NEAR-OPTIMAL ESTIMATION OF JUMP ACTIVITY IN SEMIMARTINGALES

By Adam D. Bull[1]

University of Cambridge

In quantitative finance, we often model asset prices as semimartingales, with drift, diffusion and jump components. The jump activity index measures the strength of the jumps at high frequencies, and is of interest both in model selection and fitting, and in volatility estimation. In this paper, we give a novel estimate of the jump activity, together with corresponding confidence intervals. Our estimate improves upon previous work, achieving near-optimal rates of convergence, and good finite-sample performance in Monte-Carlo experiments.

1. Introduction. In quantitative finance, we often wish to model asset prices, for example, to price options or evaluate investment strategies. Typically, we assume that asset log-prices are given by a semimartingale; in other words, the sum of drift, diffusion and jump processes. In the following, we will be interested in the jump activity index, a parameter which determines the strength of the jump process at high frequencies.

The jump activity is important for two reasons. First, any semimartingale model will make claims about the jump activity; typically, the activity is either assumed known and fixed, or is a free parameter to be estimated. Knowledge of the jump activity thus informs our choice of model, and may allow us to fit it more accurately.

Second, the jump activity controls the difficulty of estimating another parameter of interest, the volatility. This parameter measures the strength of the diffusion component of price movements, and is often a key target for financial modellers. It is known that under high jump activity, the volatility becomes harder to estimate; this problem can be avoided using specialised volatility estimates, but at the cost of making stronger assumptions.

Knowledge of the jump activity is thus important both for the analysis of individual price records, to inform the choice of volatility estimate; and more generally in research, to guide the development of future estimates. In the following, we will therefore investigate the problem of accurately estimating the jump activity.

Previous attempts to estimate the jump activity of semimartingales have either achieved poor rates of convergence, or worked only under restrictive assumptions. In this paper, we will describe a new jump activity estimate, which achieves near-optimal rates of convergence in a general setting, along with improved finite-sample performance. We will further provide limiting distributions for our estimate, validated by Monte-Carlo experiments.

We begin by discussing in more detail the nature of the problem, and relevant work in the literature. We will suppose we have a log-price process given by a semimartingale X_t on $[0, 1]$, and make n observations

$$X_{j/n}, \qquad j = 0, \dots, n - 1.$$

We then define the jump activity index

$$\beta = \inf\{r \in [0, 2] : S(r) < \infty \text{ a.s.}\}, \qquad S(r) = \sum_{s \in [0,1]} |\Delta X_s|^r,$$

letting $\Delta X_s = X_s - X_{s-}$ denote the jumps of X_t, and using the convention $0^0 = 0$.

As semimartingales have finite quadratic variation, we have $S(2) < \infty$ almost surely, and so the jump activity $\beta \in [0, 2]$. When the sample path of X_t has finitely many jumps, $\beta = 0$; when it may have infinitely many jumps, but the jumps are of finite variation, $\beta \in [0, 1]$; and when the jumps may be of infinite variation, $\beta \in [1, 2]$. The more activity X_t has in its small jumps, the larger we will have to choose r to make $S(r)$ finite, and the larger β will be.

From Lemma 3.2.1 of Jacod and Protter (2012), we can equivalently define

$$\beta = \inf\{r \in [0, 2] : I(r) < \infty \text{ a.s.}\}, \qquad I(r) = \int_0^1 \int_{\mathbb{R}} 1 \wedge |x|^r \nu(dx, ds),$$

letting $\nu(dx, ds)$ denote the compensator of the jump measure of X_t. When X_t is a Lévy process, β is thus the Blumenthal–Getoor index [Blumenthal and Getoor (1961)]; for example, if X_t is a stable process, then β is its stability parameter. More generally, β gives an extension of the Blumenthal–Getoor index to semi-martingales.

The jump activity β is thus a parameter of interest when choosing models for the log-price process X_t. Many common models assume either that no jumps are present, or that there are finitely-many jumps almost-surely; in either case, we therefore assume that $\beta = 0$. This includes all Itô process models, as well as the Merton, Kou and Bates models, for example.

Some models allow positive values of β; for example, the (time-changed) normal-inverse Gaussian, Meixner and generalised hyperbolic models assume $\beta = 1$, while the (time-changed) CGMY or tempered-stable model includes β as a free parameter to be estimated. Knowledge of β thus allows us to better decide between competing models, and in the latter case also to fit these models to price data. [For definitions of the models, see Cont and Tankov (2004), Papapantoleon (2008).]

Further interest in the jump activity arises from the problem of volatility estimation. Let X_t^c denote the continuous part of X_t. Then the integrated volatility of X_t over $[0, 1]$, given by the quadratic variation $[X^c]_1$, is a parameter of much interest in options pricing or risk modelling, and its estimation has been extensively studied.

When X_t is continuous, the integrated volatility can be estimated by the observed quadratic variation; however, price data is widely accepted to contain jumps, which must be accounted for explicitly. Methods for doing so include thresholding [Mancini (2001, 2009)], bipower variation [Barndorff-Nielsen and Shephard (2004), Barndorff-Nielsen et al. (2006)], and characteristic functions [Todorov and Tauchen (2012a, 2012b)].

Unfortunately, the convergence rates of these methods suffer when the jumps are of infinite variation. While this can be avoided if we assume the jumps are driven by a stable-like process [Jacod and Todorov (2014)], or that prices are given by a time-changed process [Bull (2014)], it is known that in general, poor rates are unavoidable [Jacod and Reiss (2014)].

When estimating volatility, we would therefore like to know whether the jumps are of infinite variation, and if so, how active they are; equivalently, we would like to know whether β is greater than 1, and if so, by how much. This question is of interest both when choosing an estimator to apply to particular price data, and also more generally when planning research on volatility estimation.

Previous authors have attempted to recover β in a variety of settings, including when no diffusion component is present [Todorov and Tauchen (2010), Woerner (2011), Zhao and Wu (2009)], or when testing if β is greater than zero [Aït-Sahalia and Jacod (2011), Lee and Hannig (2010)] or one [Cont and Mancini (2011)]. In the following, however, we will concentrate on estimating β in general, when a diffusion term may also be present.

In this context, Aït-Sahalia and Jacod (2009) provide an estimate of β based upon jump counting. While Aït-Sahalia and Jacod cannot prove results for all semimartingales, they do provide convergence rates under the additional assumption that the jumps are dominated by a stochastic integral of a stable-like process. Similar assumptions have also been considered by Jing, Kong and Liu (2011) and Jacod and Todorov (2014), for example, and are satisfied by many common models of price data.

Under these conditions, Aït-Sahalia and Jacod (2009) show that their estimate of the jump activity β converges at a rate $n^{-\beta/10}$. Related estimates have been considered also by Jing, Kong and Liu (2011), Aït-Sahalia and Jacod (2012) and Jing et al. (2012); the best convergence is obtained by the latter, who achieve the rate $n^{-\beta/8}$. However, this still falls short of the corresponding lower bound of $n^{-\beta/4} \log(n)^{-(1-\beta/4)}$, given by Aït-Sahalia and Jacod (2012).

If we assume not only stable-like jumps, but also that log-prices are given by a Lévy process, Reiß (2013) shows we can estimate β at the near-optimal rate

$n^{-\beta/4+\varepsilon}$, for any $\varepsilon > 0$. However, the assumption of Lévy behaviour is quite restrictive in a financial context, and unfortunately the approach of Reiß does not easily generalise to semimartingales.

In the following, we will therefore describe a new estimate of the jump activity β, using a multi-scale jump-counting approach. We will show that by combining jump-counting estimates across different time-scales, we will be able to cancel out the bias in these estimates, obtaining improved accuracy.

In a similar setting to that of Aït-Sahalia and Jacod (2009), with no assumption of Lévy behaviour, our estimate will obtain the near-optimal convergence rates $n^{-\beta/4+\varepsilon}$, as well as improved finite-sample performance. We will also give limiting distributions, validated by Monte-Carlo experiments.

In Section 2, we describe our estimates in full, and in Section 3, discuss their theoretical properties. In Section 4, we then perform our Monte-Carlo experiments, and in Section 5, give proofs.

2. Jump activity estimates. We now describe our estimate of the jump activity β. In the following, we will suppose that $\beta > 0$; we note the case $\beta = 0$ can be tested for separately, for example, using the methods of Lee and Hannig (2010) or Aït-Sahalia and Jacod (2011).

Our approach builds upon the work of Aït-Sahalia and Jacod (2009), who estimate β by counting jumps in X_t. The authors define the jump counts

$$\widetilde{A}_n(\tau) = \sum_{j=0}^{n-2} 1_{\tau|X_{(j+1)/n} - X_{j/n}| \geq 1},$$

which for suitable $\tau > 0$, approximate the number of jumps in X_t of size at least τ^{-1}.

For $\rho > 1$, Aït-Sahalia and Jacod then estimate β by

$$\widehat{\beta}_n^{\mathrm{AJ}} = \log_\rho\left(\frac{\widetilde{A}_n(\rho\tau_n)}{\widetilde{A}_n(\tau_n)}\right),$$

using the convention $0/0 = 1$. If the jumps of X_t are dominated by a stochastic integral of a stable-like process, then as $n \to \infty$, for suitable sequences τ_n, we can expect

$$\widetilde{A}_n(\tau_n) \approx C\tau_n^\beta,$$

for some quantity $C > 0$. We would then have that

$$\widehat{\beta}_n^{\mathrm{AJ}} \approx \log_\rho\left(\frac{C(\rho\tau_n)^\beta}{C\tau_n^\beta}\right) = \beta.$$

Unfortunately, Aït-Sahalia and Jacod were not able to provide good convergence rates for this method, as the estimates $\widehat{\beta}_n^{\mathrm{AJ}}$ are too biased when τ_n is large.

In the following, we will therefore provide an improved version of this method, which corrects for the bias in $\widehat{\beta}_n^{AJ}$, achieving near-optimal rates of convergence.

We will use three techniques to correct for this bias. First, will we symmetrise the data, correcting for bias due to high-activity, asymmetrically-distributed jumps. Second, we will smooth the jump counts, correcting for bias due to the roughness of the indicator function $1_{|x|\geq 1}$. Finally, and most importantly, we will eliminate the remaining bias by cancelling between estimates at different time-scales.

We first describe a procedure to symmetrise the process X_t, as given, for example, in Jacod and Todorov (2014). For $j = 0, \ldots, n-3$, we define random variables

$$\Delta X_{j,n} = (X_{(j+2)/n} - X_{(j+1)/n}) - (X_{(j+1)/n} - X_{j/n}).$$

We note that when X_t is a Lévy process, the random variables $\Delta X_{j,n}$ are symmetric, even if the increments of X_t are not. More generally, we may think of the $\Delta X_{j,n}$ as symmetrised increments of the process X_t, across time intervals of length $2/n$.

In the following, we will wish to work with increments of X_t across different time-scales simultaneously. For $k = 0, 1, \ldots, j = 0, \ldots, n - 2k - 1$, we therefore also define random variables

$$\Delta X_{j,k,n} = \sum_{l=0}^{k-1} \Delta X_{j+2l,n}.$$

We can similarly consider the $\Delta X_{j,k,n}$ to be symmetrised increments of X_t, now across time intervals of length $2k/n$.

Next, we will replace the indicator function $1_{|x|\geq 1}$ with a smooth function, similarly to Jing et al. (2012). We will use a smooth function $1 - K(x)$, where the kernel $K : \mathbb{R} \to [0, 1]$ is an even Schwartz function, equal to one in a neighbourhood of the origin. For example, in our experimental results, we will choose

$$K(x) = \begin{cases} 1, & |x| \leq 1, \\ \left(1 + \exp\left(\dfrac{1}{2 - |x|} - \dfrac{1}{|x| - 1}\right)\right)^{-1}, & 1 \leq |x| \leq 2, \\ 0, & |x| \geq 2. \end{cases}$$

We will also fix a constant $m \in \mathbb{N}$, giving the number of time-scales to use for bias correction.

For $\tau > 0$, we then define the jump counts

$$\widehat{A}_n(\tau) = 0 \vee \widehat{A}_n'(\tau), \qquad \widehat{A}_n'(\tau) = \sum_{j=0}^{n-2m-1} \widehat{a}_{j,n}(\tau),$$

where for $j = 0, \ldots, n - 2m - 1$, we set

$$\widehat{a}_{j,n}(\tau) = \sum_{k=1}^{m} w_k\big(1 - K(\tau \Delta X_{j,k,n})\big), \qquad w_k = \frac{(-1)^{k+1}}{2k}\binom{m}{k}.$$

For a constant $\rho > 1$, and sequence $\tau_n > 0$, we finally estimate β by

$$\widehat{\beta}_n = 0 \vee \log_\rho \left(\frac{\widehat{A}_n(\rho \tau_n)}{\widehat{A}_n(\tau_n)} \right) \wedge 2,$$

using the convention $0/0 = 1$.

When $m = 1$, this estimate is similar to the jump-counting estimate of Aït-Sahalia and Jacod (2009): we replace the increments $X_{(j+1)/n} - X_{j/n}$ with symmetrised increments $\Delta X_{j,n}$; replace the indicator function $1_{|x| \geq 1}$ with a smooth function $1 - K(x)$; and clip the estimate $\widehat{\beta}_n$ to the interval $[0, 2]$. When $m > 1$, we additionally replace $\widehat{A}_n(\tau)$ with a linear combination of jump counts across different time-scales, clipped to be nonnegative.

We note the clipping of $\widehat{A}_n(\tau)$ and $\widehat{\beta}_n$ ensures that the estimate $\widehat{\beta}_n$ is always reasonable, even when the jump counts $A'_n(\tau)$ may be inaccurate. While this step makes no contribution to the asymptotic behaviour of $\widehat{\beta}_n$, it does reduce its error in finite time.

In the following sections, we will show that our changes reduce the bias in the estimate $\widehat{\beta}_n$, providing both theoretical and experimental improvements to accuracy. We will also use these results to motivate the selection of parameters in our estimate: the number of time-scales m, inverse jump threshold τ_n, and threshold ratio ρ.

We will further give limiting distributions for $\widehat{\beta}_n$, allowing us to build confidence intervals for β. Define the constants

$$C_{\beta,\rho} = K_{\beta,\rho} / \rho^\beta \log(\rho)^2 K_\beta^2,$$

$$K_\beta = \int_{\mathbb{R}} (1 - K(x)) |x|^{-(1+\beta)} \, dx,$$

$$K_{\beta,\rho} = \int_{\mathbb{R}} (K(x) - K(\rho x))^2 |x|^{-(1+\beta)} \, dx,$$

and for $\widehat{\beta}_n \in (0, 2]$, the random variables

$$\widehat{U}_n(\beta) = \frac{\tau_n^{\widehat{\beta}_n/2}(\widehat{\beta}_n - \beta)}{\widehat{\sigma}_{\rho,n}}, \qquad \widehat{\sigma}_{\rho,n}^2 = \frac{C_{\widehat{\beta}_n,\rho} K_{\widehat{\beta}_n} \tau_n^{\widehat{\beta}_n}}{\widehat{A}_n(\tau_n)}.$$

When $\widehat{\beta}_n = 0$, likewise define

$$\widehat{U}_n(\beta) = -\infty.$$

We note that the random variables $\widehat{U}_n(\beta)$ are always well defined, as $\widehat{\beta}_n$ must lie within $[0, 2]$.

We will be able to show that, under suitable conditions, the standardised errors

$$\widehat{U}_n(\beta) \xrightarrow{d} N(0, 1).$$

We will therefore be able to define γ-level confidence intervals for β,

$$\widehat{I}_n(\gamma) = \{\beta \in (0, 2) : |\widehat{U}_n(\beta)| \leq \Phi^{-1}(\tfrac{1}{2}(1 + \gamma))\},$$

where Φ denotes the standard Gaussian distribution function.

We note that the integrals K_β and $K_{\beta,\rho}$ can usually be computed numerically. In the case where $\widehat{\beta}_n$ is very small but nonzero, the integration of K_β may be slow to converge, and it may be preferable to instead take $\widehat{\beta}_n = 0$. In our experimental tests, we did so for $\widehat{\beta}_n < 10^{-3}$.

3. Theoretical results.

To describe our theoretical results, we must first state our assumptions. The assumptions will be very similar to those made by Jacod and Todorov (2014), and essentially require that the jumps of the log-price process X_t are dominated by a stochastic integral of a stable-like process. Similar assumptions have also been made by Aït-Sahalia and Jacod (2009) and Jing, Kong and Liu (2011), and are satisfied by many common models of price data; we refer to Jacod and Protter (2012) for definitions and notation.

ASSUMPTION 1. We first assume we have a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with filtration $\mathcal{F}_t$, and a jump activity index $\beta \in (0, 2)$. We then assume the log-price process

$$X_t = \int_0^t b_s \, ds + \int_0^t c_s \, dB_s + \int_0^t \gamma_s^+ \, dL_s^+ + \int_0^t \gamma_s^- \, dL_s^- + \int_0^t \int_\mathbb{R} \delta_s(x) \mu(dx, ds),$$

where:

(i) B_t is an adapted Brownian motion;

(ii) the adapted Poisson random measure $\mu(dx, ds)$ has intensity $dx \, ds$, and is independent of B_t;

(iii) the Lévy processes

$$L_t^\pm = \int_0^t \int_\mathbb{R} \delta^\pm(x)(\mu(dx, ds) - 1_{\delta^\pm(x) < 1} \, dx \, ds),$$

for disjointly-supported functions $\delta^\pm(x) \geq 0$, $\int_\mathbb{R} 1 \wedge \delta^\pm(x)^2 \, dx < \infty$;

(iv) the predictable processes b_s and $\gamma_s^\pm$ are locally bounded; and

(v) the predictable function $\delta_s(x)$ has $\int_\mathbb{R} 1 \wedge |\delta_s(x)|^{\upsilon_1} \, dx$ locally bounded, for a parameter $\upsilon_1 \in (0, \beta/2)$.

We additionally assume the volatility process

$$c_t = c_0 + \int_0^t b_s^c \, ds + \int_0^t H_s \, dB_s + \int_0^t H_s' \, dB_s'$$

$$+ \int_0^t \int_\mathbb{R} \delta_s^c(x)(\mu(dx, ds) - 1_{|\delta_s^c(x)| < 1} \, dx \, ds),$$

where:

(i) the adapted Brownian motion B_t' is independent of B_t and $\mu(dx, ds)$;
(ii) the predictable processes b_s^c, H_s and H_s' are locally bounded; and
(iii) the predictable function $\delta_s^c(x)$ has $\int_{\mathbb{R}} 1 \wedge \delta_s^c(x)^2 \, dx$ locally bounded.

We next assume the processes $L_t^\pm$ are close to one-sided β-stable processes. Let

$$F^\pm(U) = \int_{\delta^\pm(x) \in U} dx$$

denote the Lévy measures of the processes $L_t^\pm$, and for $x > 0$, let

$$\overline{F}^\pm(x) = F^\pm\big((x, \infty)\big)$$

denote their upper Lévy distribution functions. We then require that for $x \in (0, 1)$,

$$\big|\overline{F}^\pm(x) - \beta^{-1}x^{-\beta}\big| = O\big(x^{-\upsilon_2}\big),$$

for a parameter $\upsilon_2 < \beta - 1$.

Finally, we assume that the characteristics b, H and $\gamma^\pm$ are smooth in quadratic mean: we assume there are stopping times $T_n \to \infty$, such that for $V = b$, H or $\gamma^\pm$, and any $0 \le t \le t + h \le 1$,

$$\mathbb{E}\big[(V_{(t+h) \wedge T_n} - V_{t \wedge T_n})^2 | \mathcal{F}_t\big] = O(h),$$

uniformly in t.

In other words, we assume that the log-price process X_t and volatility process c_t are Itô semimartingales; that the jumps of X_t are dominated by stochastic integrals against Lévy processes $L_t^\pm$, whose Lévy distribution functions approach those of a β-stable process; and that the drift process b_t, leverage process H_t, and jump integrands $\gamma_t^\pm$ exhibit smoothness behaviour typical of Itô semimartingales.

We note that the jump processes in our assumptions are all described using a Grigelionis representation, as integrals against a common Poisson random measure μ; however, this condition is not restrictive, as any collection of jump processes can be expressed in this form [Jacod and Protter (2012), Theorem 2.1.2]. We likewise note that while our assumptions choose a specific normalisation for the jump processes $L_t^\pm$, this is not restrictive, as the processes can always be rescaled by the terms $\gamma_t^\pm$.

While the driving Lévy processes $L_t^\pm$ must have stable-like behaviour, our model allows for deviations from stability both in the Lévy distribution functions $\overline{F}^\pm$, which must be close to stable only for small jumps; and in the idiosyncratic jumps described by $\delta_t(x)$, which can account for any additional jump activity. The presence of two separate one-sided Lévy processes $L_t^\pm$ also allows us to describe processes with asymmetric jump activity.

We further allow the volatility c_t to contain jumps and leverage, and the other characteristic processes b_t, H_t and $\gamma_t^\pm$ to display a wide range of semimartingale

behaviour. Finally, we note that when the processes $\gamma_t^{\pm}$ are not both almost-surely zero, the parameter β in our assumptions agrees with the jump activity index as defined in the Introduction.

Under these assumptions, we will be able to provide limiting distributions for the estimates $\widehat{\beta}_n$, and standardised errors $\widehat{U}_n(\beta)$; we begin by defining the appropriate notion of convergence. Let $Z_n \in \mathbb{R}^d$ be random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and $Z \in \mathbb{R}^d$ a random variable defined on a suitable extension $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$. We will say the Z_n converge stably in distribution to Z, $Z_n \overset{\text{sd}}{\to} Z$, if

$$\mathbb{E}[Yf(Z_n)] \to \widetilde{\mathbb{E}}[Yf(Z)],$$

for all random variables $Y \in \mathbb{R}$ on Ω, and bounded continuous functions $f : \mathbb{R}^d \to \mathbb{R}$ [Jacod and Protter (2012), Section 2.2.1].

We note that stable convergence in distribution is stronger than the usual convergence in distribution, and requires convergence to hold even after conditioning on events in $\mathcal{F}$. Under suitable conditions, this stronger notion of convergence will allow us to show not only that the estimates $\widehat{\beta}_n$ converge to unbiased Gaussian mixtures, but also that the standardised errors $\widehat{U}_n(\beta)$ converge to standard Gaussians.

To be precise, we first define the jump activity processes

$$\overline{\Gamma}_t = \int_0^t \overline{\gamma}_s \, ds, \qquad \overline{\gamma}_t = \tfrac{1}{2}(|\gamma_t^+|^{\beta} + |\gamma_t^-|^{\beta});$$

we note that the process $\overline{\gamma}_t$ measures the instantaneous stable-like jump activity at time t, while $\overline{\Gamma}_1$ counts the total stable-like jump activity over the interval $[0, 1]$. We then have the following results.

THEOREM 1.   *Under Assumption 1, let $m \in \mathbb{N}$, $\alpha = m/2(m + 1)$, $\tau_n = Cn^{\alpha}$ for some $C > 0$, and $\rho > 0$. Then on the event $\overline{\Gamma}_1 > 0$,*

$$\tau_n^{\beta/2}(\widehat{\beta}_n - \beta) \overset{\text{sd}}{\to} \sigma_{\beta,\rho}Z, \qquad \widehat{\sigma}_{\rho,n}^2 \overset{p}{\to} \sigma_{\beta,\rho}^2, \qquad \widehat{U}_n(\beta) \overset{\text{sd}}{\to} Z,$$

where the variance

$$\sigma_{\beta,\rho}^2 = C_{\beta,\rho}/\overline{\Gamma}_1,$$

and Z is a random variable defined, on a suitable extension of the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, to be standard Gaussian given $\mathcal{F}$.

COROLLARY 1.   *In the setting of Theorem 1, let $\gamma \in (0, 1)$. If $\overline{\Gamma}_1$ is not almost surely zero, then*

$$\mathbb{P}(\beta \in \widehat{I}_n(\gamma)|\overline{\Gamma}_1 > 0) \to \gamma,$$

and on the event $\overline{\Gamma}_1 > 0$, $\widehat{I}_n(\gamma)$ has diameter $O_p(\tau_n^{-\beta/2})$.

We conclude that on the event that X_t includes any stable-like jump activity, the estimate $\widehat{\beta}_n$ converges at a rate

$$n^{-\beta m/4(m+1)};$$

of course, we cannot expect convergence when no stable-like jumps are present. Furthermore, on this event the $\widehat{I}_n(\gamma)$ are indeed γ-level confidence intervals for β, contracting at the correct rate.

A single-scale procedure, with $m = 1$, can thus converge at a rate $n^{-\beta/8}$, recovering the results of Jing et al. (2012). Moreover, by choosing m large enough, a multi-scale procedure can achieve a rate $n^{-\beta/4+\varepsilon}$, for any $\varepsilon > 0$. Indeed, this rate is near-optimal: a corresponding lower bound rate of $n^{-\beta/4} \log(n)^{-(1-\beta/4)}$ is given by Aït-Sahalia and Jacod (2012).

Asymptotically, the rate of convergence will always be improved by choosing m larger, and we should therefore choose m as large as possible. In finite time, however, a larger choice of m may take longer to reach the asymptotic regime, and so be less accurate in practice.

The optimal finite-time choice of m may be difficult to compute, and depends on a number of unknown quantities. However, in our Monte-Carlo experiments, we found the choice $m = 3$ performed well, and should already provide improvements over a single-scale estimate. If more accuracy is desired, users may wish to perform a simulation study to select m, or compare estimates for a number of different choices of m.

Another practical consideration comes from microstructure noise. When observing price data at high frequencies, it is now widely accepted that observations of efficient prices are corrupted by noise. Similarly to Aït-Sahalia and Jacod (2009), for reasonable choices of τ_n the estimates $\widehat{\beta}_n$ are sensitive only to large jumps in prices, and so will not be much affected by noise. However, modifications to account for noise asymptotically are also possible, as in Jing, Kong and Liu (2011) or Bull (2014), and may be left for future work.

4. Monte-Carlo experiments. We now perform Monte-Carlo tests of our multi-scale estimates $\widehat{\beta}_n$, comparing them to the jump-counting estimates $\widehat{\beta}_n^{\mathrm{AJ}}$ of Aït-Sahalia and Jacod (2009). We note that as $\widehat{\beta}_n^{\mathrm{AJ}}$ can sometimes be very large, its RMSE can be distorted by the small chance of a large error. To provide a fair comparison, we will therefore consider the clipped estimates

$$\widetilde{\beta}_n = 0 \vee \widehat{\beta}_n^{\mathrm{AJ}} \wedge 2,$$

defined similarly to $\widehat{\beta}_n$; we note that this clipping can only reduce the error in $\widehat{\beta}_n^{\mathrm{AJ}}$.

We will also compare our confidence intervals $\widehat{I}_n(\gamma)$ to similar ones defined in terms of $\widetilde{\beta}_n$. From Theorem 3 of Aït-Sahalia and Jacod (2009), and arguing as in our Theorem 1, we have that the

$$\widetilde{I}_n(\gamma) = \left\{ \beta \in (0, 2) : \left| \widetilde{U}_n(\beta) \right| \leq \Phi^{-1}\left(\tfrac{1}{2}(1 + \gamma) \right) \right\}$$

are γ-level confidence intervals for β, where

$$\widetilde{U}_n(\beta) = \log(\rho)\big(\widetilde{A}_n(\tau_n)^{-1} - \widetilde{A}_n(\rho\tau_n)^{-1}\big)^{-1/2}(\widetilde{\beta}_n - \beta).$$

We may thus compare the $\widehat{I}_n(\gamma)$ to the $\widetilde{I}_n(\gamma)$.

In each run of our simulation, we will generate $n = 23{,}400$ observations, corresponding to observations taken every second of a typical 6.5-hour trading day. Our observations will be drawn from a log-price process

$$X_t = B_t + \gamma_t R_t, \qquad t \in [0, 1],$$

where B_t is a standard Brownian motion; the deterministic scaling process

$$\gamma_t = (2t - 1) \vee 0;$$

and the jump process

$$R_t = \theta_1 S_t^{\beta} + \theta_2 S_t^{\beta - 0.2},$$

for constants $\theta_1, \theta_2 > 0$, and independent α-stable processes S_t^{α}.

The process X_t thus models a price process with both diffusion and jump components. Its jumps are driven by a β-stable process, with time-varying intensity γ_t, but also contain a nuisance component, with jump activity $\beta - 0.2$.

The relative strengths of these jumps are given by the constants θ_1 and θ_2, which we will set in terms of a parameter $p \in (0, 1)$. To set θ_1, we will require that an increment $\theta_1(S_{j/n}^{\beta} - S_{(j-1)/n}^{\beta})$ contains a jump larger than 0.2 with probability p. To set θ_2, we will likewise require this condition holds for $\theta_2 S_t^{\beta - 0.2}$, with probability $0.05p$.

To model the microstructure noise present at one-second time scales, we will generate observations

$$Z_j = X_{j/n} + 0.01\varepsilon_j, \qquad j = 0, \ldots, n - 1,$$

where the independent noises $\varepsilon_j \sim N(0, 1)$. As noted in Aït-Sahalia and Jacod (2009) and in Section 3, the estimates $\widetilde{\beta}_n$ and $\widehat{\beta}_n$ can be expected to be robust to the presence of such noise, and we will compute them as though the observations Z_j were noiseless.

The estimates $\widetilde{\beta}_n$ and $\widehat{\beta}_n$ then depend on a number of parameters; we begin by considering the inverse thresholds $\tau_n = Cn^{\alpha}$. As noted in Aït-Sahalia and Jacod (2009), τ_n should be chosen to ensure our jump counts $\widetilde{A}_n$ or $\widehat{A}_n$ will be zero when no jump is present; the constant C should thus be chosen relative to the size of the diffusion component of X_t, as measured for example by its integrated volatility.

In our simulations, we know that the integrated volatility of X_t is equal to one, and so we may choose our parameters accordingly. In general, the volatility will not be equal to one; however, we can achieve a similar effect by first renormalising

TABLE 1
Simulated means and standard deviations of the estimates $\widetilde{\beta}_n$ and $\widehat{\beta}_n$, as well as coverages of the 95% confidence intervals $\widetilde{I}_n(0.95)$ and $\widehat{I}_n(0.95)$

β	p	$\widetilde{\beta}$			$\widehat{\beta}$		
		0.5%	1%	2%	0.5%	1%	2%
0.4	mean	0.39	0.39	0.39	0.39	0.39	0.39
	std. dev.	0.18	0.13	0.09	0.13	0.09	0.07
	95% cov.	0.90	0.91	0.93	0.91	0.92	0.91
0.8	mean	0.85	0.82	0.80	0.81	0.80	0.79
	std. dev.	0.49	0.34	0.23	0.26	0.18	0.13
	95% cov.	0.88	0.92	0.93	0.92	0.93	0.93
1.2	mean	1.13	1.21	1.23	1.22	1.22	1.20
	std. dev.	0.75	0.59	0.46	0.40	0.29	0.21
	95% cov.	0.80	0.90	0.93	0.93	0.94	0.92
1.6	mean	0.91	1.26	1.44	1.54	1.58	1.57
	std. dev.	0.92	0.80	0.63	0.43	0.36	0.31
	95% cov.	0.53	0.77	0.89	0.93	0.93	0.92

the observations to have estimated integrated volatility equal to one. Such an estimate could be provided by the method of Podolskij and Vetter (2009), for example, although we will not pursue this further here.

In any case, we may now choose our parameters without worrying about issues of scale. With $\widetilde{\beta}$, Aït-Sahalia and Jacod (2009) recommend a threshold rate $\alpha = 1/5$; with $\widehat{\beta}$, we will instead use the rate α given by Theorem 1. It remains to choose the constants C, ρ, and for $\widehat{\beta}$ also m; in our tests, we found the values $C = 0.05$, $\rho = 2$, and $m = 3$ worked well.

Table 1 then gives the mean and standard deviation of 10,000 simulated estimates $\widetilde{\beta}_n$ or $\widehat{\beta}_n$, for a number of choices of β, p and m. The table also gives the simulated coverage of the 95% confidence intervals $\widetilde{I}_n(0.95)$ or $\widehat{I}_n(0.95)$. We see that the multi-scale estimate $\widehat{\beta}_n$ has reduced bias and variance compared with the single-scale estimate $\widetilde{\beta}_n$, while the confidence intervals $\widehat{I}_n(0.95)$ retain good coverage, improving upon $\widetilde{I}_n(0.95)$ when β is large.

Figure 1 plots the RMSE of the estimates $\widetilde{\beta}_n$ and $\widehat{\beta}_n$; in the case $p = 1\%$, Figure 2 further gives the full simulated distribution of $\widehat{\beta}_n$. Again, we can see the multi-scale estimate $\widehat{\beta}_n$ is more accurate than the single-scale estimate $\widetilde{\beta}_n$. While the accuracy of $\widehat{\beta}_n$ suffers when β is large, it remains good enough to distinguish between different values of β.

Finally, Figure 3 plots the simulated distribution of the standardised errors $\widehat{U}_n$, together with the density of a standard Gaussian distribution, shown as a solid line. We can see that even in the finite-sample case, for $\beta = 0.4, 0.8, 1.2$, the errors $\widehat{U}_n$ show good agreement with their asymptotic distributions.
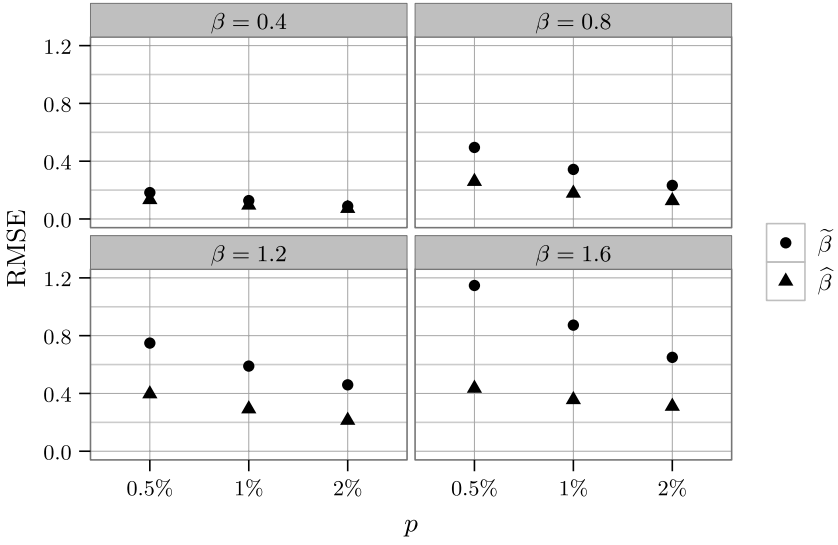
FIG. 1.   *Simulated RMSEs of the estimates $\widetilde{\beta}_n$ and $\widehat{\beta}_n$.*

In the case $\beta = 1.6$, we see a strong deviation from Gaussian on the right tail of $\widehat{U}_n$, due to the clipping of $\widehat{\beta}_n$ at 2. This clipping, however, serves only to reduce the error in the estimate $\widehat{\beta}_n$, and so does not harm the coverage of the confidence intervals $\widehat{I}_n(\gamma)$. Furthermore, the effect can be expected to disappear as n tends to infinity.
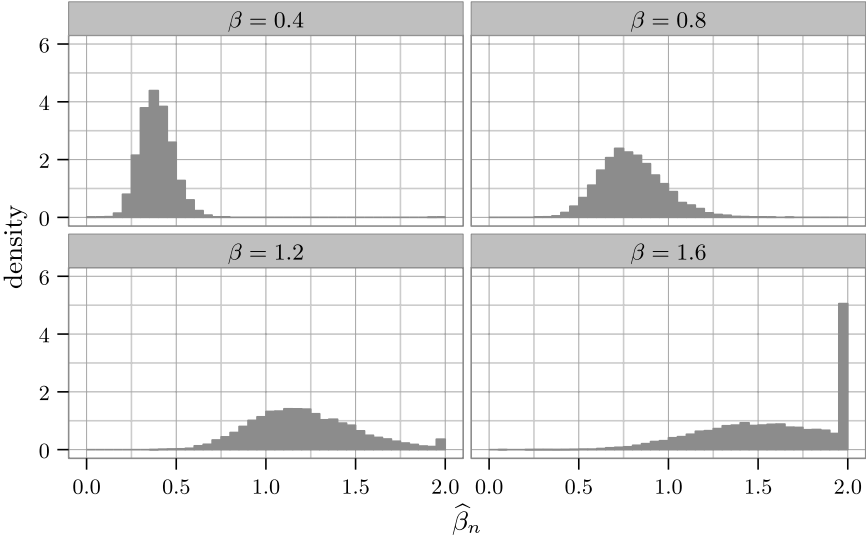

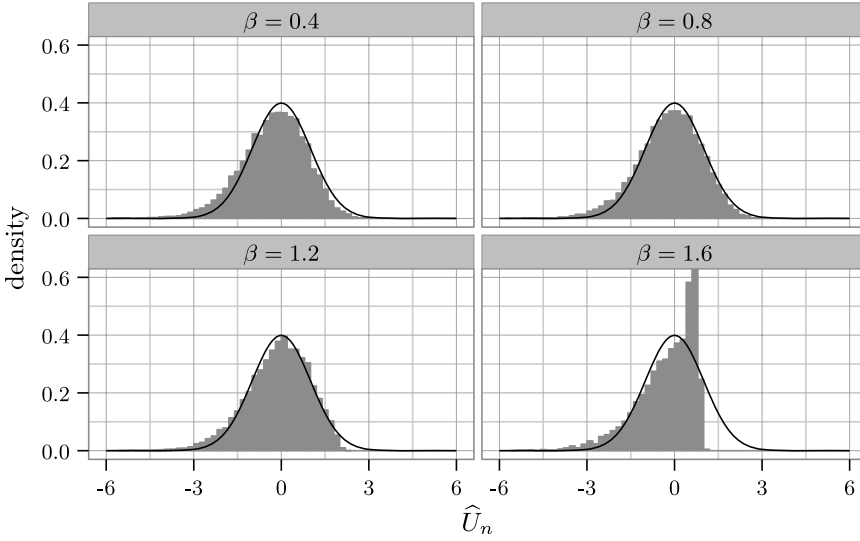FIG. 2.   *Simulated distributions of the estimates $\widehat{\beta}_n$, $p = 1\%$.*

FIG. 3. *Simulated distributions of the standardised errors $\widehat{U}_n$, $p = 1\%$.*

5. Proofs. We now give a proof of Theorem 1. In Section 5.1, we will state the technical results we require; in Section 5.2, prove our main results; and in the supplementary material [Bull (2015b)], give the remaining technical proofs.

5.1. *Technical results.* We begin with a technical lemma bounding various stochastic integrals, similarly to Jacod and Protter (2012).

LEMMA 1. *Let B_s be a Brownian motion, $\mu(dx, ds)$ a Poisson jump measure with intensity $dx \, ds$, a_s a predictable process, $f_s(x)$ a predictable function, $t \in [0, 1]$, and $\kappa_p > 0$ denote constants depending only on $p \geq 1$.*

(i) *If $\int_0^t |a_s| \, ds < \infty$,*

$$\left| \int_0^t a_s \, ds \right|^p \leq t^{p-1} \int_0^t |a_s|^p \, ds.$$

(ii) *If a_s is locally bounded,*

$$\mathbb{E}\left[\left| \int_0^t a_s \, dB_s \right|^p \right] \leq \kappa_p \mathbb{E}\left[\left(\int_0^t a_s^2 \, ds \right)^{p/2} \right].$$

(iii) *If $\int_{\mathbb{R}} f_s(x)^2 \, dx$ is locally bounded, and $p \in [1, 2]$, then*

$$\mathbb{E}\left[\left| \int_0^t \int_{\mathbb{R}} f_s(x)(\mu(dx, ds) - dx \, ds) \right|^p \right] \leq \kappa_p \mathbb{E}\left[\int_0^t \int_{\mathbb{R}} |f_s(x)|^p \, dx \, ds \right].$$

(iv) *If $\int_{\mathbb{R}} 1 \wedge |f_s(x)|\, dx$ is locally bounded, then*

$$\mathbb{E}\left[1 \wedge \left|\int_0^t \int_{\mathbb{R}} f_s(x)\mu(dx, ds)\right|^p\right] \leq \kappa_p \mathbb{E}\left[\int_0^t \int_{\mathbb{R}} 1 \wedge |f_s(x)|\, dx\, ds\right].$$

(v) *If a_s and $\int_{\mathbb{R}} 1 \wedge f_s(x)^2\, dx$ are locally bounded, $p \in [1, 2]$, and $\alpha \geq 0$, then*

$$\mathbb{E}\left[1 \wedge \left|t^{-\alpha} \int_0^t a_s \int_{\mathbb{R}} f_s(x)\big(\mu(dx, ds) - 1_{|f_s(x)|<1}\, dx\, ds\big)\right|^p\right]$$

$$\leq \kappa_p \mathbb{E}\left[\int_0^t |t^{-\alpha} a_s|^p \int_{|f_s(x)|<t^\alpha} |f_s(x)|^p\, dx\, ds + \int_0^t \int_{|f_s(x)|\geq t^\alpha} dx\, ds\right.$$

$$\left. + t^{p-1} \int_0^t \left|t^{-\alpha} a_s \int_{t^\alpha \leq |f_s(x)|<1} f_s(x)\, dx\right|^p\, ds\right].$$

PROOF.   Parts (i) and (ii) are immediate from the Hölder and Burkholder–Davis–Gundy inequalities, respectively. Part (iii) follows from Lemma 2.1.5 of Jacod and Protter (2012), and part (iv) likewise follows from their Lemma 2.1.8, noting that the left-hand side is decreasing in p.

Finally, let W denote the left-hand side of part (v). We make the decomposition

$$W \leq \kappa_p \mathbb{E}\left[\left|\int_0^t \int_{\mathbb{R}} g_{1,s}(x)\big(\mu(dx, ds) - dx\, ds\big)\right|^p\right.$$

$$\left. + 1 \wedge \left|\int_0^t \int_{\mathbb{R}} g_{2,s}(x)\mu(dx, ds)\right|^p + \left|\int_0^t \int_{\mathbb{R}} g_{3,s}(x)\, dx\, ds\right|^p\right],$$

where the terms

$$g_{i,s}(x) = t^{-\alpha} a_s f_s(x) 1_{J_i}\big(|f_s(x)|\big),$$

for intervals

$$J_1 = [0, t^\alpha), \qquad J_2 = [t^\alpha, \infty), \qquad J_3 = [t^\alpha, 1).$$

We deduce that

$$W \leq \kappa_p \mathbb{E}\left[\int_0^t \int_{\mathbb{R}} |g_{1,s}(x)|^p\, dx\, ds + \int_0^t \int_{\mathbb{R}} 1 \wedge |g_{2,s}(x)|\, dx\, ds\right.$$

$$\left. + t^{p-1} \int_0^t \left|\int_{\mathbb{R}} g_{3,s}(x)\, dx\right|^p\, ds\right],$$

using parts (i), (iii) and (iv). The desired result follows.   □

Next, we give a technical result on the characteristic exponents of one-sided stable processes.

LEMMA 2. *Let* $\mathcal{F}[f](u) = \int_{\mathbb{R}^d} \exp(i \langle u, x \rangle) f(x)\, dx$ *denote the Fourier transform,* $\Gamma(x)$ *the gamma function,* $\beta \in (0, 2)$, $u \in \mathbb{R}$, *and*

$$C_\beta = \begin{cases} -2\Gamma(-\beta)\cos(\beta\pi/2), & \beta \neq 1, \\ \pi, & \beta = 1. \end{cases}$$

We then have:

(i) $\int_0^\infty (1 - \cos(ux)) x^{-(1+\beta)}\, dx = \frac{1}{2} C_\beta |u|^\beta$; *and*
(ii) $C_\beta \int_{\mathbb{R}} \mathcal{F}[K](u) |u|^\beta\, du = 2\pi K_\beta$.

PROOF. We show each result in turn.

(i) This is a well-known result on stable processes; see, for example, Lemma 14.11 of Sato (1999).

(ii) For $\beta \neq 1$, using generalised functions, we have

$$C_\beta \int_{\mathbb{R}} \mathcal{F}[K](u) |u|^\beta\, du = C_\beta \int_{\mathbb{R}} K(x) \mathcal{F}[|u|^\beta](x)\, dx$$

$$= 2\pi \int_{\mathbb{R}} (1 - K(x)) |x|^{-(1+\beta)}\, dx,$$

since K is symmetric, and $K(0) = 1$. For $\beta = 1$, the same holds by analytic continuation. $\square$

Using these lemmas, we will be able to prove several Lévy approximations to the behaviour of random variables $\int_t^{t+h} a_s\, dX_s$. These approximations will hold under a localisation assumption; by standard techniques, we will be able to assume the following.

ASSUMPTION 2. Assumption 1 holds, the processes b_t, b_t^c, c_t, H_t, H_t', $\gamma_t^\pm$, $\int_{\mathbb{R}} 1 \wedge |\delta_t(x)|^{\upsilon_1}\, dx$ and $\int_{\mathbb{R}} 1 \wedge \delta_t^c(x)^2\, dx$ are uniformly bounded, and the stopping time $T_1 = \infty$.

We now state our Lévy approximation results; proofs of these results will be given in the supplementary material [Bull (2015b)]. Our first result bounds the error in approximating variables $\int_t^{t+h} a_s\, dX_s$ by Lévy integrals.

LEMMA 3. *Under Assumption 2, let* $0 \leq t \leq t + h \leq 1$, *set*

$$\xi_{t+h} = \int_t^{t+h} a_s\, dX_s,$$

for a deterministic real-valued process a_s *satisfying* $|a_s| \leq 1$, *and define the Lévy approximation*

$$\overline{\xi}_{t+h} = \int_t^{t+h} a_s \big(b_t\, ds + c_t\, dB_s + \gamma_t^+\, dL_s^+ + \gamma_t^-\, dL_s^-\big).$$

Then the approximation error

$$\xi_{t+h} - \overline{\xi}_{t+h} = Y_1 + Y_2,$$

where the random variable

$$Y_1 = \int_t^{t+h} a_s \left(H_t B_s + \int_t^s H_r' \, dB_r' + \int_t^s \int_{|\delta_r^c(x)|<1} \delta_r^c(x) \big(\mu(dx, dr) - dx\, dr\big) \right) dB_s,$$

and for $\alpha \in (0, \frac{1}{2})$, $u = O(h^{-\alpha})$, and some $\varepsilon > 0$, we have

$$\mathbb{E}\big[|u Y_1|^2 | \mathcal{F}_t\big] = O\big(h^{1+\varepsilon}\big), \qquad \mathbb{E}\big[1 \wedge |u Y_2| | \mathcal{F}_t\big] = O\big(h^{1+\varepsilon-\alpha\beta/2}\big),$$

uniformly over a_s and t.

Next, we state a result on the characteristic functions of random variables $\int_t^{t+h} a_s \, dX_s$. Our argument will follow Lemmas 11 and 12 of Jacod and Todorov (2014), although we give a tighter bound than in those results.

LEMMA 4. *In the setting of Lemma 3, suppose also that $|a_s| = 1$, and $\int_t^{t+h} a_s \, ds = 0$. Then for some $\varepsilon > 0$, we have*

$$\mathbb{E}\big[\cos(u\xi_{t+h}) | \mathcal{F}_t\big] = \exp\left(-\int_t^{t+h} \theta_t(a_s u) \, ds\right) + O\big(h^{1+\varepsilon-\alpha\beta/2}\big),$$

uniformly over a_s and t, where

$$\theta_t(u) = \tfrac{1}{2}(c_t u)^2 + C_\beta \overline{\gamma}_t |u|^\beta.$$

Our final technical result gives a large-jump approximation to functions of integrals $\int_t^{t+h} a_s \, dX_s$.

LEMMA 5. *In the setting of Lemma 3, suppose $|a_s| = 1$, let $t' \in [t, t + h]$, and set $h' = t + h - t'$. Also let f be a bounded even function, constant in a neighbourhood of the origin, whose derivative f' is a Schwartz function. Then*

$$\mathbb{E}\big[f(u\xi_{t+h}) | \mathcal{F}_{t'}\big] = f(u\xi_{t'}) + h' |u|^\beta \overline{\gamma}_{t'} \int_{\mathbb{R}} \big(f(x) - f(0)\big) |x|^{-(1+\beta)} \, dx + Y,$$

for a term Y satisfying $\mathbb{E}[|Y| | \mathcal{F}_t] = o(h^{1-\alpha\beta})$, uniformly in a_s, t and t'.

5.2. *Main proofs.* We now prove our main results. In the following, we will use the shorthand

$$t_j = j/n, \qquad t_{j,k} = (j + 2k)/n.$$

Our next lemma then bounds the means of our jump counts $\widehat{a}_{j,n}(\tau)$.

LEMMA 6. *Under Assumption* 2, *for* m *and* τ_n *as in the statement of Theorem* 1, *we have*

$$\mathbb{E}[\widehat{a}_{j,n}(\tau_n)|\mathcal{F}_{t_j}] = \tau_n^\beta K_\beta n^{-1}\overline{\gamma}_{t_j} + o\big(n^{-(1-\alpha\beta/2)}\big),$$

uniformly in $j = 0, \ldots, n - 2m - 1$.

PROOF. We can equivalently define the constants w_k by

$$(1) \qquad w_k = \sum_{l=k\vee 1}^{m} (-1)^{k+1}(2l)^{-1}\binom{l}{k},$$

letting the above also define a new constant w_0. We then have

$$2\pi\widehat{a}_{j,n}(\tau_n) = 2\pi\sum_{k=0}^{m} w_k\big(1 - K(\tau_n\Delta X_{j,k,n})\big),$$

since the summand vanishes for $k = 0$,

$$= -2\pi\sum_{k=0}^{m} w_k K(\tau_n\Delta X_{j,k,n}),$$

since $\sum_{k=0}^m w_k = -\sum_{l=1}^m (2l)^{-1}(1-1)^l = 0$,

$$= -\int_{\mathbb{R}} \mathcal{F}[K](u)\sum_{k=0}^{m} w_k \cos(u\tau_n\Delta X_{j,k,n})\,du,$$

by Fourier inversion,

$$= -\int_{|u|\le n^\varepsilon} \mathcal{F}[K](u)\sum_{k=0}^{m} w_k \cos(u\tau_n\Delta X_{j,k,n})\,du + O\big(n^{-1}\big),$$

for any $\varepsilon > 0$, since K is Schwartz.

For small enough ε, setting $\theta_{j,n}(u) = n^{-1}\theta_{t_j}(\tau_n u)$, we deduce

$$2\pi\mathbb{E}[\widehat{a}_{j,n}(\tau_n)|\mathcal{F}_{t_j}]$$

$$= -\int_{|u|\le n^\varepsilon} \mathcal{F}[K](u)\sum_{k=0}^{m} w_k \exp\big(-2k\theta_{j,n}(u)\big)\,du + o\big(n^{-(1-\alpha\beta/2)}\big),$$

using Lemma 4,

$$= \int_{|u|\le n^\varepsilon} \mathcal{F}[K](u)\sum_{l=1}^{m}(2l)^{-1}\big(1 - \exp(-2\theta_{j,n}(u))\big)^l\,du$$

$$+ o\big(n^{-(1-\alpha\beta/2)}\big),$$

from (1),

$$= \int_{|u| \leq n^{\varepsilon}} \mathcal{F}[K](u)\big(\theta_{j,n}(u) + O\big(\theta_{j,n}(u)^{m+1}\big)\big)\, du + o\big(n^{-(1-\alpha\beta/2)}\big),$$

considering the Taylor series of $\log(1 - x)$,

$$= \int_{\mathbb{R}} \mathcal{F}[K](u)\theta_{j,n}(u)\, du + o\big(n^{-(1-\alpha\beta/2)}\big),$$

since K is Schwartz, and for $|u| \leq n^{\varepsilon}$, $\theta_{j,n}(u) = O(n^{-(1-2(\alpha+\varepsilon))})$,

$$= n^{-1}\tau_n^{\beta}\overline{\gamma}_{t_j} C_{\beta} \int_{\mathbb{R}} \mathcal{F}[K](u)|u|^{\beta}\, du + o\big(n^{-(1-\alpha\beta/2)}\big),$$

since K is constant in a region of the origin, and so $\mathcal{F}[K]$ is orthogonal to polynomials vanishing at the origin,

$$= 2\pi \tau_n^{\beta} K_{\beta} n^{-1} \overline{\gamma}_{t_j} + o\big(n^{-(1-\alpha\beta/2)}\big),$$

using Lemma 2(ii). $\square$

We next prove a lemma giving the variance of terms like $K(\tau_n \Delta X_{j,k,n})$. To begin, for $\beta \in (0, 2)$, $\rho > 0$, we define the constants

$$\overline{K}_{\beta,\rho} = \rho^{-\beta/2} \int_{\mathbb{R}} \big(1 - K(x)\big)\big(1 - K(\rho x)\big)|x|^{-(1+\beta)}\, dx.$$

We then have the following result.

LEMMA 7.   *Under Assumption* 2, *for m and τ_n as in the statement of Theorem* 1, *let $j, j' = 0, \ldots, n - 2m - 1$, and $k, k' = 1, \ldots, m$. Also let a_s, a'_s be deterministic processes with $|a_s| = |a'_s| = 1$, let $\max(t_j, t_{j'}) \leq t \leq t + h \leq \min(t_{j,k}, t_{j',k'})$, and set*

$$\xi_{t'} = \int_{t_j}^{t'} a_s\, dX_s, \qquad \xi'_{t'} = \int_{t_{j'}}^{t'} a'_s\, dX_s,$$

$$V = \mathbb{E}\big[K(\tau_n \xi_{t_{j,k}})|\mathcal{F}_{t+h}\big], \qquad V' = \mathbb{E}\big[K\big(\rho \tau_n \xi'_{t_{j',k'}}\big)|\mathcal{F}_{t+h}\big].$$

Then

$$\mathbb{Cov}\big[V, V'|\mathcal{F}_t\big] = h\tau_n^{\beta}\rho^{\beta/2}\overline{K}_{\beta,\rho}\overline{\gamma}_t + Y,$$

for a term Y satisfying $\mathbb{E}[|Y||\mathcal{F}_{\min(t_j, t_{j'})}] = o(n^{-(1-\alpha\beta)})$, uniformly.

PROOF.   In the following, let Y denote any term satisfying

$$\mathbb{E}\big[|Y||\mathcal{F}_{\min(t_j, t_{j'})}\big] = o\big(n^{-(1-\alpha\beta)}\big).$$

Repeatedly applying Lemma 5, we have

$$\mathbb{E}[VV'|\mathcal{F}_t]$$
$$= \mathbb{E}\big[\big(K(\tau_n\xi_{t+h}) - (t_{j,k} - t - h)\tau_n^\beta\overline{\gamma}_{t+h}K_\beta\big)$$
$$\times \big(K(\rho\tau_n\xi'_{t+h}) - (t_{j',k'} - t - h)(\rho\tau_n)^\beta\overline{\gamma}_{t+h}K_\beta\big)|\mathcal{F}_t\big] + Y$$
$$= \mathbb{E}\big[K(\tau_n\xi_{t+h})K(\rho\tau_n\xi'_{t+h})|\mathcal{F}_t\big]$$
$$- \big((t_{j,k} - t - h) + \rho^\beta(t_{j',k'} - t - h)\big)\tau_n^\beta\overline{\gamma}_t K_\beta + Y$$
$$= K(\tau_n\xi_t)K(\rho\tau_n\xi'_t)$$
$$- h\tau_n^\beta\overline{\gamma}_t \int_{\mathbb{R}}(1 - K(x)K(\rho x))|x|^{-(1+\beta)}\,dx$$
$$- \big((t_{j,k} - t - h) + \rho^\beta(t_{j',k'} - t - h)\big)\tau_n^\beta\overline{\gamma}_t K_\beta + Y.$$

Again applying Lemma 5, we deduce that

$$\mathbb{C}\mathrm{ov}[V, V'|\mathcal{F}_t]$$
$$= \mathbb{E}[VV'|\mathcal{F}_t] - \mathbb{E}[V|F_t]\mathbb{E}[V'|\mathcal{F}_t]$$
$$= h\tau_n^\beta\overline{\gamma}_t\bigg((1 + \rho^\beta)K_\beta - \int_{\mathbb{R}}(1 - K(x)K(\rho x))|x|^{-(1+\beta)}\,dx\bigg) + Y$$
$$= h\tau_n^\beta\rho^{\beta/2}\overline{\gamma}_t\overline{K}_{\beta,\rho} + Y. \qquad \square$$

Next, we prove a lemma bounding the covariation of terms $K(\tau_n\Delta X_{j,k,n})$ with other martingales.

LEMMA 8. *Under Assumption 2, for m and τ_n as in the statement of Theorem 1, let $t \in [0, 1]$, and $k = 1, \ldots, m$. Then*

$$\sum_{j=0}^{\lfloor nt \rfloor - 2m} \mathbb{E}\big[(1 - K(\tau_n\Delta X_{j,k,n}))(M_{t_{j,k}} - M_{t_j})|\mathcal{F}_{t_j}\big] = o_p(n^{\alpha\beta/2}),$$

where M is either:

  (i) *equal to B; or*
 (ii) *a bounded martingale orthogonal to B*.

PROOF. We prove each claim in turn.

 (i) For $p, q > 1$, $1/p + 1/q = 1$, we have

$$\big|\mathbb{E}\big[(1 - K(\tau_n\Delta X_{j,k,n}))(B_{t_{j,k}} - B_{t_j})|\mathcal{F}_{t_j}\big]\big|$$
$$\leq \mathbb{E}\big[|1 - K(\tau_n\Delta X_{j,k,n})|^p|\mathcal{F}_{t_j}\big]^{1/p}\mathbb{E}\big[|B_{t_{j,k}} - B_{t_j}|^q|\mathcal{F}_{t_j}\big]^{1/q},$$

using Hölder's inequality,

$$= O(n^{-1/2})\mathbb{E}[1 - K(\tau_n \Delta X_{j,k,n})|\mathcal{F}_{t_j}]^{1/p},$$

using Lemma 1(ii), and since K takes values in $[0, 1]$,

$$= O(n^{-1/2-(1-\alpha\beta)/p}),$$

using Lemma 5,

$$= o(n^{-(1-\alpha\beta/2)}),$$

for small enough p. Summing this result, we conclude that

$$\sum_{j=0}^{\lfloor nt \rfloor - 2m} \mathbb{E}[(1 - K(\tau_n \Delta X_{j,k,n}))(B_{t_{j,k}} - B_{t_j})|\mathcal{F}_{t_j}] = o(n^{\alpha\beta/2}).$$

(ii) Using Lemma 3, for fixed k and n, we can write

$$\Delta X_{j,k,n} = \overline{\xi}_{t_{j,k}}^{(j)} + Y_1^{(j)} + Y_2^{(j)},$$

for a Lévy approximation $\overline{\xi}_t^{(j)}$, and error terms $Y_1^{(j)}$, $Y_2^{(j)}$. We can then write

$$\sum_{j=0}^{\lfloor nt \rfloor - 2m} \mathbb{E}[(1 - K(\tau_n \Delta X_{j,k,n}))(M_{t_{j,k}} - M_{t_j})|\mathcal{F}_{t_j}]$$

$$= \sum_{j=0}^{\lfloor nt \rfloor - 2m} \mathbb{E}[(K(\tau_n \overline{\xi}_{t_{j,k}}^{(j)}) - K(\tau_n \Delta X_{j,k,n}))(M_{t_{j,k}} - M_{t_j})|\mathcal{F}_{t_j}]$$

$$+ \sum_{j=0}^{\lfloor nt \rfloor - 2m} \mathbb{E}[(1 - K(\tau_n \overline{\xi}_{t_{j,k}}^{(j)}))(M_{t_{j,k}} - M_{t_j})|\mathcal{F}_{t_j}],$$

where we will bound separately the two sums on the right-hand side.

For the first sum, we have

$$\sum_{j=0}^{\lfloor nt \rfloor - 2m} \mathbb{E}[(K(\tau_n \overline{\xi}_{t_{j,k}}^{(j)}) - K(\tau_n \Delta X_{j,k,n}))(M_{t_{j,k}} - M_{t_j})|\mathcal{F}_{t_j}]$$

$$= O(1) \sum_{j=0}^{\lfloor nt \rfloor - 2m} \mathbb{E}[(1 \wedge |\tau_n(Y_1^{(j)} + Y_2^{(j)})|)|M_{t_{j,k}} - M_{t_j}||\mathcal{F}_{t_j}]$$

since $K(x + y) = K(x) + O(1 \wedge |y|)$,

$$= O(1) \sum_{j=0}^{\lfloor nt \rfloor - 2m} (\mathbb{E}[|\tau_n Y_1^{(j)}|^2|\mathcal{F}_{t_j}]^{1/2}\mathbb{E}[(M_{t_{j,k}} - M_{t_j})^2|\mathcal{F}_{t_j}]^{1/2}$$

$$+ \mathbb{E}[1 \wedge |\tau_n Y_2^{(j)}||\mathcal{F}_{t_j}]),$$

by Cauchy–Schwarz, and since M is bounded,

$$= o_p(n^{-1/2}) \sum_{j=0}^{\lfloor nt \rfloor - 2m} \mathbb{E}\big[(M_{t_{j,k}} - M_{t_j})^2\big]^{1/2} + o(n^{\alpha\beta/2}),$$

using Lemma 3,

$$= o_p(1)\Bigg(\sum_{j=0}^{\lfloor nt \rfloor - 2m} \mathbb{E}\big[(M_{t_{j,k}} - M_{t_j})^2\big] \Bigg)^{1/2} + o(n^{\alpha\beta/2}),$$

using Cauchy–Schwarz,

$$= o_p(1)\mathbb{E}\big[(M_1 - M_0)^2\big]^{1/2} + o(n^{\alpha\beta/2}),$$

as M is a martingale,

$$= o_p(n^{\alpha\beta/2}),$$

as M is bounded.

It remains to bound the second sum. Given $\mathcal{F}_{t_j}$, $\overline{\xi}_t^{(j)}$ is a function of the Brownian motion B and Poisson random measure μ, so we may apply Theorem III.4.34 of Jacod and Shiryaev (2003). We deduce that

$$K\big(\tau_n \overline{\xi}_{t_{j,k}}^{(j)}\big) - \mathbb{E}\big[K\big(\tau_n \overline{\xi}_{t_{j,k}}^{(j)}\big)|\mathcal{F}_{t_j}\big]$$
$$= \int_{t_j}^{t_{j,k}} G_s^{(j)} \, dB_s + \int_{t_j}^{t_{j,k}} \int_{\mathbb{R}} G_s'^{(j)}(x)\big(\mu(dx,ds) - dx \, ds\big),$$

for a predictable process $G_s^{(j)}$, and predictable function $G_s'^{(j)}(x)$. Likewise, by their Lemma III.4.24, we have

$$M_t - M_0 = \int_0^t G_s''(x)\big(\mu(dx,ds) - dx \, ds\big) + \overline{M}_t,$$

for a predictable function $G_s''(x)$, and a martingale $\overline{M}_t$ orthogonal to B and μ.

Now, as K is bounded, so is $G_s'^{(j)}$; furthermore, by considering the quadratic variation, we have

$$\mathbb{E}\Bigg[\int_{t_j}^{t_{j,k}} \int_{\mathbb{R}} G_s'^{(j)}(x)^2 \, dx \, ds \Big| \mathcal{F}_{t_j} \Bigg]$$

(2)
$$\leq \mathbb{V}\mathrm{ar}\big[K\big(\tau_n \overline{\xi}_{t_{j,k}}^{(j)}\big)|\mathcal{F}_{t_j}\big]$$
$$= O(1)\big(\mathbb{V}\mathrm{ar}\big[K(\tau_n \Delta_{j,k,n})|\mathcal{F}_{t_j}\big] + \mathbb{E}\big[1 \wedge |\tau_n(Y_1^{(j)} + Y_2^{(j)})|^2|\mathcal{F}_{t_j}\big]\big)$$
$$= O\big(n^{-(1-\alpha\beta)}\big),$$

using Lemmas 3 and 7. We likewise have

$$\mathbb{E}\left[\int_0^1 \int_{\mathbb{R}} G_s''(x)^2 \, dx \, ds\right] \leq \mathbb{E}[(M_t - M_0)^2]$$

(3)

$$= O(1),$$

as M is bounded.

Setting $\varepsilon_n = n^{-\alpha\beta/4}$, we thus obtain

$$\mathbb{E}\big[(1 - K(\tau_n \overline{\xi}_{t_{j,k}}^{(j)}))(M_{t_{j,k}} - M_{t_j})|\mathcal{F}_{t_j}\big]$$

$$= \mathbb{E}\left[\int_{t_j}^{t_{j,k}} \int_{\mathbb{R}} G_s'^{(j)}(x) G_s''(x) \, dx \, ds \Big| \mathcal{F}_{t_j}\right],$$

applying Itô's lemma,

$$\leq \mathbb{E}\left[\int_{t_j}^{t_{j,k}} \int_{\mathbb{R}} G_s'^{(j)}(x)^2 \, dx \, ds \Big| \mathcal{F}_{t_j}\right]^{1/2}$$

$$\times \mathbb{E}\left[\int_{t_j}^{t_{j,k}} \int_{|G_s''(x)| \leq \varepsilon_n} G_s''(x)^2 \, dx \, ds \Big| \mathcal{F}_{t_j}\right]^{1/2}$$

$$+ O(1)\mathbb{E}\left[\int_{t_j}^{t_{j,k}} \int_{|G_s''(x)| > \varepsilon_n} |G_s''(x)| \, dx \, ds \Big| \mathcal{F}_{t_j}\right],$$

using Cauchy–Schwarz, and since $G_s'^{(j)}$ is bounded,

$$= O(n^{-(1-\alpha\beta)/2})\mathbb{E}\left[\int_{t_j}^{t_{j,k}} \int_{|G_s''(x)| \leq \varepsilon_n} G_s''(x)^2 \, dx \, ds \Big| \mathcal{F}_{t_j}\right]^{1/2}$$

$$+ O(n^{\alpha\beta/4})\mathbb{E}\left[\int_{t_j}^{t_{j,k}} \int_{|G_s''(x)| > \varepsilon_n} G_s''(x)^2 \, dx \, ds \Big| \mathcal{F}_{t_j}\right],$$

using (2). We thus have

$$\sum_{j=0}^{\lfloor nt \rfloor - 2m} \mathbb{E}\big[(1 - K(\tau_n \overline{\xi}_{t_{j,k}}^{(j)}))(M_{t_{j,k}} - M_{t_j})|\mathcal{F}_{t_j}\big]$$

$$= O_p(n^{\alpha\beta/2})\mathbb{E}\left[\int_0^1 \int_{|G_s''(x)| \leq \varepsilon_n} G_s''(x)^2 \, dx \, ds\right]^{1/2}$$

$$+ O_p(n^{\alpha\beta/4})\mathbb{E}\left[\int_0^1 \int_{|G_s''(x)| > \varepsilon_n} G_s''(x)^2 \, dx \, ds\right],$$

using Cauchy–Schwarz again,

$$= o_p(n^{\alpha\beta/2}),$$

using (3). □

We now prove a limit theorem for our jump counts $\widehat{A}'_n(\tau)$.

LEMMA 9. *In the setting of Theorem* 1, *for* $l = 0, 1$, *set*

$$\eta_{n,l} = \tau_{n,l}^{\beta/2} \big(\tau_{n,l}^{-\beta} \widehat{A}'_n(\tau_{n,l}) - K_\beta \overline{\Gamma}_1 \big), \qquad \tau_{n,l} = \rho^l \tau_n.$$

Then the random vector

$$\eta_n \xrightarrow{sd} \overline{\Gamma}_1^{1/2} \widetilde{Z},$$

where the random variable $\widetilde{Z}$ *is defined, on a suitable extension of the probability space* $(\Omega, \mathcal{F}, \mathbb{P})$, *to satisfy*

$$\widetilde{Z} | \mathcal{F} \sim N \left(0, \begin{pmatrix} \overline{K}_{\beta,1} & \overline{K}_{\beta,\rho} \\ \overline{K}_{\beta,\rho} & \overline{K}_{\beta,1} \end{pmatrix} \right).$$

PROOF. We first make a localisation argument, allowing us to work under Assumption 2. Since we wish to bound both c_t and its characteristics, we will localise explicitly. For $k = 1, 2, \ldots$, let ϕ_k be a smooth bounded function with bounded derivatives, equal to the identity on $[-k, k]$.

Without loss of generality, we may assume that the stopping times T_k also localise the processes $b_t, b_t^c, H_t, H'_t, \gamma_t^\pm, \int_\mathbb{R} 1 \wedge |\delta_t(x)|^{\upsilon_1} dx$ and $\int_\mathbb{R} 1 \wedge \delta_t^c(x)^2 dx$. We can then write

$$X_t^{(k)} = \int_0^t b_{s \wedge T_k} ds + \int_0^t \phi_k(c_s^{(k)}) dB_s + \int_0^t \gamma_{s \wedge T_k}^+ dL_s^+ + \int_0^t \gamma_{s \wedge T_k}^- dL_s^-$$

$$+ \int_0^t \int_\mathbb{R} \delta_{s \wedge T_k}(x) \mu(dx, ds),$$

where

$$c_t^{(k)} = c_0 + \int_0^t b_{s \wedge T_k}^c ds + \int_0^t H_{s \wedge T_k} dB_s + \int_0^t H'_{s \wedge T_k} dB'_s$$

$$+ \int_0^t \int_\mathbb{R} \delta_{s \wedge T_k}^c(x) \big(\mu(dx, ds) - 1_{|\delta_{s \wedge T_k}^c(x)| < 1} dx \, ds \big).$$

We note that $X^{(k)} = X$ eventually almost-surely, so it suffices to prove our result instead for the processes $X^{(k)}$; an application of Itô's lemma shows that these processes satisfy Assumption 2.

We next define random variables

$$\zeta_{j,n,l} = \tau_{n,l}^{-\beta/2} \mathbb{E} \left[\sum_{j'=(j-2m+1)\vee 0}^{j \wedge (n-2m-1)} \widehat{a}_{j'}(\tau_{n,l}) \bigg| \mathcal{F}_{t_{j+1}} \right]$$

$$- \tau_{n,l}^{-\beta/2} \mathbb{E} \left[\sum_{j'=(j-2m+1)\vee 0}^{j \wedge (n-2m)-1} \widehat{a}_{j'}(\tau_{n,l}) \bigg| \mathcal{F}_{t_j} \right]$$

$$- \tau_{n,l}^{\beta/2} K_\beta n^{-1} \overline{\gamma}_{t_j} 1_{j < n-2m},$$

so we may write

$$\eta_{n,l} = \sum_{j=0}^{n-1} \zeta_{j,n,l} + \psi_{n,l},$$

where the term

$$\psi_{n,l} = -\tau_{n,l}^{\beta/2} K_\beta \int_0^1 (\overline{\gamma}_s - \overline{\gamma}_{\lfloor ns \rfloor/n} 1_{s < 1 - 2m/n}) \, ds.$$

Since

$$\mathbb{E}[\psi_{n,l}] = O\big(n^{\alpha\beta/2}\big)\Big(\int_0^{1-2m/n} \mathbb{E}\big[|\overline{\gamma}_s - \overline{\gamma}_{\lfloor ns \rfloor/n}|\big] \, ds + \int_{1-2m/n}^1 \mathbb{E}[|\overline{\gamma}_s|] \, ds\Big)$$

$$= O\big(n^{\alpha\beta/2}\big)\Big(\sum_{*\in\{+,-\}} \int_0^1 \mathbb{E}\big[|\gamma_s^* - \gamma_{\lfloor ns \rfloor/n}^*|^2\big]^{(1\wedge\beta)/2} \, ds + n^{-1}\Big),$$

as the function $x \mapsto |x|^\beta$ is $(1 \wedge \beta)$-Lipschitz, and $\overline{\gamma}_s$ is bounded,

$$= O\big(n^{-(1\wedge\beta-\alpha\beta)/2}\big),$$

as the $\gamma_s^\pm$ are smooth in quadratic mean,

$$= o(1),$$

we deduce that

$$\eta_{n,l} = \sum_{j=0}^{n-1} \zeta_{j,n,l} + o_p(1).$$

The desired result then follows from Theorem 2.2.15 of Jacod and Protter (2012), provided that for $t \in [0, 1]$, $l = 0, 1$:

(i) $\sum_{j=0}^{n-1} |\mathbb{E}[\zeta_{j,n,l}|\mathcal{F}_{t_j}]| \xrightarrow{p} 0$;

(ii) $\sum_{j=0}^{\lfloor nt \rfloor - 1} \mathbb{V}\mathrm{ar}[\zeta_{j,n,l}|\mathcal{F}_{t_j}] \xrightarrow{p} \overline{K}_{\beta,1} \overline{\Gamma}_t$;

(iii) $\sum_{j=0}^{\lfloor nt \rfloor - 1} \mathbb{C}\mathrm{ov}[\zeta_{j,n,0}, \zeta_{j,n,1}|\mathcal{F}_{t_j}] \xrightarrow{p} \overline{K}_{\beta,\rho} \overline{\Gamma}_t$;

(iv) $\sum_{j=0}^{n-1} \mathbb{E}[|\zeta_{j,n,l}|^p|\mathcal{F}_{t_j}] \xrightarrow{p} 0$, for some $p > 2$; and

(v) $\sum_{j=0}^{\lfloor nt \rfloor - 1} \mathbb{E}[\zeta_{j,n,l}(M_{t_{j+1}} - M_{t_j})|\mathcal{F}_{t_j}] \xrightarrow{p} 0$, where M is either:
  (a) equal to B; or
  (b) a bounded martingale orthogonal to B.

We now prove each claim in turn.

(i) From Lemma 6, we have that for $j = 0, \ldots, n - 2m - 1$,

$$\mathbb{E}[\zeta_{j,n,l}|\mathcal{F}_{t_j}] = \tau_{n,l}^{-\beta/2} \mathbb{E}[\widehat{a}_{j,n}(\tau_{n,l})|\mathcal{F}_{t_j}] - \tau_{n,l}^{\beta/2} K_\beta n^{-1} \overline{\gamma}_{t_j}$$

$$= o(n^{-1}).$$

From the definitions, we also have that for $j = n - 2m, \ldots, n - 1$,

$$\mathbb{E}[\zeta_{j,n,l}|\mathcal{F}_{t_j}] = 0.$$

We conclude that

$$\sum_{j=0}^{n-1} |\mathbb{E}[\zeta_{j,n,l}|\mathcal{F}_{t_j}]| = o(1).$$

(ii) From Lemma 7, we have that for $j = 2m - 1, \ldots, n - 2m - 1$, and terms $Y_{j,n,l}$ satisfying $\mathbb{E}[|Y_{j,n,l}|] = o(n^{-1})$,

$$\mathbb{V}\mathrm{ar}[\zeta_{j,n,l}|\mathcal{F}_{t_j}] = \tau_{n,l}^{-\beta} \sum_{k',k''=1}^{m} w_{k'} w_{k''}$$

$$\times \sum_{j'=j-2k'+1}^{j} \sum_{j''=j-2k''+1}^{j} \mathbb{C}\mathrm{ov}\big[\mathbb{E}\big[K(\tau_{n,l}\Delta X_{j',k',n})|\mathcal{F}_{t_{j+1}}\big]$$

$$\times \mathbb{E}\big[K(\tau_{n,l}\Delta X_{j'',k'',n})|\mathcal{F}_{t_{j+1}}\big]|\mathcal{F}_{t_j}\big]$$

$$= n^{-1}\left(2\sum_{k=1}^{m} kw_k\right)^2 \overline{K}_{\beta,1}\overline{\gamma}_{t_j} + Y_{j,n,l}$$

$$= n^{-1}\overline{K}_{\beta,1}\overline{\gamma}_{t_j} + Y_{j,n,l},$$

since

$$2\sum_{k=1}^{m} kw_k = 1 - \sum_{k=0}^{m}(-1)^k \binom{m}{k}$$

(4)
$$= 1 - (1-1)^m$$

$$= 1.$$

For $j = 0, \ldots, 2m - 2$ or $j = n - 2m, \ldots, n - 1$, by a similar argument, we have the same result for terms $Y_{j,n,l}$ satisfying $\mathbb{E}[|Y_{j,n,l}|] = O(n^{-1})$. We deduce that

$$\sum_{j=0}^{\lfloor nt \rfloor - 1} \mathbb{V}\mathrm{ar}[\zeta_{j,n,l}|\mathcal{F}_{t_j}] = n^{-1}\overline{K}_{\beta,1} \sum_{j=0}^{\lfloor nt \rfloor - 1} \overline{\gamma}_{t_j} + \sum_{j=0}^{\lfloor nt \rfloor - 1} Y_{j,n,l}$$

$$= \overline{K}_{\beta,1}\overline{\Gamma}_t + o_p(1) + O_p\left(\sum_{j=0}^{\lfloor nt \rfloor - 1} \mathbb{E}[|Y_{j,n,l}|]\right)$$

$$= \overline{K}_{\beta,1}\overline{\Gamma}_t + o_p(1).$$

(iii) The result follows similar to part (ii).

(iv) Since $\zeta_{j,n,l} = O(n^{-\alpha\beta/2})$, the result is trivial for large enough p.

(v) In either case (a) or (b), we have

$$\sum_{j=0}^{\lfloor nt \rfloor - 1} \mathbb{E}[\zeta_{j,n,l}(M_{t_{j+1}} - M_{t_j})|\mathcal{F}_{t_j}]$$

$$= \tau_{n,l}^{-\beta/2} \sum_{j=0}^{\lfloor nt \rfloor - 1} \sum_{j'=(j-2m+1)\vee 0}^{j \wedge (n-2m-1)} \mathbb{E}[\widehat{a}_{j'}(\tau_{n,l})(M_{t_{j+1}} - M_{t_j})|\mathcal{F}_{t_j}],$$

since M is a martingale,

$$= \tau_{n,l}^{-\beta/2} \sum_{j=0}^{\lfloor nt \rfloor - 2m} \mathbb{E}[\widehat{a}_j(\tau_{n,l})(M_{t_{j,m}} - M_{t_j})|\mathcal{F}_{t_j}] + o_p(1),$$

since $\widehat{a}_{j,n}(\tau)$ is bounded,

$$= \tau_{n,l}^{-\beta/2} \sum_{j=0}^{\lfloor nt \rfloor - 2m} \sum_{k=1}^{m} w_k$$

$$\times \mathbb{E}[(1 - K(\tau_{n,l}\Delta X_{j,k,n}))(M_{t_{j,k}} - M_{t_j})|\mathcal{F}_{t_j}] + o_p(1),$$

from the definition of $\widehat{a}_{j,n}(\tau)$,

$$= o_p(1),$$

using Lemma 8. □

Finally, we can prove a limit theorem for $\widehat{\beta}_n$.

PROOF OF THEOREM 1. We begin by defining the variables

$$\widehat{\beta}'_n = \log_\rho\left(\frac{A'_n(\rho\tau_n)}{A'_n(\tau_n)}\right), \qquad \widehat{U}'_n(\beta) = \frac{\tau_n^{\widehat{\beta}'_n/2}(\widehat{\beta}'_n - \beta)}{\widehat{\sigma}_{\rho,n}}.$$

From Lemma 9, on the event $\overline{\Gamma}_1 > 0$, we have that

$$A'_n(\tau_{n,l}) \xrightarrow{p} \tau_{n,l} K_\beta \overline{\Gamma}_1, \qquad \widehat{\beta}'_n \xrightarrow{p} \beta.$$

Hence, with probability tending to one,

$$\widehat{\beta}'_n = \widehat{\beta}_n, \qquad \widehat{U}'_n(\beta) = \widehat{U}_n(\beta).$$

It thus suffices to prove limit theorems for the quantities $\widehat{\beta}'_n$ and $\widehat{U}'_n(\beta)$.

Next, we note we may equivalently define $\sigma^2_{\beta,\rho}$ by

$$\sigma^2_{\beta,\rho} = ((1 + \rho^{-\beta})\overline{K}_{\beta,1} - 2\rho^{-\beta/2}\overline{K}_{\beta,\rho})/\log(\rho)^2 K_\beta^2 \overline{\Gamma}_1.$$

Again using Lemma 9, on the event $\overline{\Gamma}_1 > 0$, we also have

$$
\begin{aligned}
\widehat{\beta}'_n &= \log_\rho\big(\widehat{A}'_n(\rho\tau_n)/\widehat{A}'_n(\tau_n)\big) \\
&= \beta + \log_\rho\big(1 + \eta_{n,1}/(\rho\tau_n)^{\beta/2} K_\beta \overline{\Gamma}_1\big) - \log_\rho\big(1 + \eta_{n,0}/\tau_n^{\beta/2} K_\beta \overline{\Gamma}_1\big) \\
&= \beta + \big(\rho^{-\beta/2}\eta_{n,1} - \eta_{n,0}\big)/\log(\rho)\tau_n^{\beta/2} K_\beta \overline{\Gamma}_1 + o_p(n^{-\alpha\beta/2});
\end{aligned}
$$

we deduce that

$$
\tau_n^{\beta/2}(\widehat{\beta}'_n - \beta) \xrightarrow{\text{sd}} \sigma_{\beta,\rho} Z.
$$

Similarly, we have that $\widehat{\sigma}^2_{\rho,n} \xrightarrow{p} \sigma^2_{\beta,\rho}$. Using equation (2.2.5) of Jacod and Protter (2012), we thus obtain

$$
(\tau_n^{\beta/2}(\widehat{\beta}'_n - \beta), \widehat{\sigma}^2_{\rho,n}) \xrightarrow{\text{sd}} (\sigma_{\beta,\rho} Z, \sigma^2_{\beta,\rho}).
$$

By continuous mapping, we deduce that

$$
\widehat{U}'_n(\beta) \xrightarrow{\text{sd}} Z. \qquad \square
$$

We have thus proved Theorem 1; we note that Corollary 1 then follows directly.

SUPPLEMENTARY MATERIAL

Supplement to "Near-optimal estimation of jump activity in semimartingales" (DOI: 10.1214/15-AOS1349SUPP; .pdf). We provide proofs of our technical results.

REFERENCES

AÏT-SAHALIA, Y. and JACOD, J. (2009). Estimating the degree of activity of jumps in high frequency data. *Ann. Statist.* **37** 2202–2244. MR2543690

AÏT-SAHALIA, Y. and JACOD, J. (2011). Testing whether jumps have finite or infinite activity. *Ann. Statist.* **39** 1689–1719. MR2850217

AÏT-SAHALIA, Y. and JACOD, J. (2012). Identifying the successive Blumenthal–Getoor indices of a discretely observed process. *Ann. Statist.* **40** 1430–1464. MR3015031

BARNDORFF-NIELSEN, O. E. and SHEPHARD, N. (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics* **2** 1–37.

BARNDORFF-NIELSEN, O. E., GRAVERSEN, S. E., JACOD, J. and SHEPHARD, N. (2006). Limit theorems for bipower variation in financial econometrics. *Econometric Theory* **22** 677–719. MR2283032

BLUMENTHAL, R. M. and GETOOR, R. K. (1961). Sample functions of stochastic processes with stationary independent increments. *J. Math. Mech.* **10** 493–516. MR0123362

BULL, A. D. (2014). Estimating time-changes in noisy Lévy models. *Ann. Statist.* **42** 2026–2057. MR3262476

BULL, A. D. (2015a). Software for "Near-optimal estimation of jump activity in semimartingales." Available at https://www.repository.cam.ac.uk/handle/1810/248959.

BULL, A. D. (2015b). Supplement to "Near-optimal estimation of jump activity in semimartingales."
    DOI:10.1214/15-AOS1349SUPP.

CONT, R. and MANCINI, C. (2011). Nonparametric tests for pathwise properties of semimartingales.
    Bernoulli **17** 781–813. MR2787615

CONT, R. and TANKOV, P. (2004). *Financial Modelling with Jump Processes*. Chapman &
    Hall/CRC, Boca Raton, FL. MR2042661

JACOD, J. and PROTTER, P. (2012). *Discretization of Processes. Stochastic Modelling and Applied
    Probability* **67**. Springer, Heidelberg. MR2859096

JACOD, J. and REISS, M. (2014). A remark on the rates of convergence for integrated volatility
    estimation in the presence of jumps. *Ann. Statist*. **42** 1131–1144. MR3224283

JACOD, J. and SHIRYAEV, A. N. (2003). *Limit Theorems for Stochastic Processes*, 2nd ed.
    Grundlehren der Mathematischen Wissenschaften **288**. Springer, Berlin. MR1943877

JACOD, J. and TODOROV, V. (2014). Efficient estimation of integrated volatility in presence of infi-
    nite variation jumps. *Ann. Statist*. **42** 1029–1069. MR3210995

JING, B.-Y., KONG, X.-B. and LIU, Z. (2011). Estimating the jump activity index under noisy
    observations using high-frequency data. *J. Amer. Statist. Assoc*. **106** 558–568. MR2847970

JING, B.-Y., KONG, X.-B., LIU, Z. and MYKLAND, P. (2012). On the jump activity index for
    semimartingales. *J. Econometrics* **166** 213–223. MR2862961

LEE, S. S. and HANNIG, J. (2010). Detecting jumps from Lévy jump diffusion processes. *Journal
    of Financial Economics* **96** 271–290.

MANCINI, C. (2001). Disentangling the jumps of the diffusion in a geometric jumping Brownian
    motion. *Giornale dell'Istituto Italiano degli Attuari* **64** 19–47.

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

PAPAPANTOLEON, A. (2008). An introduction to Lévy processes with applications in finance.
    Preprint. Available at arXiv:0804.0482.

PODOLSKIJ, M. and VETTER, M. (2009). Estimation of volatility functionals in the simultaneous
    presence of microstructure noise and jumps. *Bernoulli* **15** 634–658. MR2555193

REISS, M. (2013). Testing the characteristics of a Lévy process. *Stochastic Process*. *Appl*. **123** 2808–
    2828. MR3054546

SATO, K.-I. (1999). *Lévy Processes and Infinitely Divisible Distributions. Cambridge Studies in
    Advanced Mathematics* **68**. Cambridge Univ. Press, Cambridge. MR1739520

TODOROV, V. and TAUCHEN, G. (2010). Activity signature functions for high-frequency data anal-
    ysis. *J. Econometrics* **154** 125–138. MR2558956

TODOROV, V. and TAUCHEN, G. (2012a). Inverse realized Laplace transforms for nonparamet-
    ric volatility density estimation in jump-diffusions. *J. Amer. Statist. Assoc*. **107** 622–635.
    MR2980072

TODOROV, V. and TAUCHEN, G. (2012b). The realized Laplace transform of volatility. *Economet-
    rica* **80** 1105–1127. MR2963883

WOERNER, J. H. C. (2011). Analyzing the fine structure of continuous time stochastic processes. In
    Seminar on Stochastic Analysis, *Random Fields and Applications VI. Progress in Probability* **63**
    473–492. Birkhäuser, Basel. MR2857040

ZHAO, Z. and WU, W. B. (2009). Nonparametric inference of discretely sampled stable Lévy pro-
    cesses. *J. Econometrics* **153** 83–92. MR2558496

STATISTICAL LABORATORY
UNIVERSITY OF CAMBRIDGE
WILBERFORCE ROAD
CAMBRIDGE CB3 0WB
UNITED KINGDOM
E-MAIL: a.bull@statslab.cam.ac.uk