1}_{\{N_1 > 0\}}. \end{aligned}$$

We study the convergence of each term separately.

Step 3). The term I . From the proof of Lemma 10, we readily have

$$\left| \sum_{q \notin \mathcal{E}_n} \operatorname{sgn}(\Delta X_{T_q}) \Delta_{i(n,q)+1}^n X^c \right| \lesssim \lambda_n \Delta_n$$

in probability, so we shall replace the sum in q over \mathcal{E}_n by the sum in q over $\{1, \dots, N_1\}$ in the following. By Lemma 12, we derive

$$(30) \quad I = IV(1 + (\lambda_n \Delta_n + \sqrt{\Delta_n/\lambda_n}) \mathcal{R}_n^{(1)}) + \Delta_n \mathcal{R}_n^{(2)},$$

with

$$IV = - \frac{\sum_{q=1}^{N_1} \operatorname{sgn}(\Delta X_{T_q}) \Delta_{i(n,q)+1}^n X^c}{\int_0^1 \int_{\mathbb{R}} |x| e^{-\beta_n((\lfloor t \Delta_n^{-1} \rfloor + 1) \Delta_n - t)} \underline{p}(dt, dx)}$$

and where both $\mathcal{R}_n^{(1)}$ and $\mathcal{R}_n^{(2)}$ are bounded in probability. Next, denoting by ΔX_i^j the j -th jump in $[(i-1)\Delta_n, i\Delta_n]$, we have

$$\begin{aligned} \sum_{q=1}^{N_1} \operatorname{sgn}(\Delta X_{T_q}) \Delta_{i(n,q)+1}^n X^c &= \sum_{i=2}^n \operatorname{sgn}(\Delta X_{i-1}^1) \Delta_i^n X^c \mathbf{1}_{\{\Delta_{i-1}^n N=1\}} + \sum_{i=2}^n \sum_{j=2}^{\Delta_{i-1}^n N} \operatorname{sgn}(\Delta X_i^j) \Delta_i^n X^c \mathbf{1}_{\{\Delta_i^n N \geq 2\}} \\ &= \sum_{i=2}^n \operatorname{sgn}(\Delta X_{i-1}^1) \Delta_i^n X^c \mathbf{1}_{\{\Delta_{i-1}^n N=1\}} + \lambda_n \Delta_n \mathcal{R}_n^{(3)}, \end{aligned}$$

where $\mathcal{R}_n^{(3)}$ is bounded in probability, using the same argument as in Lemma 11. It is not difficult to see that $\sum_{i=2}^n \operatorname{sgn}(\Delta X_{i-1}^1) \Delta_i^n X^c \mathbf{1}_{\{\Delta_{i-1}^n N \geq 1\}}$ is centred with variance $\lambda_n \Delta_n \int_{\Delta_n}^1 \sigma_s^2 ds$ up to an error of order $(\lambda_n \Delta_n)^2$. Therefore, when $\sup_n \lambda_n < \infty$, we have that IV is of order $\sqrt{\Delta_n}$ and of order $\sqrt{\Delta_n/\lambda_n}$ otherwise, using that $\int_0^1 \int_{\mathbb{R}} |x| e^{-\beta_n((\lfloor t \Delta_n^{-1} \rfloor + 1) \Delta_n - t)} \underline{p}(dt, dx)$ is equivalent to $\lambda_n \Delta_n |x| \star \nu \left(\frac{1 - e^{-\beta_n \Delta_n}}{\beta_n \Delta_n} \right)$. We thus obtain the decomposition

$$(31) \quad \begin{aligned} I &= \left(\frac{\beta_n \Delta_n}{1 - e^{-\beta_n \Delta_n}} \right) \frac{\sqrt{\Delta_n} \int_0^1 \sigma_s^2 ds}{\sqrt{\lambda_n} |x| \star \nu} \mathcal{J}_3^{(n)} \\ &= e^{-\beta_n \Delta_n} \beta_n \Delta_n \mathcal{V}_3^{(n)} \mathcal{J}_3^{(n)}, \end{aligned}$$

where $\mathcal{J}_3^{(n)}$ is bounded in probability and $\mathcal{V}_3^{(n)}$ is defined in the statement of Proposition 9. We investigate further the convergence of $\mathcal{J}_3^{(n)}$. Define

$$V_n^2 = \sum_{i=2}^n \mathbb{E}[(\Delta_i^n X^c)^2 \mathbf{1}_{\{\Delta_{i-1}^n N=1\}} | \mathcal{F}_{t_{i-1}, n}^W].$$

Clearly, $\sum_{i=2}^n \operatorname{sgn}(\Delta X_{i-1}^1) \Delta_i^n X^c \mathbf{1}_{\{\Delta_{i-1}^n N=1\}}$ is centred and we claim that for every $\eta > 0$:

$$(32) \quad \sum_{i=2}^n \mathbb{E}[V_n^{-2} (\Delta_i^n X^c)^2 \mathbf{1}_{\{\Delta_{i-1}^n N=1\}} \mathbf{1}_{\{V_n^{-1} |\Delta_i^n X^c| \geq \eta\}} \mathbf{1}_{\{N_1 \geq 1\}}] \rightarrow 0$$

as $n \rightarrow \infty$. Indeed, applying successively Cauchy-Schwarz's, Markov's and Burkholder-Davis-Gundy's inequality, we obtain

$$\begin{aligned} \mathbb{E}[(\Delta_i^n X^c)^2 \mathbf{1}_{\{V_n^{-1} |\Delta_i^n X^c| \geq \eta\}} | \mathcal{F}_1^N] &\leq \mathbb{E}[(\Delta_i^n X^c)^4]^{1/2} \mathbb{P}[|\Delta_i^n X^c| \geq \eta V_n | \mathcal{F}_1^N]^{1/2} \\ &\leq \mathbb{E}[(\Delta_i^n X^c)^4]^{1/2} \frac{\mathbb{E}[(\Delta_i^n X^c)^2]^{1/2}}{\eta V_n} \lesssim \eta^{-1} V_n^{-1} \Delta_n^{3/2}. \end{aligned}$$

Next $V_n^2 = \sum_{i=2}^n \mathbf{1}_{\{\Delta_{i-1}^n N=1\}} \int_{t_{i-1},n}^{t_{i,n}} \sigma_s^2 ds \geq \underline{\sigma}^2 \Delta_n \sum_{i=2}^n \mathbf{1}_{\{\Delta_{i-1}^n N=1\}} = \underline{\sigma}^2 \Delta_n \underline{V}_n^2$ say. Summing up and taking expectation, it follows that

$$\begin{aligned} \sum_{i=2}^n \mathbb{E}[V_n^{-2} (\Delta_i^n X^c)^2 \mathbf{1}_{\{\Delta_{i-1}^n N=1\}} \mathbf{1}_{\{V_n^{-1} |\Delta_i^n X^c| \geq \eta\}}] &\lesssim \eta^{-1} \Delta_n^{3/2} \sum_{i=2}^n \mathbb{E}[V_n^{-3} \mathbf{1}_{\{\Delta_{i-1}^n N=1\}} \mathbf{1}_{\{N_1 \geq 1\}}] \\ &\lesssim \mathbb{E}[V_n^{-1} \mathbf{1}_{\{N_1 \geq 1\}}] \leq \mathbb{E}[(V_n^2)^{-1} \mathbf{1}_{\{N_1 \geq 1\}}]^{1/2} \end{aligned}$$

by Jensen's inequality. Since \underline{V}_n^2 has a Binomial distribution with parameters $(n-1, \lambda_n \Delta_n e^{-\lambda_n \Delta_n})$, we have that $\underline{V}_n^2 \rightarrow \infty$ in probability since $\lambda_n \rightarrow \infty$ and is bounded below on $\{N_1 \geq 1\}$ which has probability that converges to one, the Lindeberg condition (32) follows by dominated convergence and we further infer

$$\frac{1}{V_n} \sum_{i=2}^n \Delta_i^n X^c \operatorname{sgn}(\Delta_j^1 X) \mathbf{1}_{\{\Delta_{i-1}^n N=1\}} \rightarrow \mathcal{N}(0, 1)$$

in distribution as $n \rightarrow \infty$. Observing that $\frac{V_n^2}{\lambda_n \Delta_n} \rightarrow \int_0^1 \sigma_s^2 ds$ in probability, in view of (30), we conclude

$$\frac{1}{\sqrt{\lambda_n \Delta_n} \int_0^1 \sigma_s^2 ds} \sum_{q=1}^{N_1} \operatorname{sgn}(\Delta X_{T_q}) \Delta_{i(n,q)}^n X^c \rightarrow \mathcal{N}(0, 1)$$

in distribution as $n \rightarrow \infty$ and likewise for $\mathcal{J}_3^{(n)}$ in view of (31).

Step 4). The term II . We write, using the proof of Lemma 10 and Lemma 12,

$$\begin{aligned} &(1 - e^{-\beta_n \Delta_n})^{-1} II \\ &= 1 - \frac{\sum_{q \in \mathcal{E}_n} \operatorname{sgn}(\Delta X_{T_q}) \Delta_{i(n,q)}^n X^c}{\sum_{q \in \mathcal{E}_n} \operatorname{sgn}(\Delta X_{T_q}) \Delta_{i(n,q)}^n X} \\ &= 1 - \frac{\sum_{q \in \mathcal{E}_n} \operatorname{sgn}(\Delta X_{T_q}) \Delta_{i(n,q)}^n X^c}{\int_0^1 \int_{\mathbb{R}} |x| e^{-\beta_n ((\lfloor t \Delta_n^{-1} \rfloor + 1) \Delta_n - t)} \underline{p}(dt, dx)} (1 + \max\{\lambda_n \Delta_n, \sqrt{\lambda_n / \Delta_n}\} \mathcal{R}_n^{(1)}) \\ &= 1 - \left(\frac{\sum_{q=1}^{N_1} \operatorname{sgn}(\Delta X_{T_q}) \Delta_{i(n,q)}^n X^c}{\int_0^1 \int_{\mathbb{R}} |x| e^{-\beta_n ((\lfloor t \Delta_n^{-1} \rfloor + 1) \Delta_n - t)} \underline{p}(dt, dx)} + \Delta_n \mathcal{R}_n^{(2)} \right) (1 + \max\{\lambda_n \Delta_n, \sqrt{\lambda_n / \Delta_n}\} \mathcal{R}_n^{(1)}) \end{aligned}$$

where $\mathcal{R}_n^{(1)}$ and $\mathcal{R}_n^{(2)}$ are bounded in probability. By Step 2), we know that

$$\begin{aligned} \frac{\sum_{q=1}^{N_1} \operatorname{sgn}(\Delta X_{T_q}) \Delta_{i(n,q)}^n X^c}{\int_0^1 \int_{\mathbb{R}} |x| e^{-\beta_n ((\lfloor t \Delta_n^{-1} \rfloor + 1) \Delta_n - t)} \underline{p}(dt, dx)} &= \left(\frac{\beta_n \Delta_n}{1 - e^{-\beta_n \Delta_n}} \right) \frac{\sqrt{\Delta_n} \int_0^1 \sigma_s^2 ds}{\sqrt{\lambda_n |x| \star \nu}} \mathcal{U}^{(n)} \\ &= e^{-\beta_n \Delta_n} \left(\frac{\beta_n \Delta_n}{1 - e^{-\beta_n \Delta_n}} \right) \mathcal{V}_4^{(n)}, \end{aligned}$$

where $\mathcal{U}^{(n)}$ is bounded in probability and asymptotically normal if $\lambda_n \rightarrow \infty$ and $\mathcal{V}_4^{(n)}$ is defined in the statement of Proposition 9. Finally, we have proved

$$II = (1 - e^{-\beta_n \Delta_n}) \left(1 + e^{-\beta_n \Delta_n} \left(\frac{\beta_n \Delta_n}{1 - e^{-\beta_n \Delta_n}} \right) \mathcal{V}_4^{(n)} \mathcal{J}_4^{(n)} \right)$$

where $\mathcal{J}_4^{(n)}$ is bounded in probability and asymptotically normal if $\lambda_n \rightarrow \infty$.

Step 4'). It is easily shown that $\sum_{i=2}^n \mathbb{E}[(\Delta_i^n X^c)^2 \mathbf{1}_{\{\Delta_{i-1}N=1\}} \mathbf{1}_{\{\Delta_i^n N=1\}} \mathbf{1}_{\{N_1 \geq 1\}} | \mathcal{F}_{t_{i-1},n}] \rightarrow 0$ if $\lambda_n \rightarrow \infty$ as $n \rightarrow \infty$, so we actually have from Step 3) and Step 4) the joint convergence $(\mathcal{J}_n^{(3)}, \mathcal{J}_n^{(4)}) \rightarrow \mathcal{N}(0, \text{Id}_{\mathbb{R}^2})$ in distribution as $\lambda_n \rightarrow \infty$.

Step 5). The term *III*. By the proof of Lemma 10 and Lemma 12, we have

$$\begin{aligned} & (1 - e^{-\beta_n \Delta_n})^{-1} III \\ &= \frac{\sum_{q \in \mathcal{E}_n} \text{sgn}(\Delta X_{T_q}) (Z_{t_{i(n,q)-1,n}}^\beta - \frac{2\Delta_n}{1-e^{-\beta_n \Delta_n}} \sum_{j=1}^{q-1} \Delta X_{T_j})}{\int_0^1 \int_{\mathbb{R}} |x| e^{-\beta_n (\lfloor t \Delta_n^{-1} \rfloor + 1) \Delta_n - t} \underline{p}(dt, dx)} (1 + \max\{\lambda_n \Delta_n, \sqrt{\Delta_n / \lambda_n}\} \mathcal{R}_n^{(1)}) \\ &= \left(\frac{\sum_{q=1}^{N_1} \text{sgn}(\Delta X_{T_q}) (Z_{t_{i(n,q)-1,n}}^\beta - \frac{2\Delta_n}{1-e^{-\beta_n \Delta_n}} \sum_{j=1}^{q-1} \Delta X_{T_j})}{\int_0^1 \int_{\mathbb{R}} |x| e^{-\beta_n (\lfloor t \Delta_n^{-1} \rfloor + 1) \Delta_n - t} \underline{p}(dt, dx)} + \frac{\lambda_n^2}{\beta_n} \Delta_n \mathcal{R}_n^{(2)} \right) (1 + \max\{\lambda_n \Delta_n, \sqrt{\Delta_n / \lambda_n}\} \mathcal{R}_n^{(1)}) \end{aligned}$$

where $\mathcal{R}_n^{(1)}$ and $\mathcal{R}_n^{(2)}$ are bounded in probability. Indeed, in the same way as for Lemma 10, we have

$$\begin{aligned} & \left| \sum_{q \notin \mathcal{E}_n} \text{sgn}(\Delta X_{T_q}) (Z_{t_{i(n,q)-1,n}}^\beta - \frac{2\Delta_n}{1-e^{-\beta_n \Delta_n}} \sum_{j=1}^{q-1} \Delta X_{T_j}) \right| \\ & \leq \sum_{q \notin \mathcal{E}_n} (|Z_{t_{i(n,q)-1,n}}^\beta| + \frac{2\Delta_n}{1-e^{-\beta_n \Delta_n}} \sum_{j=1}^{q-1} |\Delta X_{T_j}|) \mathbf{1}_{\{T_{q+1} - T_q < 2\Delta_n\}} \lesssim \frac{\lambda_n^3 \Delta_n}{\beta_n} \end{aligned}$$

in probability, as follows from (29) and the computations of Lemma 10. When exactly one jump occurs in $(i(n, q) - 1, i(n, q)]$, we have $Z_{t_{i(n,q)-1,n}}^\beta = e^{-\beta_n (t_{i(n,q)-1,n} - T_q)} Z_{T_q-}^\beta$ and

$$\begin{aligned} & \sum_{q=1}^{N_1} \text{sgn}(\Delta X_{T_q}) (Z_{t_{i(n,q)-1,n}}^\beta - \frac{2\Delta_n}{1-e^{-\beta_n \Delta_n}} \sum_{j=1}^{q-1} \Delta X_{T_j}) \\ &= \int_{0 \leq s < t \leq 1} \int_{\mathbb{R}^2} y \text{sgn}(x) (e^{-\beta_n (t-s)} e^{-\beta_n (\lfloor t \Delta_n^{-1} \rfloor \Delta_n - t)} - \frac{2\Delta_n}{1-e^{-\beta_n \Delta_n}}) \underline{p}(ds, dy) \underline{p}(dt, dx) + \frac{\lambda_n^3 \Delta_n}{\beta_n} \mathcal{R}_n^{(3)}, \end{aligned}$$

where the remainder term $\mathcal{R}_n^{(3)}$ is bounded in probability and accounts for the case where more than one jump occurs in the intervals $(i(n, q) - 1, i(n, q)]$. By Fubini's theorem, the main term splits into $\mathcal{M}_n^{(1)} + \mathcal{M}_n^{(2)} + \mathcal{M}_n^{(3)}$, with

$$\begin{aligned} \mathcal{M}_n^{(1)} &= \lambda_n^2 (x \star \nu) (\text{sgn}(x) \star \nu) \int_{0 \leq s < t \leq 1} (e^{-\beta_n (t-s)} e^{-\beta_n (\lfloor t \Delta_n^{-1} \rfloor \Delta_n - t)} - \frac{2\Delta_n}{1-e^{-\beta_n \Delta_n}}) dt ds \\ &= \frac{\lambda_n^2}{\beta_n} (x \star \nu) (\text{sgn}(x) \star \nu) \left(\frac{e^{\beta_n \Delta_n} - 1}{\beta_n \Delta_n} - \Delta_n \frac{1 - e^{-\beta_n}}{1 - e^{-\beta_n \Delta_n}} - \frac{\beta_n \Delta_n}{1 - e^{-\beta_n \Delta_n}} \right), \\ \mathcal{M}_n^{(2)} &= \lambda_n \int_0^1 \int_{\mathbb{R}} g_n^{(1)}(t, x) (\underline{p} - \lambda_n \underline{q})(dt, dx), \\ \mathcal{M}_n^{(3)} &= \int_{[0,1]^2 \times \mathbb{R}^2} g_n^{(2)}(t, x, s, y) (\underline{p} - \lambda_n \underline{q})(ds, dy) (\underline{p} - \lambda_n \underline{q})(dt, dx), \end{aligned}$$

with

$$g_n^{(1)}(t, x) = x(\operatorname{sgn}(y) \star \nu) \int_{0 \leq s < t \leq 1} (e^{-\beta_n(s-t)} e^{-\beta_n(\lfloor s\Delta_n^{-1} \rfloor \Delta_n - s)} - \frac{2\Delta_n}{1-e^{-\beta_n\Delta_n}}) ds \\ + (y \star \nu) \operatorname{sgn}(x) \int_{0 \leq s < t \leq 1} (e^{-\beta_n(t-s)} e^{-\beta_n(\lfloor t\Delta_n^{-1} \rfloor \Delta_n - t)} - \frac{2\Delta_n}{1-e^{-\beta_n\Delta_n}}) ds$$

and

$$g_n^{(2)}(t, x, s, y) = y \operatorname{sgn}(x) \mathbf{1}_{\{t > s\}} (e^{-\beta_n(t-s)} e^{-\beta_n(\lfloor t\Delta_n^{-1} \rfloor \Delta_n - t)} - \frac{2\Delta_n}{1-e^{-\beta_n\Delta_n}}).$$

By standard yet tedious computations, evaluating centred terms and their variances in terms of asymptotics in n , we can infer from the previous representation that

$$\frac{\sum_{q=1}^{N_1} \operatorname{sgn}(\Delta X_{T_q}) (Z_{t_{i(n,q)-1,n}}^\beta - \frac{2\Delta_n}{1-e^{-\beta_n\Delta_n}} \sum_{j=1}^{q-1} \Delta X_{T_j})}{\int_0^1 \int_{\mathbb{R}} |x| e^{-\beta_n(\lfloor t\Delta_n^{-1} \rfloor + 1)\Delta_n - t} \underline{p}(dt, dx)} \\ = \frac{\mathcal{M}_n^{(1)}}{\lambda_n(|x| \star \nu) \frac{1-e^{-\beta_n\Delta_n}}{\beta_n\Delta_n}} + \frac{\lambda_n \int_0^1 \int_{\mathbb{R}} g_n^{(3)}(t, x) (\underline{p} - \lambda_n \underline{q})(dt, dx)}{\int_0^1 \int_{\mathbb{R}} |x| e^{-\beta_n(\lfloor t\Delta_n^{-1} \rfloor + 1)\Delta_n - t} \underline{p}(dt, dx)} \\ + \frac{\beta_n\Delta_n}{1-e^{-\beta_n\Delta_n}} \min \left\{ \sqrt{\frac{x^2 \star \nu}{2(|x| \star \nu)^2 \beta_n} \frac{1-e^{-2\beta_n\Delta_n}}{2\beta_n\Delta_n}}, \lambda_n \beta_n^{-1} \right\} \mathcal{J}_n^{(2)},$$

where $\mathcal{J}_n^{(2)}$ is bounded in probability and $g_n^{(3)}(t, x) = -\lambda_n^{-2} \frac{\mathcal{M}_n^{(1)}}{|y| \star \nu \left(\frac{1-e^{-\beta_n\Delta_n}}{\beta_n\Delta_n} \right)} |x| e^{-\beta_n(\lfloor t\Delta_n^{-1} \rfloor + 1)\Delta_n - t} + g_n^{(1)}(t, x)$. It is not difficult to check that

$$\frac{\lambda_n \int_0^1 \int_{\mathbb{R}} g_n^{(3)}(t, x) (\underline{p} - \lambda_n \underline{q})(dt, dx)}{\int_0^1 \int_{\mathbb{R}} |x| e^{-\beta_n(\lfloor t\Delta_n^{-1} \rfloor + 1)\Delta_n - t} \underline{p}(dt, dx)} = e^{-\beta_n\Delta_n} \mathcal{V}_n^{(1)} \mathcal{J}_n^{(1)}$$

where $\mathcal{V}_n^{(1)}$ is defined in the statement of Proposition 9 and $\mathcal{J}_n^{(1)}$ is bounded in probability. It follows that

$$\frac{1 - e^{-\beta_n\Delta_n}}{\beta_n\Delta_n} \frac{\sum_{q=1}^{N_1} \operatorname{sgn}(\Delta X_{T_q}) (Z_{t_{i(n,q)-1,n}}^\beta - \frac{2\Delta_n}{1-e^{-\beta_n\Delta_n}} \sum_{j=1}^{q-1} \Delta X_{T_j})}{\int_0^1 \int_{\mathbb{R}} |x| e^{-\beta_n(\lfloor t\Delta_n^{-1} \rfloor + 1)\Delta_n - t} \underline{p}(dt, dx)} \\ = \frac{\mathcal{M}_n^{(1)}}{\lambda_n(|x| \star \nu)} + e^{-\beta_n\Delta_n} \mathcal{V}_n^{(1)} \mathcal{J}_n^{(1)} + \min \left\{ \sqrt{\frac{x^2 \star \nu}{2(|x| \star \nu)^2 \beta_n} \frac{1 - e^{-2\beta_n\Delta_n}}{2\beta_n\Delta_n}}, \lambda_n \beta_n^{-1} \right\} \mathcal{J}_n^{(2)}.$$

We eventually obtain the decomposition

$$III = \beta_n \Delta_n \underline{\mathcal{M}}_n + \beta_n \Delta_n e^{-\beta_n\Delta_n} \mathcal{V}_n^{(1)} \mathcal{J}_n^{(1)} + \beta_n \Delta_n e^{-\beta_n\Delta_n} \mathcal{V}_n^{(2)} \mathcal{J}_n^{(2)}$$

with $\underline{\mathcal{M}}_n = \frac{\lambda_n}{\beta_n} \frac{(x \star \nu)(\operatorname{sgn}(x) \star \nu)}{|x| \star \nu} \left(\frac{e^{\beta_n\Delta_n} - 1}{\beta_n\Delta_n} - \Delta_n \frac{1 - e^{-\beta_n\Delta_n}}{1 - e^{-\beta_n\Delta_n}} - \frac{\beta_n\Delta_n}{1 - e^{-\beta_n\Delta_n}} \right)$ and $\mathcal{V}_n^{(2)}$ defined in the statement of Proposition 9.

Step 5') Using classical results of [21, Theorem 3], one can further investigate the asymptotic distribution of $(\mathcal{J}_n^{(1)}, \mathcal{J}_n^{(2)})$ under the additional assumption $|x|^3 \star \nu < \infty$ assuring the existence of $\int_0^1 \int_{\mathbb{R}} g_n^3(t, x) \nu(dx) dt$ and the condition

$$(\operatorname{sgn}(x) \star \nu)^2 (|x|^2 \star \nu) + (x \star \nu)^2 - 2(\operatorname{sgn}(x) \star \nu)(|x|^2 \star \nu) \neq 0,$$

otherwise the term $\mathcal{V}_n^{(1)}$ becomes negligible. We then have that $\mathcal{J}_n^{(1)}$ is asymptotically normal, and under the stronger assumption $|x|^4 \star \nu < \infty$, we even have the convergence of $(\mathcal{J}_n^{(1)}, \mathcal{J}_n^{(2)})$ towards a standard two-dimensional Gaussian distribution. We omit the details.

Step 6). Combining Step 2) and the results of Steps 3), 4) and 5), we obtain

$$\widehat{s}_n^{\text{oracle}} = (1 - e^{-\beta_n \Delta_n}) + \beta_n \Delta_n \underline{\mathcal{M}}_n + \beta_n \Delta_n e^{-\beta_n \Delta_n} \mathcal{V}_n^T \mathcal{J}_n$$

with the notation introduced in the statement of Proposition 9. Now, let $\widehat{\beta}_n^{\text{oracle}} = -\frac{1}{\Delta_n} \log(\max\{1 - \widehat{s}_n^{\text{oracle}}, \Delta_n\})$. For n large enough, a first-order Taylor's expansion entails

$$\widehat{\beta}_n^{\text{oracle}} = \beta_n + \beta_n e^{\beta_n \Delta_n} (\underline{\mathcal{M}}_n + e^{-\beta_n \Delta_n} \mathcal{V}_n^T \mathcal{J}_n)$$

where the random vector $\underline{\mathcal{J}}_n$ has the same asymptotic properties as \mathcal{J}_n . Setting $\mathcal{M}_n^{\text{oracle}} = e^{\beta_n \Delta_n} \underline{\mathcal{M}}_n$, and checking that all the terms have the right order, we obtained the desired result and Proposition 9 is proved.

Completion of proof of Theorem 6. We start by a simple approximation result.

Lemma 13. *We have*

$$\mathbb{P}(\text{sgn}(\Delta X_{T_q}) = \text{sgn}(\Delta_{i(n,q)}^n X) \text{ for every } 1 \leq q \leq N_1) \rightarrow 1.$$

Proof. Note first that $\mathbb{P}(\Omega_n \cap \{\sup_{|t-s| < 2\Delta_n} |N_t - N_s| = 1\}) \rightarrow 1$ if $\lambda_n^2 \Delta_n \rightarrow 0$. We first work under Assumption 3 (I). In this case, we have $\beta_n \Delta_n \rightarrow 0$ and we claim that

$$(33) \quad \mathbb{P}\left(\min_{1 \leq q \leq N_1} \frac{\Delta_{i(n,q)}^n X}{\Delta X_{T_q}} \leq 0\right) \rightarrow 0$$

from which Lemma 13 readily follows. For all $q \geq 1$, we have

$$\begin{aligned} \frac{\Delta_{i(n,q)}^n X}{\Delta X_{T_q}} &= 1 + \frac{\Delta_{i(n,q)}^n X^c}{\Delta X_{T_q}} - \frac{(1 - e^{-\beta_n \Delta_n}) Z_{i(n,q)-1}^\beta}{\Delta X_{T_q}} \\ &\geq 1 - \frac{|\Delta_{i(n,q)}^n X^c|}{|\Delta X_{T_q}|} - \frac{\beta_n \Delta_n |Z_{i(n,q)-1}^\beta|}{|\Delta X_{T_q}|} \\ &\geq 1 - \frac{\max_{i \in \mathcal{A}_n^c} |\Delta_i^n X^c|}{\min_{1 \leq i \leq N_1} |\Delta X_{T_i}|} - \frac{\beta_n \Delta_n \max_{1 \leq i \leq n} |Z_{t_i, n}^\beta|}{\min_{1 \leq i \leq n} |\Delta X_{T_i}|}. \end{aligned}$$

Thus,

$$\mathbb{P}\left(\min_{1 \leq q \leq N_1} \frac{\Delta_{i(n,q)}^n X}{\Delta X_{T_q}} \leq 0\right) \leq \mathbb{P}\left(\max_{i \in \mathcal{A}_n^c} |\Delta_i^n X^c| + \beta_n \Delta_n \max_{1 \leq i \leq n} |Z_{t_i, n}^\beta| \geq \min_{1 \leq i \leq N_1} |\Delta X_{T_i}| \right)$$

and this term converges to 0 in the same way as the terms (19) and (20). The convergence (33) follows and Lemma 13 is proved. The case where Assumption 3 (II) is fulfilled corresponds to showing the convergence (27) and the proof follows likewise. \square

We are ready to prove Theorem 6.

Step 1). We prove the result on $\Omega_n \cap \{\text{sgn}(\Delta X_{T_q}) = \text{sgn}(\Delta_{i(n,q)}^n X) \text{ for every } 1 \leq q \leq N_1\} \cap \{\sup_{|t-s| < 2\Delta_n} |N_t - N_s| = 1\}$ thanks to Lemma 13. On this event, the quantity

$$\widehat{s}_n = -\frac{\sum_{q=1}^{\widehat{\lambda}_n} \text{sgn}(\Delta_{\mathcal{I}_n(q)}^n X) (\Delta_{\mathcal{I}_n(q)+1}^n X + 2\Delta_n \sum_{j=1}^{q-1} \Delta_{\mathcal{I}_n(j)}^n X)}{\sum_{q=1}^{\widehat{\lambda}_n} \text{sgn}(\Delta_{\mathcal{I}_n(q)}^n X) \Delta_{\mathcal{I}_n(q)}^n X}$$

is equal to

$$-\frac{\sum_{q=1}^{N_1} \operatorname{sgn}(\Delta_{i(n,q)}^n X) (\Delta_{i(n,q)+1}^n X + 2\Delta_n \sum_{j=1}^{q-1} \Delta_{i(n,j)}^n X)}{\sum_{q=1}^{N_1} \operatorname{sgn}(\Delta_{i(n,q)}^n X) \Delta_{i(n,q)}^n X}$$

and

$$\widehat{s}_n^{\text{oracle}} = -\frac{\sum_{q=1}^{N_1} \operatorname{sgn}(\Delta X_{T_q}) (\Delta_{i(n,q)+1}^n X + 2\Delta_n \sum_{j=1}^{q-1} \Delta X_{T_j})}{\sum_{q=1}^{N_1} \operatorname{sgn}(\Delta X_{T_q}) \Delta_{i(n,q)}^n X}.$$

It follows that $\widehat{s}_n = \widehat{s}_n^{\text{oracle}} + \mathcal{R}_n^{(1)}$, with

$$\mathcal{R}_n^{(1)} = 2\Delta_n \frac{\sum_{q=1}^{N_1} \operatorname{sgn}(\Delta X_{T_q}) \sum_{j=1}^{q-1} (\Delta X_{T_j} - \Delta_{i(n,j)}^n X)}{\sum_{q=1}^{N_1} \operatorname{sgn}(\Delta X_{T_q}) \Delta_{i(n,q)}^n X} = I + II,$$

with

$$I = 2\Delta_n \frac{\sum_{q=1}^{N_1} \operatorname{sgn}(\Delta X_{T_q}) \sum_{j=1}^{q-1} \Delta_{i(n,j)}^n \int_0^t (1 - e^{-\beta_n(t-s)}) x \underline{p}(dt, dx)}{\sum_{q=1}^{N_1} \operatorname{sgn}(\Delta X_{T_q}) \Delta_{i(n,q)}^n X}$$

and

$$II = -2\Delta_n \frac{\sum_{q=1}^{N_1} \operatorname{sgn}(\Delta X_{T_q}) \sum_{j=1}^{q-1} \Delta_{i(n,j)}^n X^c}{\sum_{q=1}^{N_1} \operatorname{sgn}(\Delta X_{T_q}) \Delta_{i(n,q)}^n X}.$$

say.

Step 2). We quickly study each term separately. The term I further splits into $I = III + IV$, with

$$III = 2\Delta_n \frac{\sum_{q=1}^{N_1} \operatorname{sgn}(\Delta X_{T_q}) \sum_{j=1}^{q-1} \Delta X_{T_j} (1 - e^{-\beta_n(t_{i(n,j),n} - T_j)})}{\sum_{q=1}^{N_1} \operatorname{sgn}(\Delta X_{T_q}) \Delta_{i(n,q)}^n X},$$

$$IV = 2\Delta_n \frac{\sum_{q=1}^{N_1} \operatorname{sgn}(\Delta X_{T_q}) \sum_{j=1}^{q-1} Z_{t_{i(n,j)-1,n}} (1 - e^{-\beta_n \Delta_n})}{\sum_{q=1}^{N_1} \operatorname{sgn}(\Delta X_{T_q}) \Delta_{i(n,q)}^n X}.$$

With the same kind of arguments as developed in the proof of Proposition 9, it is not difficult to see that $|IV| \lesssim \Delta_n^2 \lambda_n^2$ in probability. For the term III , we use the same kind of arguments and obtain

$$III = \widetilde{\mathcal{M}}_n + \beta_n \Delta_n (\sqrt{\lambda_n} \Delta_n + \sqrt{\Delta_n} \beta_n^{-1}) \mathcal{R}_n^{(2)},$$

with

$$\widetilde{\mathcal{M}}_n = \frac{2\Delta_n \lambda_n \operatorname{sgn}(x) \star \nu(x \star \nu) \left(\frac{1}{2} - \frac{1 - e^{-\beta_n \Delta_n}}{2\beta_n \Delta_n} + \frac{1 + e^{-\beta_n \Delta_n}}{2\beta_n} - \frac{1 - e^{-\beta_n \Delta_n}}{\beta_n^2 \Delta_n} \right)}{(|x| \star \nu) \frac{1 - e^{-\beta_n \Delta_n}}{\beta_n \Delta_n}}$$

and $\mathcal{R}_n^{(2)}$ is bounded in probability. We also have $|II| \lesssim \Delta_n^{3/2}$ in probability. We omit the details.

Step 3). Finally we obtain the decomposition

$$\widehat{s}_n = \widehat{s}_n^{\text{oracle}} + \widetilde{\mathcal{M}}_n + \beta_n \Delta_n (\sqrt{\lambda_n} \Delta_n + \sqrt{\Delta_n} \beta_n^{-1}) \mathcal{R}_n^{(2)}$$

and we conclude by applying Proposition 9, replacing \widehat{s}_n by $\widehat{s}_n^{\text{oracle}}$ and studying the order of each term carefully.

Acknowledgements. This research is supported by the department OSIRIS (Optimization, Simulation, RIsk and Statistics for Energy Markets) of EDF and by the FiME (Finance for Energy Markets) Research Initiative which is gratefully acknowledged.

REFERENCES

- [1] René Aid, Luciano Campi, Adrien Nguyen Huu, and Nizar Touzi. A structural risk-neutral model of electricity prices. *International Journal of Theoretical and Applied Finance*, 12(7):925–947, 2009.
- [2] Yacine Aït-Sahalia and Jean Jacod. Testing for jumps in a discretely observed process. *The Annals of Statistics*, 37(1):184–222, 2009.
- [3] Yacine Aït-Sahalia and Jean Jacod. *High-frequency financial econometrics*. Princeton University Press, 2014.
- [4] Ole Eiler Barndorff-Nielsen, Neil Shephard, and Matthias Winkel. Limit theorems for multipower variation in the presence of jumps. *Stochastic processes and their applications*, 116(5):796–806, 2006.
- [5] Fred Espen Benth, Jan Kallsen, and Thilo Meyer-Brandis. A non-Gaussian Ornstein–Uhlenbeck process for electricity spot price modeling and derivatives pricing. *Applied Mathematical Finance*, 14(2):153–169, 2007.
- [6] Fred Espen Benth, Rüdiger Kiesel, and Anna Nazarova. A critical empirical study of three electricity spot price models. *Energy Economics*, 34(5):1589–1616, 2012.
- [7] Hermine Biermé and Agnès Desolneux. A Fourier approach for the level crossings of shot noise processes with jumps. *Journal of Applied Probability*, 49(1):100–113, 2012.
- [8] Alvaro Cartea and Marcelo G Figueroa. Pricing in electricity markets: a mean reverting jump diffusion model with seasonality. *Applied Mathematical Finance*, 12(4):313–335, 2005.
- [9] Rama Cont and Peter Tankov. *Financial modelling with jump processes*, volume 2. CRC press, 2003.
- [10] Olivier Féron and Elias Daboussi. *Commodities, Energy and Environmental Finance*, chapter Calibration of electricity price models, pages 183–207. Springer, 2015.
- [11] Hans Föllmer and Martin Schweizer. Hedging of contingent claims under incomplete information. *Applied stochastic analysis, Stochastic Monographs*, 5:389–414, 1991.
- [12] Hélyette Geman and Andrea Roncoroni. Understanding the fine structure of electricity prices. *The Journal of Business*, 79(3):1225–1261, 2006.
- [13] Jhonny Gonzalez, John Moriarty, and Jan Palczewski. Bayesian calibration and number of jump components in electricity spot price models. *Energy Economics*, 65:375–388, 2017.
- [14] Jean Jacod and Philip Protter. *Discretization of processes*, volume 67. Springer Science & Business Media, 2011.
- [15] Rüdiger Kiesel, Gero Schindlmayr, and Reik H Börger. A two-factor model for the electricity forward market. *Quantitative Finance*, 9(3):279–287, 2009.
- [16] Julio J Lucia and Eduardo S Schwartz. Electricity prices and power derivatives: Evidence from the Nordic power exchange. *Review of derivatives research*, 5(1):5–50, 2002.
- [17] Cecilia Mancini. Estimation of the characteristics of the jumps of a general Poisson-diffusion model. *Scandinavian Actuarial Journal*, 2004(1):42–52, 2004.
- [18] Robert C Merton. Option pricing when underlying stock returns are discontinuous. *Journal of financial economics*, 3(1-2):125–144, 1976.
- [19] Thilo Meyer-Brandis and Peter Tankov. Multi-factor jump-diffusion models of electricity prices. *International Journal of Theoretical and Applied Finance*, 11(5):503–528, 2008.
- [20] Manuel Moreno, Pedro Serrano, and Winfried Stute. Statistical properties and economic implications of jump-diffusion processes with shot-noise effects. *European Journal of Operational Research*, 214(3):656–664, 2011.
- [21] Giovanni Peccati and Murad S Taqqu. Central limit theorems for double Poisson integrals. *Bernoulli*, 14(3):791–821, 2008.
- [22] Philip Protter and Kazuhiro Shimbo. No arbitrage and general semimartingales. In *Markov processes and related topics: a Festschrift for Thomas G. Kurtz*, pages 267–283. Institute of Mathematical Statistics, 2008.
- [23] Thorsten Schmidt. Modelling energy markets with extreme spikes. In *Mathematical Control Theory and Finance*, pages 359–375. Springer, 2008.
- [24] Martin Schweizer. *Handbooks in mathematical finance: Option pricing, interest rates and risk management*, chapter A guided tour through quadratic hedging approaches, pages 538–574. Cambridge University Press, 2001.
- [25] Almut Elisabeth Dorothea Veraart. Inference for the jump part of quadratic variation of Itô semimartingales. *Econometric Theory*, 26(2):331–368, 2010.
- [26] Xavier Warin. Gas storage hedging. In *Numerical Methods in Finance*, pages 421–445. Springer, 2012.

THOMAS DESCHATRE, EDF LAB PARIS-SACLAY AND UNIVERSITÉ PARIS-DAUPHINE PSL, CNRS, CEREMADE, 75016 PARIS, FRANCE

E-mail address: deschatre@ceremade.dauphine.fr

OLIVIER FÉRON, EDF LAB PARIS-SACLAY AND FIME, LABORATOIRE DE FINANCE DES MARCHÉS DE L'ÉNERGIE,
91120 PALAISEAU, FRANCE.

E-mail address: `olivier-2.feron@edf.fr`

MARC HOFFMANN, UNIVERSITÉ PARIS-DAUPHINE PSL, CNRS, CEREMADE, 75016 PARIS, FRANCE. (TO WHOM
CORRESPONDENCE SHOULD BE ADDRESSED.)

E-mail address: `hoffmann@ceremade.dauphine.fr`