

# ESTIMATING THE DEGREE OF ACTIVITY OF JUMPS IN HIGH FREQUENCY FINANCIAL DATA

BY YACINE AÏT-SAHALIA<sup>1</sup> AND JEAN JACOD

*Princeton University and UPMC (Université Paris-6)*

We define a generalized index of jump activity, propose estimators of that index for a discretely sampled process, and derive the estimators' properties. These estimators are applicable despite the presence of Brownian volatility in the process, which makes it more challenging to infer the characteristics of the small, infinite activity, jumps. When the method is applied to high frequency stock returns, we find evidence of infinitely active jumps in the data and estimate their index of activity.

**1. Introduction.** Using high frequency financial data, which are now widely available, we can hope to answer a number of questions regarding the characteristics of the process that drives asset returns. Let us model the log-price  $X$  of some asset as a 1-dimensional process, which we will observe over a fixed time interval  $[0, T]$  at discrete times  $0, \Delta_n, 2\Delta_n, \dots$  with a time interval  $\Delta_n$  between successive observations that is small: this is the essence of high frequency data. Let us further assume that this process is an Itô semimartingale, meaning that its characteristics are absolutely continuous with respect to Lebesgue measure: so it has a drift, a continuous martingale part which is the integral of a possibly stochastic process with respect to a Brownian motion, and we will also let it have jumps, with a possibly stochastic Lévy measure.

For modeling purposes one would like to infer the characteristics of  $X$ , that is, its drift, its volatility and its Lévy measure, from the observations. When the time interval  $\Delta_n$  goes to 0, it is well known that one can infer consistently the volatility, under very weak assumptions. But such consistent inference is impossible for the drift or the Lévy measure, if the overall time interval  $[0, T]$  is kept fixed.

In fact, even in the unrealistic case where the whole path of  $X$  is observed over  $[0, T]$ , one can infer neither the drift nor the Lévy measure. One can however hope to be able to characterize the behavior of the Lévy measure near 0: first whether it does not explode near 0, meaning that the number of jumps is finite; second, when this number is infinite, we would like to be able to say something about the concentration of small jumps.

---

<sup>1</sup>Supported in part by NSF Grant DMS-0532370.

*AMS 2000 subject classifications.* Primary 62F12, 62M05; secondary 60H10, 60J60.

*Key words and phrases.* Jumps, index of activity, Blumenthal-Gettoor index, discrete sampling, high frequency.

For this purpose, let us define

$$(1) \quad B(r)_t = \sum_{s \leq t} |\Delta X_s|^r, \quad I_t = \{r \geq 0 : B(r)_t < \infty\}, \quad \beta_t = \inf(I_t),$$

where  $\Delta X_s = X_s - X_{s-}$  is the size of the jump at time  $s$ , and  $r \geq 0$ , with the convention  $0^0 = 0$ . Necessarily, the (random) set  $I_t$  contains the interval  $(\beta_t, \infty)$ , whereas it may contain  $\beta_t$  itself or not. Moreover  $2 \in I_t$  always, and of course  $t \mapsto \beta_t$  is non-decreasing. Hence if we observe the whole path of  $X$  over  $[0, T]$ , we know the sets  $I_t(\omega)$  and the numbers  $\beta_t(\omega)$  for all  $t \leq T$ .

We call  $\beta_T(\omega)$  the *jump activity index* for the path  $t \mapsto X_t(\omega)$  at time  $T$  (or, more precisely, up to time  $T$ ). We define this index in analogy with the special case where  $X$  is a Lévy process: in this case,  $I_t$  and  $\beta_t$  are no longer random, and further they do not depend on the time  $t$ , and  $I_t$  is also the set of all  $r \geq 0$  such that  $\int_{\{|x| \leq 1\}} |x|^r F(dx) < \infty$ , where  $F$  is the Lévy measure. This property shows that, for a Lévy process, the jump activity index coincides with the Blumenthal-Gettoor index of the process (see Blumenthal and Gettoor (1961)). In the further special case where  $X$  is a stable process, then  $\beta$  is also the stable index of the process.

When  $X$  is a Lévy process, the interval  $I$  and the index  $\beta$  are of course only tiny elements of the whole Lévy measure  $F$ , conveying approximately the same information ( $I$  gives slightly more information than  $\beta$ ). However, the value of  $\beta$  is probably the most informative knowledge one can draw about  $F$  from the observation of the path  $t \mapsto X_t$  for all  $t \leq T$  when  $T$  is finite. Things are very different when  $T \rightarrow \infty$ , though, since observing  $X$  over  $[0, \infty)$  completely specifies  $F$ . But when the time horizon  $T$  is kept fixed, and with the whole path observed over  $[0, T]$ , we can infer only the behavior of the Lévy measure  $F$  near 0 (because we need potentially infinitely many observations for consistent estimation). Then,  $\beta$  captures an essential qualitative feature of  $F$ , which is its level of activity: when  $\beta$  increases, the (small) jumps tend to become more and more frequent.

$\beta$  is related to the “degree of activity” of jumps. All Lévy measures put finite mass on the set  $(-\infty, -\varepsilon] \cup [\varepsilon, +\infty)$  for any arbitrary  $\varepsilon > 0$ , so if the process has infinite jump activity it must be because of the “small” jumps, defined as those smaller than  $\varepsilon$ . If  $F([-\varepsilon, \varepsilon]) < \infty$ , the process has finite activity and  $0 \in I$ , or equivalently  $\beta = 0$ . But if  $F([-\varepsilon, \varepsilon]) = \infty$ , the process has infinite activity, and in addition  $\beta > 0$  as long as the Lévy measure  $F([-\varepsilon, \varepsilon])$  diverges near 0 at a rate larger than a power  $\varepsilon^{-a}$  for some  $a > 0$ . The higher  $\beta$  gets (up to 2), the more active the small jumps become. The same remarks apply also for general semimartingales. These properties are what motivate our calling  $\beta$  a jump activity index, and our interest in estimating it.

In the more realistic situation where the semimartingale  $X$  is only observed at times  $i\Delta_n$  over  $[0, T]$ , the estimation problem is made more challenging by the presence in  $X$  of a continuous martingale part. By its very nature,  $\beta_T$  characterizes the behavior of  $F$  near 0. Hence it is natural to expect that the small increments of the process are going to be the ones that are most informative about  $\beta_T$ . But those

small increments are precisely the ones where the contribution from the continuous martingale part of the process is inexorably mixed with the contribution from the small jumps. Being able to “see through” the continuous part of the semimartingale in order to say something about the number and concentration of small jumps is going to be the challenge we face as we attempt to estimate  $\beta_T$ .

Related to this paper are Woerner (2006), who proposes an estimator of the jump activity index and of the Hurst exponent, but in the absence of a continuous Brownian part to the semimartingale, and Cont and Mancini (2007), who propose a test for the finiteness of the variation of the jump part. Another related problem is the estimation of the index  $\beta$  of a stable process (see e.g., DuMouchel (1983)). But our situation here is fundamentally different from those in that we also have a continuous part in the semimartingale. The situation is also different from that in Aït-Sahalia and Jacod (2008), where we studied Fisher’s information for the parameters of a Brownian plus stable pure jump process, but there the jump process was the dominant component. Here, the continuous part of the semimartingale dominates the small increments, and we wish we estimate the activity index of a pure jump process where the dominant component is a Brownian motion.

The aim of this paper is to construct estimators  $\hat{\beta}_n(T)$  for  $\beta_T$  which are consistent when  $\Delta_n \rightarrow 0$ , and provide rates of convergence and asymptotic distributions. Ideally, we would also like to have estimators that are, as much as possible, model-free, in the sense that they behave well without too strong assumptions on the form of the drift, the volatility or the Lévy measure.

As it turns out, a fully model-free behavior of the estimators may be too much to ask: the assumptions we make below on the drift and the volatility process are quite weak, but obtaining rates of convergence are going to require more specific assumptions on the Lévy measure. In particular, we will assume that the main part of the Lévy measure near 0 behaves locally like the Lévy measure of a stable process and we will provide estimators and their properties when the index is  $\beta > 0$ . This assumption seems to be unavoidable since, as we shall also see, even when  $X$  is a Lévy process, strong assumptions on the Lévy measure are necessary.

The paper is organized as follows: in Section 2 we present the main results for semimartingales. Section 3 is devoted to the special and simpler case of a symmetric stable process, and Section 4 is about more general Lévy processes. We propose a small sample bias correction in Section 5. We present the results of Monte Carlo simulations in Section 6, and in Section 7 compute our estimators over all 2005 transactions of the Dow Jones stocks, focusing in particular on Intel and Microsoft. Section 8 is devoted to technical results and to the proof of the main theorems, which apply to Itô semimartingales, under suitable assumptions on the Lévy measure.

**2. The model and main results.** Our structural assumption is that  $X$  is a 1-dimensional Itô semimartingale on some filtered space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ , which means that its characteristics  $(B, C, \nu)$  are absolutely continuous with respect to Lebesgue measure (see Jacod and Shiryaev (2003) for all notions not explained

here). In other words, the characteristics of  $X$  have the form

$$(2) \quad B_t = \int_0^t b_s ds, \quad C_t = \int_0^t \sigma_s^2 ds, \quad \nu(dt, dx) = dt F_t(dx).$$

Here,  $b = (b_t)$  and  $\sigma = (\sigma_t)$  are real-valued optional processes, and  $F_t = F_t(\omega, dx)$  is a predictable random measure, meaning that for all Borel sets  $A$  in  $\mathbb{R}$  the process  $(F_t(A))$  is predictable (possibly taking the value  $+\infty$ ). This model is quite general: for instance, the drift, volatility and jump measures can be stochastic and jump themselves.

There are other ways of expressing this assumption, for example through a Wiener process  $W$  and a Poisson random measure  $\underline{\mu}$  with compensator  $\underline{\nu}(dt, dx) = dt \otimes dx$  (up to a possible enlargement of the space), as

$$(3) \quad \begin{aligned} X_t &= X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s \\ &+ \int_0^t \int_{\mathbb{R}} \delta(s, x) 1_{\{|\delta(s, x)| \leq 1\}} (\underline{\mu} - \underline{\nu})(ds, dx) \\ &+ \int_0^t \int_{\mathbb{R}} \delta(s, x) 1_{\{|\delta(s, x)| > 1\}} \underline{\mu}(ds, dx). \end{aligned}$$

In this formulation,  $b$  and  $\sigma$  are the same as in (2), and  $\delta = \delta(\omega, t, x)$  is a predictable function, and the connection with  $F_t$  is that  $F_t(\omega, dx)$  is the restriction to  $\mathbb{R} \setminus \{0\}$  of the image of the Lebesgue measure by the map  $x \mapsto \delta(\omega, t, x)$ . However, it is easier for the problem at hand to express the assumptions on  $F_t$  rather than on  $\delta$ , which moreover is not unique (whereas  $F_t$  is uniquely defined, up to null sets).

Below, for any measure  $H$  on  $\mathbb{R}$  we denote by  $\overline{H}$  its (symmetrical) tail function:

$$(4) \quad x > 0 \quad \mapsto \quad \overline{H}(x) = H([-x, x]^c).$$

Observe that  $B(r)_t < \infty$  if and only if  $B'(r)_t < \infty$ , where  $B'(r)_t = \sum_{s \leq t} |\Delta X_s|^r \wedge 1$ , and the process  $B'(r)$  is finite-valued if and only if it is locally integrable. In other words, divergence when it occurs is caused by the small jumps. Moreover for any stopping time  $T$  we have

$$\mathbb{E}(B'(r)_T) = \mathbb{E} \left( \int_0^T ds \int_{\mathbb{R}} (|x|^r \wedge 1) F_s(dx) \right).$$

We call *instantaneous jump activity index* at time  $t$  the (random) number

$$(5) \quad \beta_t^i = \inf \left\{ r > 0 : \int_{\mathbb{R}} (|x|^r \wedge 1) F_s(dx) < \infty \right\},$$

In light of (5), this is a natural generalization of the notion of Blumenthal-Gettoor index for Lévy processes (see Blumenthal and Gettoor (1961)).  $\beta^i$  is a predictable

process taking its values in  $[0, 2]$ . This process is also characterized by the following property: for any  $\varepsilon > 0$  we have

$$(6) \quad \lim_{x \rightarrow 0} x^{\beta_t^i + \varepsilon} \overline{F}_t(x) = 0, \quad \limsup_{x \rightarrow 0} x^{\beta_t^i - \varepsilon} \overline{F}_t(x) = \infty.$$

The “limsup” above is usually not a limit.

Note finally that  $\beta_t^i = 0$  does not necessarily imply that the process has finite jump activity since it is possible for the Lévy measure to diverge slowly near 0, at a sub-geometric speed: an example would be the Gamma process, which has  $F(dx) = (\eta \exp(-\kappa x) 1_{\{x > 0\}}/x) dx$  so that the Lévy measure  $\overline{F}_t(\varepsilon)$  diverges at a logarithmic rate in  $\varepsilon$ . Finite activity processes (compound Poisson) will have  $\beta_t^i = 0$  a.s., however.

*2.1. Assumptions.* We make two assumptions. The first one, on the drift  $b$  and volatility  $\sigma$ , is quite mild:

**Assumption 1.** *The processes  $b$  and  $\sigma$  are locally bounded.*

The second assumption, on the Lévy measures  $F_t$ , is more specific. Essentially, we split  $F_t$  as  $F_t = F_t' + F_t''$ , where

- $F_t'$  is very close to the Lévy measure of a  $\beta$ -stable process, restricted to a random interval  $(-z_t^{(-)}, z_t^{(+)})$  around 0, with some  $\beta$  that is not random; the random interval may be empty for some  $(\omega, t)$ , but not for all.
- $F_t''$  is another Lévy measure with jump activity index less than some  $\beta' < \beta$ .

The precise statement of the assumption is:

**Assumption 2.** *There are three (non-random) numbers  $\beta \in (0, 2)$  and  $\beta' \in [0, \beta)$  and  $\gamma > 0$ , and a locally bounded process  $L_t \geq 1$ , such that we have for all  $(\omega, t)$ :*

$$(7) \quad F_t = F_t' + F_t'',$$

where

a)  $F_t'$  has the form

$$(8) \quad F_t'(dx) = \frac{1 + |x|^\gamma f(t, x)}{|x|^{1+\beta}} \left( a_t^{(+)} 1_{\{0 < x \leq z_t^{(+)}\}} + a_t^{(-)} 1_{\{-z_t^{(-)} \leq x < 0\}} \right) dx,$$

for some predictable non-negative processes  $a_t^{(+)}, a_t^{(-)}, z_t^{(+)}$  and  $z_t^{(-)}$  and some non-negative predictable function  $f(\omega, t, x)$ , satisfying:

$$(9) \quad \frac{1}{L_t} \leq z_t^{(+)} \leq 1, \quad \frac{1}{L_t} \leq z_t^{(-)} \leq 1, \quad a_t^{(+)} + a_t^{(-)} \leq L_t, \quad f(t, x) \leq L_t,$$

b)  $F_t''$  is a measure which is singular with respect to  $F_t'$  and satisfies

$$(10) \quad x \in (0, 1] \quad \Rightarrow \quad \int_{\mathbb{R}} (|x|^{\beta'} \wedge 1) F_t''(dx) \leq L_t, .$$

We will also need the following increasing and locally bounded process:

$$(11) \quad \bar{A}_t = \int_0^t A_s ds \quad \text{where} \quad A_t = \frac{a_t^{(+)} + a_t^{(-)}}{\beta}.$$

**Remark 1.** *In view of (6), the instantaneous index at time  $t$  “due to” part  $F_t'$  of the Lévy measure is  $\beta$  on the set  $\{A_t > 0\}$ , and 0 otherwise, whereas the one due to  $F_t''$  is everywhere smaller than  $\beta'$ . Hence, outside a null set, we have  $\beta_t = \beta$  on the set  $\{\bar{A}_t > 0\}$  and  $\beta_t \leq \beta'$  otherwise.*

**Remark 2.** *One could formulate Assumption 2 slightly differently, writing  $F_t = F_t^1 + F_t^2$ , where  $F_t^1$  is given by (8) with  $f(t, x) = 0$  and  $F_t^2$  satisfying (10) with some  $\beta''$ , and further the restriction of  $F_t^2$  to  $[-z^{(-)}, z^{(+)}]$  has an absolutely continuous part with a density of the form  $f(t, x)|x|^{\gamma-1-\beta}$  with (9). The two formulations are equivalent, provided we take  $\beta' = \beta'' \vee (\beta - \gamma)$ .*

**Remark 3.** *Take any process of the form*

$$(12) \quad dX_t = b_t dt + \sigma_t dW_t + \delta_{t-} dY_t + \delta'_{t-} dY'_t$$

where  $\delta$  and  $\delta'$  are càdlàg adapted processes,  $Y$  is  $\beta$ -stable and  $Y'$  is any other Lévy process whose Lévy measure integrates  $|x|^{\beta'}$  near the origin and has an absolutely continuous part whose density is smaller than  $K|x|^{\gamma-1-\beta}$  on  $[-1, 1]$  for some  $\gamma > 0$  (for example a stable process with index strictly smaller than  $\beta'$ ). Then  $X$  will satisfy Assumption 2. For instance, when further  $Y$  is symmetrical with Lévy density  $D|x|^{1+\beta}$ , then it is satisfied with the two numbers  $\beta$  and  $\beta'$ , and  $\gamma$  as above, and  $f(t, x) = 0$ , and  $z_t^{(-)} = z_t^{(+)} = 1$ , and  $a_t^{(-)} = a_t^{(+)} = D\delta_{t-}^\beta$ .

2.2. *The estimators.* Recall that we observe  $X_{i\Delta_n}$  for  $i = 0, 1, \dots, [T/\Delta_n]$ . While the processes  $B(r)$  are defined from the jumps  $\Delta X_s = X_s - X_{s-}$  of  $X$ , we do not observe these jumps directly. Rather, all we observe are the discrete increments

$$(13) \quad \Delta_i^n X = X_{i\Delta_n} - X_{(i-1)\Delta_n}.$$

From these increments, we could try to evaluate  $B(r)_T$  and then infer  $\beta$ . Finding consistent estimators for  $B(r)_T$  is easy, but deducing from them an estimator for  $\beta$  is almost impossible, because we need to decide whether  $B(r)_T$  is infinite, or not, based on a finite sample.

So we propose the following idea. For fixed  $\varpi > 0$  and  $\alpha > 0$ , we write

$$(14) \quad U(\varpi, \alpha)_t^n = \sum_{i=1}^{[t/\Delta_n]} \mathbf{1}_{\{|\Delta_i^n X| > \alpha \Delta_n^\varpi\}}.$$

for the number of increments whose magnitude is greater than  $\alpha \Delta_n^\varpi$ . In all cases below, we will set  $\varpi < 1/2$ .

To better understand our rationale for doing this, consider the special case  $X = \sigma W + Y$ , where  $Y$  is a  $\beta$ -stable process, so  $\beta_t(\omega) = \beta$ . Any increment  $\Delta_i^n X = X_{i\Delta_n} - X_{(i-1)\Delta_n}$  satisfies  $\Delta_i^n X = \sigma \Delta_n^{1/2} W_1 + \Delta_n^{1/\beta} Y_1$  (equality in law). Then, recalling that  $\beta < 2$  and  $\Delta_n \rightarrow 0$ , with a large probability  $\Delta_i^n X$  is close to  $\sigma \Delta_n^{1/2} W_1$  in law. Those increments give essentially no information on  $Y$ , and are of order of magnitude  $\Delta_n^{1/2}$ . However if  $Y$  has a “big” jump at time  $s$ , the corresponding increment is close to  $\Delta Y_s$ . Hence, one has to throw away all the “small” increments. However,  $\beta$  is related to the behavior of  $F$  near 0, hence to the “very small” jumps of  $Y$ . This is why we will use only increments bigger than a cutoff level  $\alpha \Delta_n^\varpi$  for some  $\varpi \in (0, 1/2)$ . Asymptotically, those increments are big because, since  $\Delta_n^{1/2} \ll \Delta_n^\varpi$ , the main contribution is due to  $Y$ ; those increments mostly contain a single “big” jump, of size of order at least  $\Delta_n^\varpi$ ; and we still get some information on small jumps, because  $\Delta_n^\varpi \rightarrow 0$ .

So, by using the statistic  $U$  which simply counts the number of large increments – those greater than  $\alpha \Delta_n^\varpi$  – we are retaining only those increments of  $X$  that are not predominantly made of contributions from its continuous semimartingale part, which are  $O_p(\Delta_n^{1/2})$ , and instead are predominantly made of contributions due to a jump. The same heuristics work for more general Itô semimartingales.

As we will see later, the key property of the functionals  $U(\varpi, \alpha, \Delta_n)$  is their convergence in probability

$$(15) \quad \Delta_n^{\varpi\beta} U(\varpi, \alpha)_t^n \xrightarrow{\mathbb{P}} \frac{\bar{A}_t}{\alpha^\beta}$$

which we will show holds under Assumption 2. This property leads us to propose an estimator of  $\beta$  at each stage  $n$ . Fix  $0 < \alpha < \alpha'$  and define

$$(16) \quad \hat{\beta}_n(t, \varpi, \alpha, \alpha') = \frac{\log(U(\varpi, \alpha)_t^n / U(\varpi, \alpha')_t^n)}{\log(\alpha' / \alpha)}$$

which is at least consistent for estimating  $\beta$  on the set  $\{\bar{A}_t > 0\}$ . If either value of  $U$  in (16)-(17) is 0, then by convention we set the estimator to be 0.

$\hat{\beta}_n$  is constructed from a suitably scaled ratio of two  $U$ s evaluated on the same time scale  $\Delta_n$  but at two different fixed levels of truncation of the increments,  $\alpha$  and  $\alpha'$ . In a way, this construction is in the same spirit as the classical estimator of Hill (1975), who conducts inference about the tails of a distribution based on ratios of various extremes.

We can also propose a second estimator defined as

$$(17) \quad \hat{\beta}'_n(t, \varpi, \alpha) = \frac{\log(U(\varpi, \alpha)_t^n / U_2(\varpi, \alpha)_t^n)}{\varpi \log 2},$$

where  $U_2(\varpi, \alpha)_t^n$  is defined analogously to  $U(\varpi, \alpha)_t^n$  in (14) except that sampling at  $\Delta_n$  is replaced by sampling at  $2\Delta_n$ . That is,  $\hat{\beta}'_n$  is constructed from a suitably scaled ratio of two  $U$ s evaluated at the same level of truncation  $\alpha$ , but on two separate time scales,  $\Delta_n$  and  $2\Delta_n$ .

One could also look at a third estimator  $\hat{\beta}_n''$  obtained from two  $U$ s evaluated at two different rates of truncation  $\varpi$  and  $\varpi'$ . And one could further consider estimators based not just on counting the increments that exceed a certain cutoff, but also on the magnitude of these increments, as in the case of power variations truncated to use only the large increments.

In the rest of the paper, we will focus on the properties of the estimator  $\hat{\beta}_n$ , noting that a similar type of analysis yields the consistency and asymptotic distribution of the other two estimators. In general, the asymptotic variance of  $\hat{\beta}_n$  is smaller than that of  $\hat{\beta}_n'$  and  $\hat{\beta}_n''$ .

Before studying the properties of the estimator  $\hat{\beta}_n$ , let us make a few remarks:

**Remark 4.** *Asymptotically, as  $n \rightarrow \infty$ , the above estimators behave well. However, for any given  $n$  it may happen that they are not informative, because too few increments are retained, up to the extreme case where  $U(\varpi, \alpha)_t^n = U(\varpi, \alpha')_t^n = 0$ . If this is the case, one should take smaller values of  $\alpha$  and  $\alpha'$ .*

**Remark 5.** *Even when the estimators are well defined, they may take a value bigger than or equal to 2. In this case, the estimation is not reliable, and it may be an indication that Assumption 2 is simply not satisfied, which would be the case for example if there is no jump at all in the observed path. So it would make sense to convince oneself that jumps are present (see e.g., Ait-Sahalia and Jacod (2006)) before attempting to estimate  $\beta$ .*

**Remark 6.** *In practice, the levels of truncation  $\alpha$  and  $\alpha'$  may be set in relation to the volatility of the continuous part of the semimartingale, i.e.,  $(t^{-1} \int_0^t \sigma_s^2 ds)^{1/2}$ , since the objective is to eliminate the increments that are mainly due to the continuous part. In the presence of jumps, that volatility can be estimated for example using the small increments of the process.*

**Remark 7.** *The construction of the estimators relies on the property (15), which holds under (slightly) weaker assumptions than Assumption 2 provided that the definition of  $\bar{A}_t$  and the rate of convergence are suitably amended. As a result, the estimators given in (16) remain consistent under weaker assumptions. For example, when  $X$  has only finitely many jumps, the index is  $\beta = 0$ , and  $U(\varpi, \alpha)_t^n$  converges to the number of jumps between 0 and  $t$ , irrespective of the value of  $\alpha$ , so  $\hat{\beta}_n$  is equal to 0 for all  $n$  large enough (obviously, this rules out the possibility of a central limit theorem).*

**2.3. The main results.** Our first result states that our estimators estimate  $\beta$  on the (random) set  $\{\bar{A}_t > 0\}$  where the jump activity index is  $\beta$ , and we can state a rate of convergence:

**Theorem 1.** *Let  $0 < \alpha < \alpha'$  and  $0 < \varpi < 1/2$  and  $t > 0$ . Under Assumptions 1 and 2, we have  $\hat{\beta}_n'(t, \varpi, \alpha, \alpha') \xrightarrow{\mathbb{P}} \beta$  on the set  $\{\bar{A}_t > 0\}$ . Moreover, if*

$$(18) \quad \chi = \chi(\beta, \gamma, \beta', \varpi) = (\varpi\gamma) \wedge \frac{1 - \varpi\beta}{3} \wedge \frac{\varpi(\beta - \beta')}{1 + \beta'} \wedge \frac{1 - 2\varpi}{2} \wedge \frac{\varpi\beta}{2},$$



the estimators  $\hat{\beta}_n(t, \varpi, \alpha, \alpha')$  are  $\Delta_n^{\chi-\varepsilon}$ -rate consistent for any  $\varepsilon > 0$  on the set  $\{\bar{A}_t > 0\}$ , in the sense that in restriction to this set the sequence of variables  $\left(\frac{1}{\Delta_n^{\chi-\varepsilon}}(\hat{\beta}'_n(t, \varpi, \alpha, \alpha') - \beta)\right)_{n \geq 1}$  is bounded in probability (or, “tight”).

The number  $\chi$  is positive, but it may also be very small. If we want an associated distributional results, we need stronger assumptions, implying essentially that  $\chi = \varpi\beta/2$  above, and this requires that the activity indices  $\beta'$  and  $\beta - \gamma$  of the “non stable-like” part of the Lévy measure be sufficiently apart from the leading activity index  $\beta$ , as follows:

**Theorem 2.** *Let  $0 < \alpha < \alpha'$  and  $t > 0$ . Assume Assumptions 1 and 2 with  $\beta' \in [0, \beta/(2 + \beta))$  and  $\gamma > \beta/2$ . Then, if  $\varpi < 1/(2 + \beta) \wedge 2/(5\beta)$ , and in restriction to the set  $\{\bar{A}_t > 0\}$ , we have the following stable convergence in law to a centered normal variable independent of  $X$ :*

$$(19) \quad \frac{1}{\Delta_n^{\varpi\beta/2}} (\hat{\beta}_n(t, \varpi, \alpha, \alpha') - \beta) \xrightarrow{\mathcal{L}^{-(s)}} \mathcal{N}\left(0, \frac{\alpha'^\beta - \alpha^\beta}{\bar{A}_t(\log(\alpha'/\alpha))^2}\right),$$

The qualifier “in restriction to the set  $\{\bar{A}_t > 0\}$ ” is essential in this statement. Recall that, unlike the usual convergence in law, stable convergence in law makes it possible to restrict the convergence to a subset of  $\Omega$ , exactly as convergence in probability does. On the complement set  $\{\bar{A}_t = 0\}$ , anything can happen: on that set, the number  $\beta$  has no meaning as a jump activity index for  $X$  on  $[0, t]$ .

Moreover, the stable convergence in law allows for the convergence of standardized statistics:

**Theorem 3.** *Under the assumptions of Theorem 2, the variables*

$$(20) \quad \frac{\log(\alpha'/\alpha)}{\sqrt{\frac{1}{U(\varpi, \alpha', \Delta_n)_t} - \frac{1}{U(\varpi, \alpha, \Delta_n)_t}}} \left(\hat{\beta}_n(t, \varpi, \alpha, \alpha') - \beta\right)$$

*converge stably in law, in restriction to the set  $\{\bar{A}_t > 0\}$ , to a standard normal variable  $\mathcal{N}(0, 1)$  independent of  $X$ .*

These results are model-free in a sense, because the drift and the volatility processes are totally unspecified apart from Assumption 1, and the Lévy measures  $F_t$  are unspecified, other than the requirements specified in Assumption 2. These three theorems will be proved in Section 8 below.

**3. Stable processes.** Here we specialize the general results in an important special case, discussing in particular the efficiency of the estimators of  $\beta$  we propose. In the special case of stable processes, the model is fully specified parametrically, and we can compare the properties of efficient parametric estimators of  $\beta$  to those of the general estimators  $\hat{\beta}_n$ .

Denote by  $Y$  a symmetric stable process with index  $\beta \in (0, 1/2)$ . We study the two situations where  $X_t = Y_t$  (the simplest of all since there is no continuous part!),

and  $X_t = bt + \sigma W_t + Y_t$  where  $\sigma > 0$  and  $b \in \mathbb{R}$  and  $W$  is a Brownian motion. The Lévy measure depends on a scale parameter  $A > 0$  and the index  $\beta$ . It has the form

$$(21) \quad F(dx) = \frac{A\beta}{2|x|^{1+\beta}} dx, \quad \text{hence } \bar{F}(x) := F([-x, x]^c) = \frac{A}{x^\beta} \text{ for } x > 0.$$

The law of  $Y_1$  has an even density  $g$  and a tail function  $\bar{G}(x) = \mathbb{P}(|Y_1| > x)$  satisfying, as  $x \rightarrow \infty$  (see Zolotarev (1986), Theorems 2.4.2 and Corollary 2 of Theorem 2.5.1):

$$(22) \quad g(x) = \frac{A\beta}{2|x|^{1+\beta}} + O\left(\frac{1}{x^{1+2\beta}}\right), \quad \bar{G}(x) = \frac{A}{x^\beta} + O\left(\frac{1}{x^{2\beta}}\right).$$

In both cases  $X = Y$  and  $X_t = bt + \sigma W_t + Y_t$ , we obviously have Assumptions 1 and 2, with  $F_t = F'_t = F$  not depending on  $(\omega, t)$ , and with  $F''_t = F'''_t = 0$ , and  $\Xi = \Omega \times (0, \infty)$ , and  $\beta' = 0$ , and  $f_t(x) = 0$ , and finally  $A_t(\omega) = A$ , the constant in (21). Then we can apply the previous results, which further hold on the whole set  $\Omega$  (because here  $\bar{A}_t = tA > 0$  for all  $(t, \omega)$ ). The results are much easier to prove in this special case, and also the requirements on  $\varpi$  are significantly weaker, thus allowing for faster rates of convergence (the larger  $\varpi$ , the faster the convergence in Theorem 1). But these improved results are no longer model-free, since the structure of jumps is completely specified in this stable model up to the unknown parameters  $A$  and  $\beta$ .

3.1. *The case  $X = Y$ .* Consider first the case where  $X$  has no continuous part. Then the general results on  $\hat{\beta}_n$  can be improved to yield:

**Theorem 4.** *Assume that  $X = Y$ . Let  $0 < \alpha < \alpha'$  and  $\varpi > 0$  and  $t > 0$ .*

- a) *If  $\varpi < 1/\beta$ , the estimators  $\hat{\beta}_n(t, \varpi, \alpha, \alpha')$  converge in probability to  $\beta$ .*
- b) *If further  $\varpi < 2/(3\beta)$ , we have stable convergence in law, over the whole set  $\Omega$ , as described in Theorems 1 (with  $\bar{A}_t = tA$ ) and 3.*

Note that in part (a) of the theorem, the closer  $\beta$  is to 2 the stronger the constraint on the truncation rate  $\varpi$ .

These estimators are not rate-efficient, however. To see that, one can recall from Aït-Sahalia and Jacod (2008) that the parametric model in which one observes the values  $X_{i\Delta_n}$  for  $i\Delta_n \leq t$  is regular, and its Fisher information for estimating  $\beta$  is asymptotically of the form

$$(23) \quad I_n \sim \frac{\log(1/\Delta_n)}{\Delta_n} C_\beta t$$

for some constant  $C_\beta$ . We can thus hope for estimators which, after centering by  $\beta$  and normalization by  $\sqrt{\log(1/\Delta_n)}/\sqrt{\Delta_n}$  are  $\mathcal{N}(0, 1/C_\beta)$ , and in fact the MLE does this.

Where is the loss of efficiency coming from? In order to compute our general estimators  $\hat{\beta}_n$ , we are forced by the presence of a continuous part in  $X$  to discard a very sizeable portion of the data – this is the effect of truncating away the small increments of  $X$ . But here, if somehow we knew from the start that there is no

continuous part in  $X$ , there would no longer be a need to do that. It is clear that better estimators of  $\beta$  could then be constructed.

And if further the law of  $Y$  has a fully-specified parametric form, as is the case here, it would be possible to improve the estimators even more. In this example, we would simultaneously estimate  $\beta$  and  $A$ , but the rates would be unchanged and it would be even more model-dependent. So this is the kind of estimator which we do not want to use, since we have no hope of extending such an estimator to the general semimartingale situation (or in fact even to more general Lévy processes than the stable ones).

3.2. *The case  $X_t = bt + \sigma W_t + Y_t$ .* We now study the situation where  $Y$  is a stable process, but  $X$  now also contains a continuous part. The distributional properties of the estimators follow directly in this special case. Indeed, for this model,  $U(\varpi, \alpha)_t^n$  is essentially the same as, or close to, the number  $V(\varpi, \alpha)_t^n$  of jumps of  $Y$  which are bigger than  $\alpha \Delta_n^\varpi$  in the interval  $[0, t]$ . But  $V(\varpi, \alpha)_t^n$  is a Poisson random variable with parameter  $Ct/\alpha^\beta \Delta_n^{\beta\varpi}$  where  $C$  is a constant. Hence

$$(24) \quad \Delta_n^{\beta\varpi} V(\varpi, \alpha)_t^n \xrightarrow{\mathbb{P}} C/\alpha^\beta,$$

$$(25) \quad \frac{1}{\Delta_n^{\beta\varpi/2}} \left( \Delta_n^{\beta\varpi} V(\varpi, \alpha)_t^n - C/\alpha^\beta \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, C/\alpha^\beta).$$

These properties carry over to  $U(\varpi, \alpha)_t^n$  and this leads to the following improvement to Theorem 1:

**Theorem 5.** *Assume that  $X_t = bt + \sigma W_t + Y_t$ . Let  $0 < \alpha < \alpha'$  and  $\varpi > 0$  and  $t > 0$ .*

a) *If  $\varpi < 1/2$ , the estimators  $\hat{\beta}_n(t, \varpi, \alpha, \alpha')$  converge in probability to  $\beta$ .*

b) *If  $\varpi < 1/(2 + \beta)$ , we have the stable convergences in law, over the whole set  $\Omega$ , as described in Theorems 1 (with  $\bar{A}_t = tA$ ) and 3.*

The estimators  $\hat{\beta}_n$  are again not rate-efficient, although they do come close. In fact, using the methods of Aït-Sahalia and Jacod (2008), we can show that Fisher's information for estimating  $\beta$ , at stage  $n$ , satisfies

$$(26) \quad I_n \sim \frac{A (\log(1/\Delta_n))^{2-\beta/2}}{\sigma^\beta \Delta_n^{\beta/2}} C'_\beta t$$

for another constant  $C'_\beta$ . Furthermore, in the (partial) statistical model where we observe the increments provided they are bigger than  $\alpha \Delta_n^\varpi$  and discard all others (here  $\alpha > 0$  and  $0 < \varpi < 1/2$ ), Fisher's information now satisfies

$$(27) \quad I_n \sim \frac{A(1-\varpi)^2 (\log(1/\Delta_n))^2}{\alpha^\beta \Delta_n^{\varpi\beta}} C''_\beta t.$$

So our general estimators are almost (up to a  $\log(1/\Delta_n)$  factor) rate-efficient for the partial parametric statistical model. As for the “complete” model, the rate approaches the true rate, by taking  $\varpi$  close to  $1/2$ ; but we cannot take  $\varpi$  bigger than  $1/(2 + \beta)$ , and since in practice  $\beta$  is unknown other than being less than 2, a conservative choice may be  $\varpi = 1/4$ .

**4. General Lévy processes.** Let us now consider the case where  $X$  is a general Lévy process. Its characteristics are of the form (2) with  $b_t = b$ ,  $\sigma_t = \sigma$  and  $F_t = F$  deterministic and not depending on  $t$ . Then Assumption 1 holds. As to Assumption 2, it may or may not hold, but if it does it takes a slightly simpler form because then everything is independent of  $(\omega, t)$ ; in particular,  $\bar{A}_t = At$  for some constant  $A > 0$ . The two theorems 1 and 3 hold without modification, except that either  $\{\bar{A}_t > 0\} = \Omega$  for all  $t > 0$ , or  $\{\bar{A}_t > 0\} = \emptyset$  for all  $t$ , in which case those theorems are void of content.

What is important here, though, is that those results fail when the assumptions we made are not satisfied, even with such a simple probabilistic structure for  $X$ .

In order to see why Assumption 2 is needed, let us consider a simpler but closely related statistical model. More precisely, suppose that we observe all “big” jumps of  $X$  up to time  $t$ , that is  $\Delta X_s$  with  $|\Delta X_s| > \alpha \Delta_n^\varpi$  for all  $s \leq t$ . A priori, this should give us more information on the Lévy measure than the original observation scheme where only increments (as opposed to jumps) are observed and only those bigger than  $\alpha \Delta_n^\varpi$  are taken into consideration.

In this statistical setting, the estimators (16) have no meaning, but we can replace  $\hat{\beta}_n$  with

$$(28) \quad \bar{\beta}_n(t, \varpi, \alpha, \alpha') = \frac{\log(\bar{U}(\varpi, \alpha)_t^n / \bar{U}(\varpi, \alpha')_t^n)}{\log(\alpha' / \alpha)},$$

where we have set

$$(29) \quad \bar{U}(\varpi, \alpha)_t^n = \sum_{s \leq t} 1_{\{|\Delta X_s| > \alpha \Delta_n^\varpi\}}.$$

The estimators  $\bar{\beta}_n$  are of course only virtual since there is no hope of actually observing the exact jumps of the process. But in the rest of this section, we study the behavior of the estimators  $\bar{\beta}_n$  in order to gain some insight on the necessity of making a restrictive assumption on the Lévy measure  $F_t$  if one is to estimate  $\beta$ : we will see that such an assumption is needed even under these idealized circumstances. Set

$$(30) \quad \gamma_n(\varpi, \alpha) = \bar{F}(\alpha \Delta_n^\varpi).$$

**Lemma 1.** *Let*

$$(31) \quad M(\varpi, \alpha)_t^n = \frac{1}{\sqrt{\gamma_n(\varpi, \alpha)}} (\bar{U}(\varpi, \alpha)_t^n - \gamma_n(\varpi, \alpha) t).$$

a) *Each sequence of processes  $M(\varpi, \alpha)^n$  converges stably in law to a standard Wiener process, independent of  $X$ .*

b) *If  $\alpha < \alpha'$  all limit points of the sequence  $\gamma_n(\varpi, \alpha') / \gamma_n(\varpi, \alpha)$  are in  $[0, 1]$ . If further this sequence converges to  $\gamma$  then the pair  $(M(\varpi, \alpha)^n, M(\varpi, \alpha')^n)$  of processes converges stably in law to a process  $(\bar{W}, \bar{W}')$ , independent of  $X$ , and which is a 2-dimensional Wiener process with unit variances 1 and unit covariance  $\sqrt{\gamma}$ .*

*Proof.* The processes  $M^n = M(\varpi, \alpha)^n$  and  $M'^n = M(\varpi, \alpha')^n$  are Lévy processes and martingales, with jumps going uniformly to 0, and with predictable brackets

$$\langle M^n, M^n \rangle_t = \langle M'^n, M'^n \rangle_t = t, \quad \langle M^n, M'^n \rangle_t = \frac{\sqrt{\gamma_n(\varpi, \alpha')}}{\sqrt{\gamma_n(\varpi, \alpha)}} t.$$

Observe also that  $\alpha' \Delta_n^\varpi \geq \alpha \Delta_n^\varpi$ , hence  $\gamma_n(\varpi, \alpha') \leq \gamma_n(\varpi, \alpha)$ . The remaining results then follow (see Jacod and Shiryaev (2003), Chapter VII).  $\square$

**Theorem 6.** *If  $\alpha' > \alpha$  and if  $\frac{\gamma_n(\varpi, \alpha')}{\gamma_n(\varpi, \alpha)} \rightarrow \gamma \in [0, 1]$ , then the sequence*

$$(32) \quad \sqrt{\gamma_n(\varpi, \alpha')} \left( \bar{\beta}_n(t, \varpi, \alpha, \alpha') - \frac{\log(\gamma_n(\varpi, \alpha)/\gamma_n(\varpi, \alpha'))}{\log(\alpha'/\alpha)} \right)$$

*converges stably in law to an  $\mathcal{N}\left(0, \frac{1-\gamma}{t(\log(\alpha'/\alpha))^2}\right)$  variable, independent of  $X$ .*

This result is a simple consequence of the previous lemma, and its proof is the same as the proof of Theorem 1 once the CLT for the processes  $U(\varpi, \alpha)^n$  is established, which we will do later.

So the situation seems hopeless in general: these estimators are not even consistent for estimating the activity index  $\beta$  of  $F$ , because of bias, and to remove the bias we have to know the ratio  $\gamma_n(\varpi, \alpha')/\gamma_n(\varpi, \alpha)$  (or at least its asymptotic behavior in a precise way), and further there is no CLT if this ratio does not converge (a fact which we do not know a priori, of course).

The major difficulty comes from the possible erratic behavior of  $\bar{F}$  near 0. Indeed, we have (6) with  $\beta$  instead of  $\beta_t^i$ , but there are Lévy measures  $F$  satisfying this, and such that for any  $r \in (0, \beta)$  we have  $x_n^r \bar{F}(x_n) \rightarrow 0$  for a sequence  $x_n \rightarrow 0$  (depending on  $r$ , of course). If  $F$  is such, the sequence  $\gamma_n(\varpi, \alpha')/\gamma_n(\varpi, \alpha)$  may have the whole of  $[0, 1]$  as limit points, depending on the parameter values  $\varpi, \alpha, \alpha'$ , and in a completely uncontrolled way for the observer.

So we need some additional assumption on  $F$ . Let us consider two assumptions (the second one is stronger than the first one):

**Assumption 3.**  $\bar{F}$  is regularly varying at 0, with index  $\beta \in (0, 2)$ .

**Assumption 4.** We have

$$(33) \quad \bar{F}(x) = \frac{A}{x^\beta} + o\left(\frac{1}{x^{\beta/2}}\right)$$

as  $x \rightarrow 0$ , for some  $A > 0$ .

**Theorem 7.** a) Under Assumption 3, we have  $\bar{\beta}_n(t, \varpi, \alpha, \alpha') \xrightarrow{\mathbb{P}} \beta$ .

b) Under Assumption 4, the variables  $\Delta_n^{-\varpi\beta} (\bar{\beta}_n(t, \varpi, \alpha, \alpha') - \beta)$  converge stably in law to a  $\mathcal{N}\left(0, \frac{\alpha'^\beta - \alpha^\beta}{tA^2(\log(\alpha'/\alpha))^2}\right)$  variable, independent of  $X$ .

*Proof.* Assumption 3 implies  $\gamma_n(\varpi, \alpha) \rightarrow \infty$  and  $\gamma_n(\varpi, \alpha)/\gamma_n(\varpi, \alpha') \rightarrow (\alpha'/\alpha)^\beta$ , so the previous theorem yields (a). Assumption 4 clearly implies

$$\sqrt{\gamma_n(\varpi, \alpha)} \frac{\log(\gamma_n(\varpi, \alpha)/\gamma_n(\varpi', \alpha'))}{\log(\alpha'/\alpha)} \rightarrow \beta,$$

and also  $\gamma_n(\varpi, \alpha) \sim A/\alpha^\beta \Delta_n^{\varpi\beta}$ , so (b) follows again from the previous theorem.  $\square$

It may of course happen that Assumptions 3 or 4 fail and nevertheless the conclusions of the previous theorem hold for a particular choice of the parameters  $(\varpi, \alpha, \alpha')$ , or for a particular choice of the sequence  $\Delta_n$ . But in view of Theorem 6 and of the previous proof these assumptions are *necessary* if we want those conclusions to hold for *all choices* of  $(\varpi, \alpha, \alpha')$ .

Now, coming back to the original realistic problem, for which only increments of  $X$  are observed, Assumption 2 when  $F_t(\omega, dx) = F(dx)$  for all  $(\omega, t)$  is obviously stronger than Assumption 4, but not much more. The need of stronger assumptions for the original problem comes from the fact that, although when we observe a “large” increment  $\Delta_{\frac{1}{i}}^n X$  it is with a high probability almost equal to a “large” jump, nevertheless the observation of this jump is blurred by the Brownian component and also by a sum of very small jumps. This fact is also the reason why we need some restriction on  $\varpi$  for the original problem, whereas here  $\varpi$  can be arbitrarily large.

**5. Small sample bias correction.** By construction, we are forced by the presence of a continuous semimartingale to rely on a small fraction of the sample – those increments larger than  $\alpha\Delta_n^\varpi$  – for the purpose of estimating  $\beta$ . As a result, the effective sample size utilized by the estimator  $\hat{\beta}_n$  is small, even if we sample at a relatively high frequency. This situation calls for an analysis of the small sample behavior of the estimator.

Such a small sample analysis is out of reach in general but it can be carried out explicitly for the model  $X_t = \sigma W_t + \theta Y_t$  studied in Section 3, where  $Y$  is a symmetric  $\beta$ -stable process and  $W$  is a Wiener process. Let  $g$  denote the density of  $Y_1$ . Here the process  $Y$  is standardized by  $\mathbb{E}(e^{iuY_t}) = e^{-t|u|^\beta/2}$ , so the limit  $\beta \rightarrow 2$  corresponds to the standard normal density,  $\phi$ .

One additional step in the expansion (22) yields as  $x \rightarrow +\infty$  :

$$(34) \quad g(x) = \frac{c_\beta}{x^{\beta+1}} + \frac{d_\beta}{x^{2\beta+1}} + O\left(\frac{1}{x^{3\beta+1}}\right)$$

and for the tail of the distribution

$$(35) \quad \bar{G}(x) = \mathbb{P}(|Y_1| > x) = 2 \int_x^{+\infty} g(v) dv = \frac{2c_\beta}{\beta x^\beta} + \frac{d_\beta}{\beta x^{2\beta}} + O\left(\frac{1}{x^{3\beta}}\right)$$

where the coefficients of the expansion are

$$(36) \quad c_\beta = \frac{\Gamma(\beta+1)}{2\pi} \sin\left(\frac{\pi\beta}{2}\right) \quad \text{and} \quad d_\beta = -\frac{\Gamma(2\beta+1)}{8\pi} \sin(\pi\beta).$$

This parametrization corresponds in terms of the general notation of the paper to

$$(37) \quad A_t = A = 2\theta^\beta c_\beta / \beta.$$

Consider now the tail probability  $P_n$  at the cutoff level  $\alpha\Delta_n^\varpi$ . The probability  $P_n$  determines the limiting behavior of the  $U$ s, since  $U(\varpi, \alpha)_t^n \sim (t/\Delta_n)P_n$ . We have:

$$\begin{aligned} P_n &= 2 \int_{-\infty}^{+\infty} \int_{\alpha\Delta_n^\varpi}^{+\infty} \frac{1}{\theta\Delta_n^{1/\beta}} g\left(\frac{x-y}{\theta\Delta_n^{1/\beta}}\right) dx \frac{1}{\sigma\Delta_n^{1/2}} \phi\left(\frac{y}{\sigma\Delta_n^{1/2}}\right) dy \\ &= 2 \int_{-\infty}^{+\infty} \int_{\frac{\alpha}{\theta}\Delta_n^{\varpi-1/\beta}}^{+\infty} \left(1 - \frac{\sigma}{\alpha}\Delta_n^{1/2-\varpi}u\right) g(v) dv \phi(u) du \\ &= \int_{-\infty}^{+\infty} \overline{G}\left(\frac{\alpha}{\theta}\Delta_n^{\varpi-1/\beta}\left(1 - \frac{\sigma}{\alpha}\Delta_n^{1/2-\varpi}u\right)\right) \phi(u) du. \end{aligned}$$

So, with

$$(38) \quad \left(1 - \frac{\sigma}{\alpha}\Delta_n^{1/2-\varpi}u\right)^{-\beta} = 1 + \frac{u\beta\sigma\Delta_n^{1/2-\varpi}}{\alpha} + \frac{u^2\beta(\beta+1)\sigma^2\Delta_n^{1-2\varpi}}{2\alpha^2} + O(\Delta_n^{3/2-3\varpi})$$

and  $\int_{-\infty}^{+\infty} u\phi(u) du = 0$  and  $\int_{-\infty}^{+\infty} u^2\phi(u) du = 1$ , we see from (35) that

$$(39) \quad \begin{aligned} P_n &= \frac{2c_\beta\theta^\beta}{\beta\alpha^\beta}\Delta_n^{1-\varpi\beta} \left(1 + \frac{\beta(\beta+1)\sigma^2}{2\alpha^2}\Delta_n^{1-2\varpi} + \frac{d_\beta\theta^\beta}{2c_\beta\alpha^\beta}\Delta_n^{1-\varpi\beta}\right) \\ &+ \text{smaller terms.} \end{aligned}$$

The behavior of  $P_n$  suggested by the leading term in that expression,  $(2c_\beta\theta^\beta/(\beta\alpha^\beta))\Delta_n^{1-\varpi\beta}$ , is the one we have used to define the estimator  $\hat{\beta}_n$  in (16) by exploiting the dependence of that leading term on  $\alpha$ . The first correction term in (39),  $(\beta(\beta+1)\sigma^2/(2\alpha^2))\Delta_n^{1-2\varpi}$ , is due to the interaction between the Wiener and the stable processes, while the second,  $(d_\beta\theta^\beta/(2c_\beta\alpha^\beta))\Delta_n^{1-\varpi\beta}$ , is due to the more accurate approximation of the tail of the stable process in (34) compared to the leading order term in (22).

To understand intuitively the need for the first term, suppose that the cutoff level corresponds to seven standard deviations of the continuous part of the semimartingale. There is very little probability that the Wiener process, alone, will generate an increment that large. On the other hand, when we count the increments due to the jump process alone, we are missing increments of the sum of the continuous and discontinuous parts where, say, the Wiener process is responsible for a one standard deviation move, and the jump process for a six standard deviation move, of the same sign. We are also missing increments where the jump process gives an eight standard deviation move and the Wiener process a one standard deviation move, of the opposite sign. The two effects partly compensate each other and indeed the term in  $u$  in (38) leads to an integral whose value is zero. But the next effect, in  $u^2$ , leads

to a net increase in the total number of increments that are larger than the cutoff when the interaction between the Wiener and jump processes is accounted for.

Asymptotically, the first of the two correcting terms in (39) is the largest, since  $1 \gg \Delta_n^{1-2\varpi} \gg \Delta_n^{1-\varpi\beta}$ , but in small samples a large value of the scaling parameter  $\theta$  relative to  $\sigma$  can make their magnitudes comparable. Using (39) at two different values  $\alpha$  and  $\alpha'$ , we obtain

$$(40) \quad \hat{\beta}_n \sim \beta + \frac{1}{\log(\alpha'/\alpha)} \left\{ \frac{\beta(\beta+1)\sigma^2}{2} \left( \frac{1}{\alpha^2} - \frac{1}{\alpha'^2} \right) \Delta_n^{1-2\varpi} + \frac{d_\beta \theta^\beta}{2c_\beta} \left( \frac{1}{\alpha^\beta} - \frac{1}{\alpha'^\beta} \right) \Delta_n^{1-\varpi\beta} \right\}.$$

This suggests a small sample bias correction for the estimator  $\hat{\beta}_n$ , obtained by subtracting from  $\hat{\beta}_n$  an estimator of the two correction terms on the right hand side of (40). As we will see in simulations below, the two correction terms are quite effective in practice.

Further, we note that the two correction terms in (40), of respective orders  $\Delta_n^{1-2\varpi}$  and  $\Delta_n^{1-\varpi\beta}$ , are asymptotically negligible at the rate  $\Delta_n^{-\varpi\beta/2}$  at which the central limit occurs. This is due to the restrictions on the choice of  $\varpi$  imposed by Theorem 2. Consequently, the bias-corrected estimator has the same asymptotic distribution as the original estimator.

More generally, we have

$$(41) \quad \hat{\beta}_n \sim \beta + \frac{1}{\log(\alpha'/\alpha)} \left\{ \frac{\int_0^t A_s \sigma_s^2 ds}{\bar{A}_t} \frac{\beta(\beta+1)}{2} \left( \frac{1}{\alpha^2} - \frac{1}{\alpha'^2} \right) \Delta_n^{1-2\varpi} + \frac{\int_0^t A_s^2 ds}{\bar{A}_t} \frac{\beta d_\beta}{4c_\beta^2} \left( \frac{1}{\alpha^\beta} - \frac{1}{\alpha'^\beta} \right) \Delta_n^{1-\varpi\beta} \right\}.$$

To implement the bias correction in practice, we need to estimate the terms  $(1/\bar{A}_t) \int_0^t A_s \sigma_s^2 ds$  and  $(1/\bar{A}_t) \int_0^t A_s^2 ds$ . In the case of a stable symmetric process  $A_s = A$  and so  $(1/\bar{A}_t) \int_0^t A_s^2 ds = A = 2\theta^\beta c_\beta / \beta$ . We can then replace  $(1/\bar{A}_t) \int_0^t A_s \sigma_s^2 ds$  by  $(1/t) \int_0^t \sigma_s^2 ds$ , and use any standard estimator of the integrated volatility. In general, we have

$$(42) \quad U(\varpi, \alpha)_t^n \sim \frac{1}{\alpha^\beta} \Delta_n^{-\varpi\beta} \left( \bar{A} + \frac{1}{\alpha^2} \overline{A\sigma^2} G_1(\beta) \Delta_n^{1-2\varpi} + \frac{1}{\alpha^\beta} \bar{A}^2 G_2(\beta) \Delta_n^{1-\varpi\beta} \right)$$

$$(43) \quad \Delta_n^{\varpi\beta} U(\varpi, \alpha)_t^n \sim a_0 \frac{1}{\alpha^\beta} + a_1 \frac{1}{\alpha^{2+\beta}} + a_2 \frac{1}{\alpha^{2\beta}}$$

where  $\bar{A} = \int_0^t A_s \sigma_s^2 ds$ ,  $\overline{A\sigma^2} = \int_0^t A_s \sigma_s^2 ds$  and  $\bar{A}^2 = \int_0^t A_s^2 ds$ . We can estimate the unknown coefficients  $a_0$ ,  $a_1$  and  $a_2$  in expression (43) by a straightforward linear regression of  $\Delta_n^{\varpi\beta} U(\varpi, \alpha)_t^n$  on  $1/\alpha^\beta$ ,  $1/\alpha^{2+\beta}$  and  $1/\alpha^{2\beta}$ . For the purpose of running that regression, we use different cutoffs levels  $\alpha$  and compute the corresponding



Sampling $\Delta_n$		1 sec	1 sec	1 sec	1 sec	5 sec
Tail Probability		0.25%	0.5%	1.0%	2.5%	1.0%
$\beta = 1.5$	Sample Mean	1.52	1.51	1.50	1.52	1.53
	Sample StDev	(0.26)	(0.18)	(0.13)	(0.08)	(0.25)
	Asymp StDev	(0.26)	(0.18)	(0.13)	(0.08)	(0.24)
$\beta = 1.25$	Sample Mean	1.27	1.26	1.25	1.26	1.27
	Sample StDev	(0.23)	(0.16)	(0.11)	(0.07)	(0.19)
	Asymp StDev	(0.23)	(0.16)	(0.11)	(0.07)	(0.19)
$\beta = 1.0$	Sample Mean	1.01	1.01	1.00	1.00	1.01
	Sample StDev	(0.19)	(0.14)	(0.10)	(0.06)	(0.14)
	Asymp StDev	(0.19)	(0.14)	(0.10)	(0.06)	(0.14)
$\beta = 0.75$	Sample Mean	0.76	0.76	0.75	0.75	0.76
	Sample StDev	(0.16)	(0.11)	(0.08)	(0.05)	(0.11)
	Asymp StDev	(0.16)	(0.11)	(0.08)	(0.05)	(0.11)
$\beta = 0.5$	Sample Mean	0.51	0.50	0.50	0.50	0.50
	Sample StDev	(0.13)	(0.09)	(0.06)	(0.04)	(0.08)
	Asymp StDev	(0.13)	(0.09)	(0.06)	(0.04)	(0.08)
$\beta = 0.25$	Sample Mean	0.25	0.25	0.25	0.25	0.25
	Sample StDev	(0.09)	(0.06)	(0.04)	(0.04)	(0.05)
	Asymp StDev	(0.09)	(0.06)	(0.04)	(0.04)	(0.05)
$\beta = 0$	Sample Mean	0.01	0.01	0.01	0.01	0.02
	Sample StDev	(0.02)	(0.01)	(0.007)	(0.005)	(0.01)

**Table 1**

*Monte Carlo Simulations of the Estimator  $\hat{\beta}_n$  Based on Two Levels of Truncation for  $\beta$ -Stable Processes and a Compound Poisson Process ( $\beta = 0$ ).*

number of increments exceeding that level,  $U(\varpi, \alpha)_t^n$  and the first-stage estimate of  $\beta$ . Given estimates of the regression coefficients, we have a generalized bias correction procedure based on subtracting from  $\hat{\beta}_n$  the terms on the right of hand side of

$$(44) \quad \hat{\beta}_n - \beta \sim \frac{1}{\log(\alpha'/\alpha)} \left\{ \frac{a_1}{a_0} \left( \frac{1}{\alpha^2} - \frac{1}{\alpha'^2} \right) + \frac{a_2}{a_0} \left( \frac{1}{\alpha^\beta} - \frac{1}{\alpha'^\beta} \right) \right\}$$

evaluated at the regression estimates of  $a_0$ ,  $a_1$  and  $a_2$ .

**6. Monte Carlo simulations.** We now report simulation results documenting the finite sample performance of the estimator  $\hat{\beta}_n$  in finite samples. We calibrate the

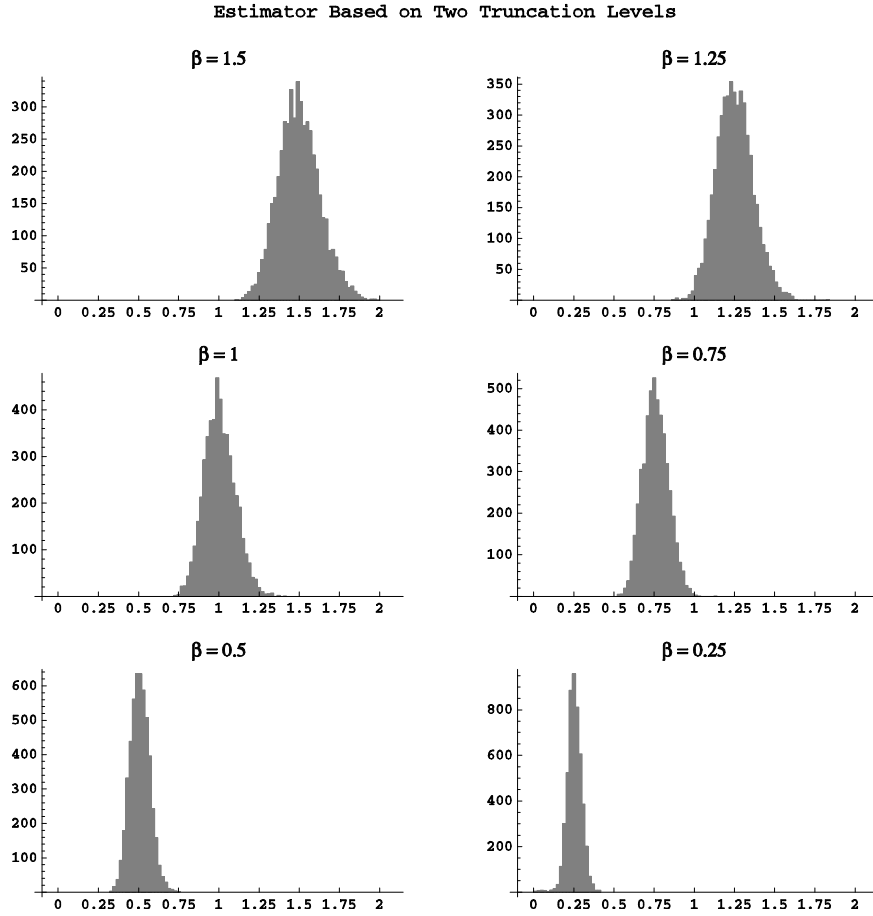


FIG. 1. Monte Carlo Distributions of the Estimator  $\hat{\beta}_n$  Based on Two Levels of Truncation for  $\beta$ -Stable Processes ( $0 < \beta < 2$ ).

values to be realistic for a liquid stock trading on the NYSE. We use an observation length of  $T = 1$  day, consisting of 6.5 hours of trading, that is  $n = 23,400$  seconds.

The averages and standard deviations of the estimator  $\hat{\beta}_n$ , which is based on two different levels of truncation  $\alpha$  and  $\alpha'$ , are reported in Table 1 for various values of  $\beta$  up to 1.5, and include a continuous (Brownian) part. The table reports the results of 5,000 simulations. The data generating process is the stochastic volatility model  $dX_t = \sigma_t dW_t + \theta dY_t$ , with  $\sigma_t = v_t^{1/2}$ ,  $dv_t = \kappa(\eta - v_t)dt + \gamma v_t^{1/2} dB_t + dJ_t$ ,  $E[dW_t dB_t] = \rho dt$ ,  $\eta^{1/2} = 0.25$ ,  $\gamma = 0.5$ ,  $\kappa = 5$ ,  $\rho = -0.5$ ,  $J$  is a compound Poisson jump process with jumps that are uniformly distributed on  $[-30\%, 30\%]$  and  $X_0 = 1$ . The jump process  $Y$  is either a  $\beta$ -stable process with  $\beta = 1.5, 1.25, 1.0, 0.75, 0.5$  and  $0.25$ , or a compound Poisson process (which has finite activity and is marked  $\beta = 0$  in the table) with fixed jump size 0.10. The estimator is implemented with  $\alpha = 5\eta$ ,

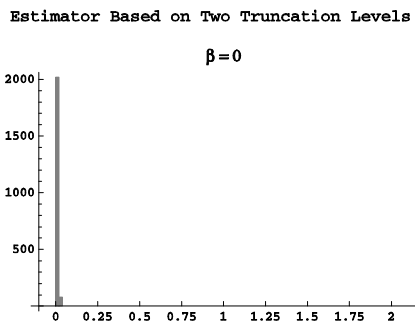


FIG. 2. Monte Carlo Distribution of the Estimator  $\hat{\beta}_n$  Based on Two Levels of Truncation for a Compound Poisson Process ( $\beta = 0$ ).

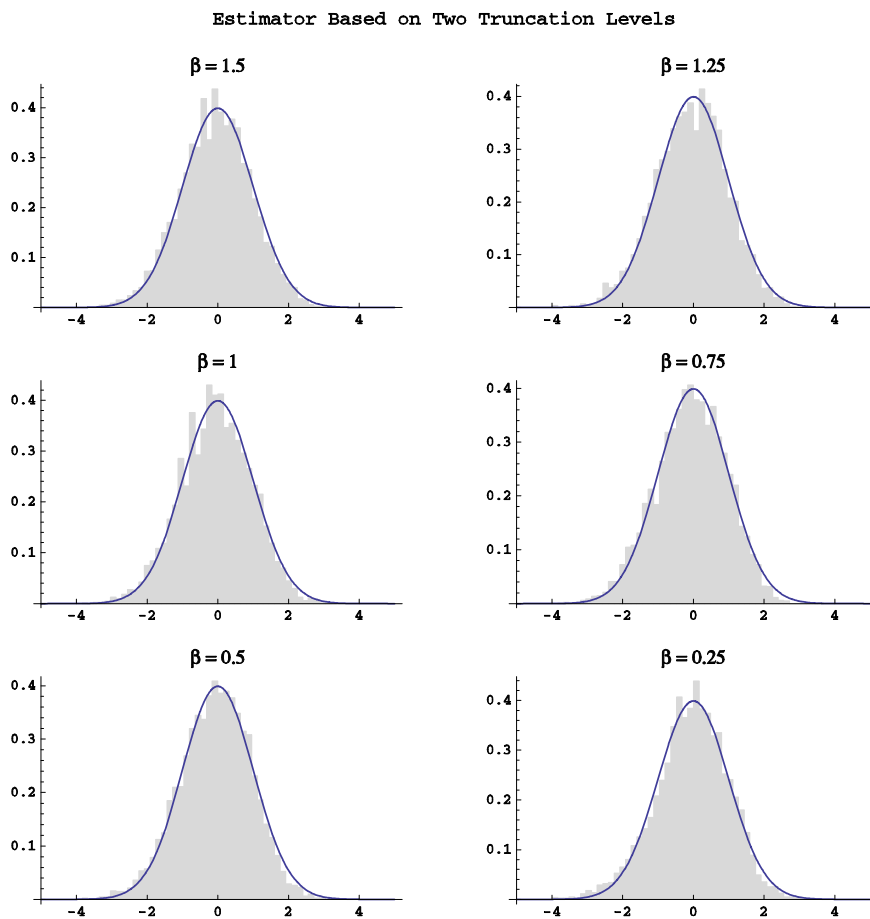


FIG. 3. Standardized Monte Carlo and Asymptotic Distributions of the Estimator  $\hat{\beta}_n$  Based on Two Levels of Truncation for  $\beta$ -Stable Processes ( $0 < \beta < 2$ ).

$\alpha' = 10\eta$  and  $\varpi = 0.20$ . Given  $\eta$  and  $\alpha$ , the scale parameter  $\theta$  (or equivalently  $A$ ) of the stable process in simulations is calibrated to deliver the various values of the tail probability  $\mathbb{P}(|\Delta Y_t| \geq \alpha \Delta_n^\varpi)$  reported in the columns of the table; for the Poisson process, it is the value of the arrival rate parameter  $\lambda$  that is set to generate the desired level of jump tail probability.

In each row, the top number is the average value of the estimator  $\hat{\beta}_n$  across the simulations, after inclusion of the bias correction discussed in Section 5, while the number below in parentheses is the standard deviation of the estimator across the same simulations; the third number in parentheses is the estimated asymptotic standard error based on the limiting distribution given in the sections above. A higher tail probability in the columns has the effect of generating more increments from the jump process that exceed the cutoff level, which makes more observations available and correspondingly reduce the standard deviation of the estimates.

As the results show,  $\hat{\beta}_n$  picks up on average fairly accurately the true value of  $\beta$ . As  $\beta$  gets too close to 2, the  $\beta$ -stable jump process starts to approximate too closely the behavior of the Brownian motion and the performance of the estimator deteriorates.

Further simulations (not reported to save space) suggest that the estimator is not overly sensitive to the selection of the truncation levels  $(\alpha, \alpha')$  within a reasonable range. Histograms of the distribution of the estimator  $\hat{\beta}_n$  are shown in Figure 1 for the same values of  $\beta$ ; the figures are based on the 1% level of jump tail probability. The histogram reports the raw, unstandardized, values of  $\hat{\beta}_n$ , which are not expected to be asymptotically normal unlike the standardized versions.

A special mention should be made about the compound Poisson case, which corresponds to  $\beta = 0$  and does not satisfy Assumption 2. As discussed in Remark 7, there is no central limit theorem in this case and  $\hat{\beta}_n$  should be equal to 0 for  $n$  large enough. This is what happens in simulations in the vast majority of cases, as shown by the histogram for  $\beta = 0$  reported in Figure 2.

The asymptotic distribution is an accurate guide for the small samples as shown in the standardized distributions in Figure 3. The estimator in the histograms is standardized according to the asymptotic distribution given in Theorem 3, and the solid curve in the figures is the limiting  $\mathcal{N}(0, 1)$  density. As the figures show, the asymptotic distribution is a fairly accurate guide for the small samples. This is in spite of the relatively small number of (large) increments that are effectively used by the estimator, combined with the facts that some large increments are kept even though they may not have contained a large jump, or conversely smaller increments may have contained two or more large, cancelling, jumps, or the Wiener process may have combined with the pure jump process to produce a larger increment. Asymptotically, these effects do not show up at the leading order in  $\Delta_n$ , but are present in small samples and appear to be effectively captured by the bias correcting term.

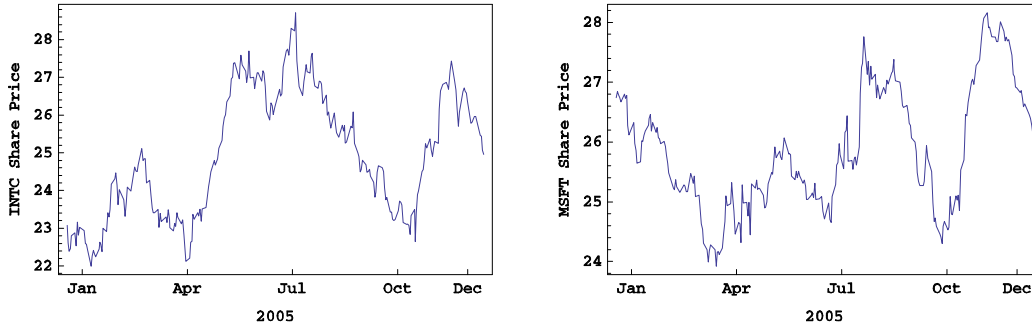


FIG. 4. Time series of INTC and MSFT stock prices, all trading days in 2005.

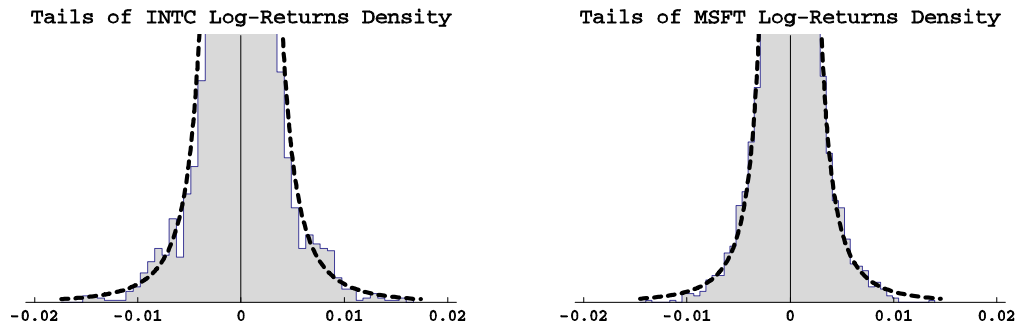


FIG. 5. Tails of the marginal density of INTC and MSFT log-returns, all trading days in 2005 sampled at 5 second intervals. The dashed curve is the leading part of the Lévy measure, that is  $F_t'(dx)$  given in (8), with jump activity index  $\beta$ , estimated for each stock by  $\hat{\beta}_n$ .

**7. Empirical application.** We now implement the estimator  $\hat{\beta}_n$  for the two most actively traded stocks in the Dow Jones Industrial Average index, Intel (INTC) and Microsoft (MSFT), and each trading day in 2005; the data source is the TAQ database. Each day, we collect all transactions on the NYSE or NASDAQ, from 9:30am until 4:00pm, for each one of these stocks. We sample in calendar time every 5 and 15 seconds. We use filters to eliminate clear data errors (price set to zero, etc.) as is standard in the empirical high frequency literature.

The two time series are plotted in Figure 4. Figure 5 contains a histogram of the tails of the unconditional densities of the log-returns from the two stocks. Comparing the two figures, we see that it is quite possible to have a standard time series plot display no obvious evidence of large moves (Figure 5), while the tails of the distribution look substantially fatter than normal (Figure 5), as confirmed by the descriptive statistics in Table 2 for the two log-returns series. All together, this ev-

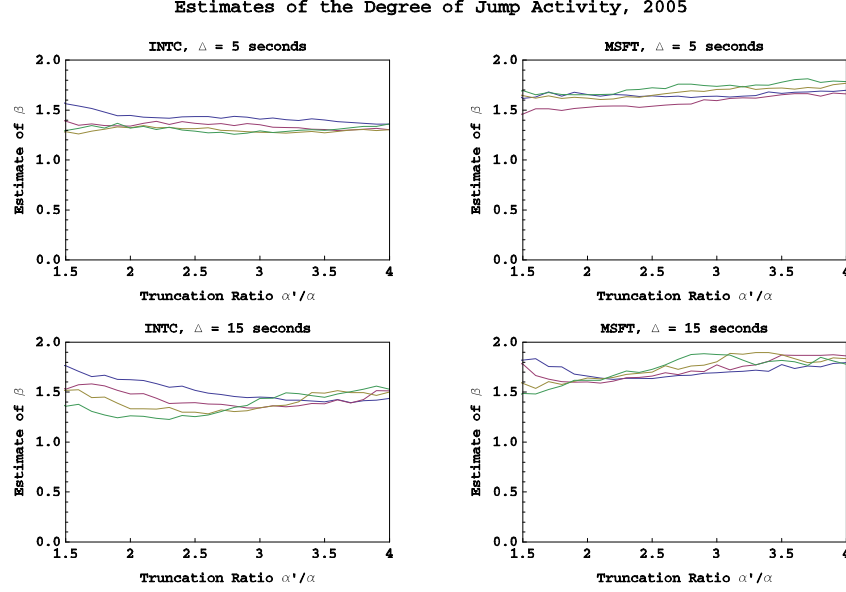


FIG. 6. Estimates of  $\beta$  obtained from  $\hat{\beta}_n$  at 5 and 15 seconds for Intel and Microsoft, all 2005 transactions.

idence points in the direction of small, active jumps, of the type that we seek to uncover using our estimator of  $\beta$ .

More formally, we compute the statistic  $\hat{S}_n$  of Aït-Sahalia and Jacod (2006) to test for the presence of jumps in the data. Over the different sampling frequencies considered (ranging from 2 seconds to 1 minute), the largest value of the statistic  $\hat{S}_n$  we obtain for the different quarters and the two stocks is 1.24. Since the asymptotic value of  $\hat{S}_n$  is 1 (resp. 2) when jumps are present (resp. absent), this provides further evidence for the presence of jumps.

Table 3 reports the estimates produced by  $\hat{\beta}_n$  for the two stocks, for the full year. For each stock and sampling intervals  $\Delta_n$  of 5 and 15 seconds, respectively, the left columns in the table use a level of truncation  $\alpha$  that is set to 7 estimated standard deviations of the continuous part of the process, a second level  $\alpha' = 2\alpha$ , and a truncation rate  $\varpi = 0.20$ . In light of the different possible choices of  $\alpha$  and  $\alpha'$ , a simple variance-reducing procedure consists in averaging the estimators obtained over the different  $(\alpha, \alpha')$ . The columns marked AVG report the results of averaging the estimates over values of  $\alpha$  ranging from 7 to 9 estimated standard deviations of the continuous part of the process) and over values of  $\alpha'$  ranging from 1.5 to 4 times  $\alpha$ .  $\hat{\beta}_n$  is the estimator defined in (16). The estimator  $\tilde{\beta}_n$  in the table denotes the estimator obtained by applying to  $\hat{\beta}_n$  the regression-based bias-correction procedure described in Section 5, implemented using the same range of truncation levels as for the average AVG. Standard errors of the estimators are in parentheses.

INTC						
	Mean	Stdev	Skew	Kurt	Min	Max
Qtr 1	$-1 \times 10^{-7}$	0.00045	1.77	204	-0.020	0.029
Qtr 2	$7 \times 10^{-7}$	0.00048	2.36	943	-0.042	0.041
Qtr 3	$-4 \times 10^{-7}$	0.00046	-13.0	2,423	-0.060	0.040
Qtr 4	$4 \times 10^{-8}$	0.00046	-2.03	925	-0.041	0.041
All Year	$5 \times 10^{-8}$	0.00046	-2.68	1,142	-0.060	0.041
MSFT						
	Mean	Stdev	Skew	Kurt	Min	Max
Qtr 1	$-6 \times 10^{-7}$	0.00037	0.62	63	-0.010	0.016
Qtr 2	$-1 \times 10^{-7}$	0.00042	0.13	954	-0.039	0.039
Qtr 3	$2 \times 10^{-7}$	0.00038	-0.35	76	-0.017	0.010
Qtr 4	$9 \times 10^{-8}$	0.00038	0.12	115	-0.018	0.018
All Year	$-4 \times 10^{-8}$	0.00039	0.11	385	-0.039	0.039

**Table 2**  
*Descriptive Statistics for Intel and Microsoft Transactions in 2005.*

$\Delta_n$	INTC				MSFT			
	5 sec		15 sec		5 sec		15 sec	
$\alpha$	7	AVG	7	AVG	7	AVG	7	AVG
$\alpha'/\alpha$	2	AVG	2	AVG	2	AVG	2	AVG
$\hat{\beta}_n$	1.34	1.32	1.48	1.40	1.53	1.66	1.60	1.73
$\tilde{\beta}_n$		1.34		1.41		1.68		1.65
	(0.04)	(0.005)	(0.05)	(0.01)	(0.05)	(0.007)	(0.06)	(0.02)

**Table 3**  
*Estimates of  $\beta$  from all Intel and Microsoft Transactions in 2005.*

We find evidence of infinitely active jumps. The estimated values of  $\beta$  are about 1.4 for INTC and 1.7 for MSFT. These correspond also to the average values  $\hat{\beta}_n$  we obtain if we average the four quarterly estimates, and also average them over different values of  $\alpha$  and  $\alpha'/\alpha$ . Clearly, the nature of these infinitely actively jumps cannot be assessed by mere visual inspection of the time series in Figure 4, as would have been the case if only large, infrequent jumps, were present.

Figure 6 shows the values of the estimates computed over the full year for various values of  $\alpha$  (the four curves on each plot correspond respectively to 6, 7, 8 and 9 estimated standard deviations of the continuous part of the process) and values of

$\alpha'$  ranging from 1.5 to 4 times  $\alpha$  (the horizontal axis represents the ratio  $\alpha'/\alpha$ ). The continuous standard deviations are computed from the small increments of the process and updated each quarter. The figure shows that, for each stock, we obtain comparable values for  $\beta$  when varying  $\alpha$  and  $\alpha'$ . The top panel of the plot shows the estimates at  $\Delta_n = 5$  seconds, the bottom one at  $\Delta_n = 15$  seconds. As expected, the lower frequency estimates are more variable, since they rely on a smaller number of large increments in total, and the variance of  $\hat{\beta}_n$  is proportional to the difference between one over the number of increments retained at the two cutoff levels.

One last comment regarding the high sampling frequency and its interaction with market microstructure noise in the data. There is no doubt that market microstructure noise is a concern at sampling intervals of a few seconds. However, in the present context, we are only using the large increments of the process, namely those greater than  $\alpha\Delta^\varpi$ . Such increments are more likely to represent a true price movement rather than noise, especially after the standard data cleaning consisting of removing price errors have been applied. Similarly, the cutoff points are well beyond the level of bid/ask bounces.

Furthermore, all we are doing is counting those increments, not exploiting their magnitude, so provided that the noise does not substantially affect the likelihood that an increment is above or below the cutoff, these estimates are unlikely to be seriously affected by the noise. But this paper represents only a first attempt at measuring the degree of jump activity. We leave to future work the development of estimators that are fully robust to market microstructure noise.

**8. Technical results and proofs.** Our basic semimartingale  $X$  satisfies (2) and Assumptions 1 and 2. Note that all three theorems 1, 2, and 3 are “local” in time, so by a standard localization procedure (see for example Aït-Sahalia and Jacod (2006)) we can indeed assume a strengthened version of our assumptions, namely:

**Assumption 5.** *The processes  $b$  and  $\sigma$  are bounded by some constant  $L$ .*

**Assumption 6.** *We have Assumption 2 with  $L_t(\omega) = L$  being constant.*

In all the sequel, the letter  $K$  denotes a constant which changes from line to line and may depend on  $X$  and its characteristics, and also on the parameters  $\varpi, \alpha, \alpha'$ , and we write  $K_p$  if we want to emphasize its dependency on some other parameter  $p$ . We use the shorthand notation  $\mathbb{E}_{i-1}^n$  and  $\mathbb{P}_{i-1}^n$ , resp.  $\mathbb{E}_t$  and  $\mathbb{P}_t$  or the conditional expectation and probability w.r.t.  $\mathcal{F}_{(i-1)\Delta_n}$ , resp.  $\mathcal{F}_t$ .

Before proceeding, we mention a number of elementary consequences of Assumption 6, to be used many times: for all  $u \in (0, 1]$  and  $v \in (0, 2]$  and  $x, y \in (0, 1]$ , we



have (recall the notation (4)):

$$(45) \quad \left\{ \begin{array}{l} \overline{F}_t''(x) \leq \frac{K}{x^{\beta'}}, \quad |\overline{F}_t(x) - \frac{A_t}{x^\beta}| \leq \frac{K}{x^{(\beta-\gamma)\vee\beta'}}, \quad \overline{F}_t(x) \leq \frac{K}{x^\beta}, \\ \int_{\{|x|\leq u\}} x^2 F_t(dx) \leq K u^{2-\beta} \\ \int_{\{|x|>u\}} (|x|^v \wedge 1) F_t(dx) \leq \begin{cases} K_v & \text{if } v > \beta \\ K_v \log(1/u) & \text{if } v = \beta \\ K_v u^{v-\beta} & \text{if } v < \beta, \end{cases} \\ \overline{F}_t(x) - \overline{F}_t(x+y) \leq \frac{K}{x^\beta} (1 \wedge \frac{y}{x} + x^{\gamma \wedge (\beta-\beta')}). \end{array} \right.$$

8.1. *Estimates for stable processes.* In this subsection we consider a symmetric stable process  $Y$  with Lévy measure (21). For each  $\delta \in (0, 1]$  we set

$$(46) \quad Y(\delta)'_t = Y_t - \sum_{s \leq t} \Delta Y_s 1_{\{|\Delta Y_s| > \delta\}}.$$

**Lemma 2.** *There is a constant  $K$  depending on  $(A, \beta)$ , such that for all  $s > 0$ :*

$$(47) \quad \mathbb{P}(|Y(\delta)'_s| > \delta/2) \leq K s^{4/3} / \delta^{4\beta/3}.$$

*Proof.* We use the notation (21) and (22). We set  $\theta = s\overline{F}(\delta/2) = sA(2/\delta)^\beta$ , and consider the processes  $Y' = Y(\delta)'$ ,  $Y'' = Y(\delta/2)'$  and  $Z_t = \sum_{r \leq t} 1_{\{|\Delta Y_r| > \delta/2\}}$ . Introduce also the sets

$$D = \{|Y_s| > \delta/2\}, \quad D' = \{|Y'_s| > \delta/2\}, \quad B = \{Z_s = 1\}, \quad B' = \{Z_s = 0\}.$$

It is of course enough to prove the result for  $s/\delta^\beta$  small, so below we assume  $\theta \leq 1/2$ .

By scaling,  $\mathbb{P}(D) = \overline{G}(\delta s^{-1/\beta}/2)$ , so (22) yields

$$(48) \quad |\mathbb{P}(D) - \theta| \leq K\theta^2.$$

Next,  $Z_s$  is a Poisson variable with parameter  $\theta \leq 1/2$ , hence

$$(49) \quad |\mathbb{P}(B) - \theta| \leq K\theta^2.$$

Since  $Y''$  is a purely discontinuous Lévy process without drift and whose Lévy measure is the restriction of  $F$  to  $[-\delta', \delta']$ , we have

$$(50) \quad \mathbb{E}((Y''_s)^2) = s \int_{\{|x| \leq \delta/2\}} x^2 F(dx) \leq K\theta\delta^2.$$

The two processes  $Y''$  and  $Z$  are independent, and conditionally on  $B$  the law of the variable  $Z_s$  is  $F$  restricted to  $\{|x| > \delta/2\}$  and normalized by  $\theta$ , hence

$$\begin{aligned}
\mathbb{P}(B \cap D^c) &= e^{-\theta} s \int_{\{|x| > \frac{\delta}{2}\}} F(dx) \mathbb{P}(|Y_s'' + x| \leq \frac{\delta}{2}) \\
&\leq e^{-\theta} s \int_{\{|x| > \delta/2\}} F(dx) \mathbb{P}(|Y_s''| \geq |x| - \frac{\delta}{2}) \\
&\leq s \left( F(\{\frac{\delta}{2} < |x| \leq \frac{\delta}{2} (1 + \theta^{1/3})\}) + F(\{|x| > \frac{\delta}{2}\}) \mathbb{P}(|Y_s''| > \frac{\delta}{2} \theta^{1/3}) \right) \\
(51) \quad &\leq K\theta \left( \theta^{1/3} + \frac{1}{\delta^2 \theta^{2/3}} \mathbb{E}((Y_s'')^2) \right) \leq K \theta^{4/3},
\end{aligned}$$

where we have used (21) and (50) for the last two inequalities.

Now, we have

$$\mathbb{P}(D \cap B^c) = \mathbb{P}(D) - \mathbb{P}(B) + \mathbb{P}(B \cap D^c).$$

Observe also that  $D \cap B' = D' \cap B'$ , and  $D'$  and  $B'$  are independent, hence

$$\mathbb{P}(D') = \frac{\mathbb{P}(D' \cap B')}{\mathbb{P}(B')} = \frac{\mathbb{P}(D \cap B')}{\mathbb{P}(B')} \leq \frac{\mathbb{P}(D \cap B^c)}{\mathbb{P}(B')} \leq K\mathbb{P}(D \cap B^c)$$

because  $\mathbb{P}(B') = e^{-\theta} \geq e^{-1/2}$ . The last two displays, plus (48) and (49) and (51) give us  $\mathbb{P}(D') \leq K\theta^{4/3}$ , hence (47).  $\square$

*8.2. Estimates for semimartingales.* Below we assume Assumptions 5 and 6, without special mention. The semimartingale  $X$  can be written as

$$X = X_0 + B + X^c + (x1_{\{|x| \leq 1\}}) \star (\mu - \nu) + (x1_{\{|x| > 1\}}) \star \mu,$$

where  $X^c$  is the continuous martingale part and  $\mu$  is the jump measure, and  $B$  and  $\nu$  are as in (2), and the " $\star$ " stands for the stochastic integral with respect to random measures (see Jacod and Shiryaev (2003)). For any  $\delta \in (0, 1]$  we set

$$(52) \quad \begin{cases} X(\delta)_t'' = \sum_{s \leq t} \Delta X_s 1_{\{|\Delta X_s| > \delta\}} \\ X(\delta)_t' = X - X(\delta)_t'' = X_0 + B + X^c + (x1_{\{|x| \leq \delta\}}) \star (\mu - \nu) - B(\delta), \end{cases}$$

where

$$(53) \quad B(\delta)_t = \int_0^t b(\delta)_s ds, \quad b(\delta)_s = \int_{\{\delta < |x| \leq 1\}} x F_s(dx).$$

By Assumption 6,  $F_t'$  and  $F_t''$  are mutually singular, hence there exists a predictable subset  $\Phi$  of  $\Omega \times (0, \infty) \times \mathbb{R}$  such that

$$(54) \quad \begin{cases} F_t'(\omega, \cdot) \text{ is supported by the set } \{x : (\omega, t, x) \notin \Phi\} \\ F_t''(\omega, \cdot) \text{ is supported by the set } \{x : (\omega, t, x) \in \Phi\}. \end{cases}$$

Observe also that if  $\beta' \leq 1$ , we can set

$$(55) \quad B(\delta)'_t = \int_0^t b(\delta)'_s ds, \quad b(\delta)'_s = \int_{\{\delta < |x| \leq 1\}} x F_s''(dx),$$

and  $|b(\delta)'_s| \leq K$ . In this case,  $B(\delta)' = (x1_{\{|x| \leq \delta\}} 1_\Phi) \star \nu$  is of finite variation, and so is  $(x1_{\{|x| \leq \delta\}} 1_\Phi) \star \mu$ . Therefore we have the decomposition

$$(56) \quad \begin{cases} X(\delta)' = X_0 + \widehat{X} + \overline{X}(\delta)^a + \overline{X}(\delta)^b - B(\delta), & \text{where} \\ \overline{X}(\delta)^a = \begin{cases} (x1_{\{|x| \leq \delta\}} 1_\Phi) \star \mu & \text{if } \beta' \leq 1 \\ (x1_{\{|x| \leq \delta\}} 1_\Phi) \star (\mu - \nu) & \text{if } \beta' > 1 \end{cases} \\ \overline{X}(\delta)^b = (x1_{\{|x| \leq \delta\}} 1_{\Phi^c}) \star (\mu - \nu), \\ \widehat{X} = \begin{cases} B + X^c - B(\delta)' & \text{if } \beta' \leq 1 \\ B + X^c & \text{if } \beta' > 1 \end{cases} \end{cases}$$

**Lemma 3.** *We have for all  $\delta \in (0, 1]$ ,  $p \geq 2$ ,  $s, t \geq 0$ :*

$$(57) \quad \begin{cases} \text{(a)} \quad |B(\delta)_{t+s} - B(\delta)_t| \leq \begin{cases} Ks & \text{if } \beta < 1 \\ Ks \log(1/\delta) & \text{if } \beta = 1 \\ Ks\delta^{1-\beta} & \text{if } \beta > 1 \end{cases} \\ \text{(b)} \quad \mathbb{E}_t(|\widehat{X}_{t+s} - \widehat{X}_t|^p) \leq K_p s^{p/2}, \\ \text{(c)} \quad \mathbb{E}_t(|\overline{X}(\delta)_{t+s}^b - \overline{X}(\delta)_t^b|^2) \leq K s \delta^{2-\beta}, \\ \text{(d)} \quad \mathbb{E}_t(|\overline{X}(\delta)_{t+s}^a - \overline{X}(\delta)_t^a|^{\beta'}) \leq K s. \end{cases}$$

*Proof.* (a) follows from (45), and (b) follows from Burkholder-Davis-Gundy inequality and Assumption 5. By (45) again we have

$$\mathbb{E}_t(|\overline{X}(\delta)_{t+s}^b - \overline{X}(\delta)_t^b|^2) = \mathbb{E}_t\left(\int_t^{t+s} dr \int_{\{|x| \leq \delta\}} |x|^2 F_r'(dx)\right) \leq K s \delta^{2-\beta},$$

which is (c). For (d) we single out the two cases  $\beta' \leq 1$  and  $\beta' > 1$ . In the first case, and since for any sequence  $(x_m)$  we have  $|\sum_m x_m|^{\beta'} \leq \sum_m |x_m|^{\beta'}$ , we get

$$\begin{aligned} \mathbb{E}_t(|\overline{X}(\delta)_{t+s}^a - \overline{X}(\delta)_t^a|^{\beta'}) &\leq \mathbb{E}_t\left(\sum_{t < r \leq t+s} |\Delta \overline{X}(\delta)_r^a|^2\right)^{\beta'/2} \\ &= \mathbb{E}_t\left(\int_t^{t+s} dr \int_{\{|x| \leq \delta\}} |x|^{\beta'} 1_\Phi(r, x) F_r''(dx)\right) \leq K s. \end{aligned}$$

In the second case, we apply Burkholder-Davis-Gundy inequality with the exponent  $1 < \beta' < 2$  to get

$$\begin{aligned} \mathbb{E}(|\overline{X}(\delta)_{t+s}^a - \overline{X}(\delta)_t^a|^{\beta'} | \mathcal{F}_t) &\leq \mathbb{E}\left(\left(\sum_{t < r \leq t+s} |\Delta \overline{X}(\delta)_r^a|^2\right)^{\beta'/2} | \mathcal{F}_t\right) \\ &\leq \mathbb{E}\left(\sum_{t < r \leq t+s} |\Delta \overline{X}(\delta)_r^a|^{\beta'} | \mathcal{F}_t\right), \end{aligned}$$

which exactly as before is smaller than  $Ks$ , and (d) is proved.  $\square$

Next, we give a general result on counting processes. Let  $N$  be a counting process (that is, right continuous with  $N_0 = 0$ , piecewise constant, with jumps equal to 1), adapted to  $(\mathcal{F}_t)$ , and with predictable compensator of the form  $\Lambda_t = \int_0^t \lambda ds$ .

**Lemma 4.** *With  $N$  and  $\Lambda$  as above, assume further that  $\lambda_t \leq u$  for some constant  $u > 0$ . Then we have*

$$(58) \quad |\mathbb{P}_t(N_{t+s} - N_t = 1) - \mathbb{E}_t(\Lambda_{t+s} - \Lambda_t)| + \mathbb{P}_t(N_{t+s} - N_t \geq 2) \leq (us)^2.$$

*Proof.* Let  $T_1, T_2, \dots$  be the successive jump times of  $N$  after time  $(i-1)\Delta_n$ . We have

$$\begin{aligned} \mathbb{P}_t(N_{t+s} - N_t \geq 1) &\leq \mathbb{E}_t(N_{t+s} - N_t) = \mathbb{E}_t(\Lambda_{t+s} - \Lambda_t) \leq us, \\ \mathbb{P}_t(N_{t+s} - N_t = 1) &= \mathbb{E}_t(N_{(t+s) \wedge T_1} - N_t) = \mathbb{E}_t(\Lambda_{(t+s) \wedge T_1} - \Lambda_t), \end{aligned}$$

$$\begin{aligned} \mathbb{E}_t(\Lambda_{t+s} - \Lambda_t) - \mathbb{P}_t(N_{t+s} - N_t = 1) &= \mathbb{E}_t(\Lambda_{t+s} - \Lambda_{(t+s) \wedge T_1}) \\ &\leq us \mathbb{P}_t(N_{t+s} - N_t \geq 1) \leq (us)^2 \end{aligned}$$

This gives us the first estimate. Next,

$$\begin{aligned} \mathbb{P}_t(N_{t+s} - N_t \geq 2) &= \mathbb{P}_t(T_2 \leq t+s) = \mathbb{E}_t\left(1_{\{T_1 < t+s\}} \mathbb{P}(T_2 \leq t+s \mid \mathcal{F}_{T_1})\right) \\ &= \mathbb{E}_t\left(1_{\{T_1 < t+s\}} \mathbb{E}\left(\int_{T_1}^{(t+s) \wedge T_2} \lambda_r dr \mid \mathcal{F}_{T_1}\right)\right) \leq us \mathbb{P}_t(N_{t+s} - N_t \geq 1) \leq (us)^2, \end{aligned}$$

hence the second estimate.  $\square$

**Lemma 5.** *With the notation  $N(\delta)_t = \sum_{s \leq t} 1_{\{|\Delta X_s| > \delta\}}$ , for all  $\delta \in (0, 1]$ ,  $\zeta \in (0, \frac{1}{2})$  and  $p \geq 2$  we have*

$$(59) \quad \mathbb{P}_t(N(\delta)_{t+s} - N(\delta)_t \geq 1, |X(\delta)'_{t+s} - X(\delta)'_t| > \delta\zeta) \leq K_p \frac{s^{p/2}}{\zeta^p \delta^p} + K \frac{s^2}{\zeta^2 \delta^{2\beta}}.$$

*Proof.* (57) and Bienaymé-Tchebycheff inequality yield

$$\mathbb{P}_t\left(|\widehat{X}_{t+s} - \widehat{X}_t| > \frac{\delta\zeta}{3}\right) \leq K_p \frac{s^{p/2}}{\zeta^p \delta^p}.$$

Next, (57) yields  $|B(\delta)_{t+s} - B(\delta)_t| \leq K_0 s / \delta$  for some constant  $K_0$ , so we have either  $|B(\delta)_{t+s} - B(\delta)_t| \leq \delta\zeta/3$  or  $K s^{p/2} / \zeta^p \delta^p \geq 1$  for some constant  $K$  independent of  $\zeta$  and  $\delta$ . Then it remains to prove that, with the notation  $M(\delta) = (x1_{\{|x| \leq \delta\}}) \star (\mu - \nu)$ ,

$$(60) \quad \mathbb{P}_t\left(N(\delta)_{t+s} - N(\delta)_t \geq 1, |M(\delta)_{t+s} - M(\delta)_t| > \frac{\delta\zeta}{3}\right) \leq K \frac{s^2}{\zeta^2 \delta^{2\beta}}.$$

For simplicity, write  $N_s = N(\delta)_{t+s} - N(\delta)_t$  and  $M_s = M(\delta)_{t+s} - M(\delta)_t$ . By Bienaymé-Tchebycheff inequality the left side of (60) is not bigger than  $9\mathbb{E}_t(N_s M_s^2) / \delta^2 \zeta^2$ . Now,  $N$  is a counting process and  $M$  is a purely discontinuous

square-integrable martingale, and they have no common jumps, so Itô's formula yields

$$N_s M_s^2 = 2 \int_0^s N_{r-} M_{r-} dM_r + \int_0^s M_{r-}^2 dN_r + \sum_{r \leq s} N_{r-} (\Delta M_r)^2.$$

Moreover, the compensator  $N$  is as in the previous lemma, with  $\lambda_s \leq K\delta^{-\beta}$ , and the predictable quadratic variation of  $M$  is  $\Lambda'_s = \int_0^s \lambda'_r dr$  with  $\lambda'_r \leq K\delta^{2-\beta}$  by (45). Then taking expectations in the above display, and since the first term of the right side above is a martingale, we get

$$\begin{aligned} \mathbb{E}_t(N_s M_s^2) &= \mathbb{E}_t \left( \int_0^s M_r^2 d\Lambda_r + \int_0^s N_r d\Lambda'_r \right) \\ &\leq K\delta^{-\beta} \int_0^s \mathbb{E}_t (M_r^2 + \delta^2 N_r) dr \\ &= K\delta^{-\beta} \int_0^s \mathbb{E}_t (\Lambda'_r + \delta^2 \Lambda_r) dr \leq K\delta^{2(1-\beta)} s^2. \end{aligned}$$

(60) then follows.  $\square$

**Lemma 6.** *Let  $\alpha > 0$ ,  $\varpi \in (0, \frac{1}{2})$  and  $\eta \in (0, \frac{1}{2} - \varpi)$ , and set*

$$(61) \quad \rho = \eta \wedge (\varpi(\beta - \beta') - \beta'\eta) \wedge (\varpi\gamma) \wedge (1 - \varpi\beta - 2\eta).$$

*There is a constant  $K$  depending on  $(\alpha, \varpi, \eta)$ , and also on the characteristics of  $X$ , such that for all  $s \in (0, \Delta_n]$  and  $t \geq 0$  we have*

$$(62) \quad \left| \mathbb{P}_t(|X_{t+s} - X_t| > \alpha\Delta_n^\varpi) - \mathbb{E}_t \left( \int_t^{t+s} \bar{F}_r(\alpha\Delta_n^\varpi) dr \right) \right| \leq K\Delta_n^{1-\varpi\beta+\rho},$$

$$(63) \quad \mathbb{P}_t(\alpha\Delta_n^\varpi < |X_{t+s} - X_t| \leq \alpha\Delta_n^\varpi(1 + \Delta_n^\eta)) \leq K\Delta_n^{1-\varpi\beta+\rho},$$

$$(64) \quad \mathbb{P}_t(|X_{t+s} - X_t| > \alpha\Delta_n^\varpi) \leq K\Delta_n^{1-\varpi\beta}.$$

*Proof.* It is clearly enough to prove the estimates for all  $\Delta_n$  small enough. We write  $\delta_n = \alpha\Delta_n^\varpi$ .

1) Apply (57) and Bienaymé-Tchebycheff inequality to obtain

$$\begin{aligned} \mathbb{P}_t(|\hat{X}_{t+s} - \hat{X}_t| > \alpha\Delta_n^{\varpi+\eta}/3) &\leq K_p \Delta_n^{p(1-2\varpi-2\eta)/2}, \\ \mathbb{P}_t(|\bar{X}(\delta_n)_{t+s}^a - \bar{X}(\delta_n)_t^a| > \alpha\Delta_n^{\varpi+\eta}/3) &\leq K\Delta_n^{1-\varpi\beta'-\eta\beta'}. \end{aligned}$$

Moreover, for all values of  $\beta$  we have  $|B(\delta_n)_{t+s} - B(\delta_n)_t| \leq K\Delta_n^{-\varpi(\beta-1)^+} s \log(1/s)$ , which is smaller than  $\alpha\Delta_n^{\varpi+\eta}/3$  as soon as  $\Delta_n$  is small enough because  $\eta < 1 - \varpi\beta$ . Since  $1 - 2\varpi - 2\eta > 0$ , by choosing  $p$  large enough we deduce from the previous estimates that for all  $\Delta_n$  small enough,

$$(65) \quad \mathbb{P}_t(|X(\delta_n)'_{t+s} + \bar{X}(\delta_n)^b_{t+s} - (X(\delta_n)'_t + \bar{X}(\delta_n)^b_t)| > \alpha\Delta_n^{\varpi+\eta}) \leq K\Delta_n^{1-\varpi\beta+\rho}.$$

2) By Assumption 6 we have  $F'_r(dx) \leq (L'/|x|^{1+\beta})dx$  for some constant  $L'$ . We fix  $n$ . For each  $\omega \in \Omega$  we endow the canonical (Skorokhod) space  $(\Omega', \mathcal{F}', (\mathcal{F}'_t))$  of all càdlàg functions on  $\mathbb{R}_+$  starting from 0 with the (unique) probability measure  $Q_\omega^n$  under which the canonical process  $X'$  is a semimartingale with characteristics  $(0, 0, \nu'_{\delta_n, \omega})$ , where

$$(66) \quad \nu'_{\delta_n, \omega}(dr, dx) = dr \mathbf{1}_{\{|x| \leq \delta_n\}} \left( \frac{L'}{|x|^{1+\beta}} dx - F'_r(\omega, dx) \right).$$

This measure does not depend on  $\omega'$ , hence under  $Q_\omega^n$  the process  $X'$  has independent increments;  $\nu'_{\delta_n, \omega}(dr, dx)$  depends measurably on  $\omega$ , hence  $Q_\omega^n(dw')$  is a transition probability from  $(\Omega, \mathcal{F})$  into  $(\Omega', \mathcal{F}')$ . Then we extend  $X, X'$  and other quantities defined on  $\Omega$  or  $\Omega'$  in the usual way (without changing the symbols) to the product  $\tilde{\Omega} = \Omega \times \Omega'$  endowed with the product  $\sigma$ -field  $\tilde{\mathcal{F}}$ , the product filtration  $(\tilde{\mathcal{F}}_t)$ , and the probability measure  $\tilde{\mathbb{P}}_n(dw, dw') = \mathbb{P}(d\omega) Q_\omega^n(dw')$ .

Because of (66) and (45), and as in Lemma 3,  $\mathbb{E}_{Q_\omega^n}(|X'_{t+s} - X'_t|^2 | \mathcal{F}'_t) \leq Ks\delta_n^{2-\beta}$ . Then we see that, for some constant  $C$  depending on  $\alpha$  and  $\beta$  but not on  $n, i$  and  $\omega$ , we have:

$$(67) \quad Q_\omega^n(|X'_{t+s} - X'_t| > \alpha \Delta_n^{\varpi+\eta} | \mathcal{F}'_t) \leq C \Delta_n^{1-\varpi\beta-2\eta}.$$

3) By well known results on extensions of spaces (see e.g. Jacod and Shiryaev (2003), Section II.7; note that the present extension of the original space is a “very good extension”),  $X'$  is a semimartingale on the extension with characteristics  $(0, 0, \nu'_{\delta_n})$ , where  $\nu'_{\delta_n}((\omega, \omega'), dr, dx) = \nu'_{\delta_n, \omega}(dr, dx)$ , and any semimartingale on the original space is a semimartingale on the extension, with the same characteristics. Moreover  $X$  and  $X'$  have almost surely no common jump, so the sum  $Y(\delta_n)' = \bar{X}(\delta_n)^b + X'$  is a semimartingale with characteristics  $(0, 0, \nu''_{\delta_n})$ , where

$$\nu''_{\delta_n}(dr, dx) = dr \mathbf{1}_{\{|x| \leq \delta_n\}} F'_r(dx) + \nu'_{\delta_n}(dr, dx) = \mathbf{1}_{\{|x| \leq \delta_n\}} \frac{L'}{|x|^{1+\beta}} dr dx,$$

where the last equality comes from (66). It follows that  $Y(\delta_n)'$  is a Lévy process with Lévy measure given above, or in other words it is a version of the process  $Y(\delta_n)'$  of (46) with  $A = L'$ . Hence we deduce from (47) and from the Lévy property of  $Y(\delta_n)'$  that, for any  $A \in \mathcal{F}_t$ ,

$$(68) \quad \tilde{\mathbb{P}}_n(A \cap \{|Y(\delta_n)'_{t+s} - Y(\delta_n)'_t| > \alpha \Delta_n^{\varpi}(1 - 2\Delta_n^\eta)\}) \leq K \Delta_n^{4/3-4\varpi\beta/3} \mathbb{P}(A).$$

Next, for all  $\Delta_n$  small enough, so that  $C \Delta_n^{1-\varpi\beta-2\eta} \leq 1/2$ , we can write:

$$\begin{aligned} & \tilde{\mathbb{P}}(A \cap \{|Y(\delta_n)'_{t+s} - Y(\delta_n)'_t| > \alpha \Delta_n^{\varpi}(1 - 2\Delta_n^\eta)\}) \\ & \geq \tilde{\mathbb{P}}\left(A \cap \left\{|\bar{X}(\delta_n)^b_{t+s} - \bar{X}(\delta_n)^b_t| > \alpha \Delta_n^{\varpi}(1 - \Delta_n^\eta)\right\} \cap \{|X'_{t+s} - X'_t| \leq \alpha \Delta_n^{\varpi+\eta}\}\right) \\ & = \tilde{\mathbb{E}}\left(\mathbf{1}_{A \cap \{|\bar{X}(\delta_n)^b_{t+s} - \bar{X}(\delta_n)^b_t| > \alpha \Delta_n^{\varpi}(1 - \Delta_n^\eta)\}} Q_\omega^n(|X'_{t+s} - X'_t| \leq \alpha \Delta_n^{\varpi+\eta})\right) \\ & \geq \frac{1}{2} \mathbb{P}\left(A \cap \left\{|\bar{X}(\delta_n)^b_{t+s} - \bar{X}(\delta_n)^b_t| > \alpha \Delta_n^{\varpi}(1 - \Delta_n^\eta)\right\}\right), \end{aligned}$$

where the last inequality comes from (67). Then by (68) and by the fact that  $A$  is arbitrary in  $\mathcal{F}_{(i-1)\Delta_n}$ , we deduce (since necessarily  $\rho \leq \frac{1-\varpi\beta}{3}$ ):

$$\mathbb{P}_t \left( |\overline{X}(\delta_n)_{t+s}^b - \overline{X}(\delta_n)_t^b| > \alpha \Delta_n^\varpi (1 - \Delta_n^\eta) \right) \leq K \Delta_n^{4/3-4\varpi\beta/3} \leq K \Delta_n^{1-\varpi\beta+\rho}.$$

In turn, combining this with (65), we readily obtain for all  $\Delta_n$  small enough:

$$\begin{aligned} \mathbb{P}_t (|X(\delta_n)'_{t+s} - X(\delta_n)'_t| > \alpha \Delta_n^\varpi) &\leq K \Delta_n^{1-\varpi\beta} \left( \Delta_n^{1/3-\varpi\beta/3} + \Delta_n^{\varpi(\beta-\beta')-2\eta} \right) \\ (69) \qquad \qquad \qquad &\leq K \Delta_n^{1-\varpi\beta+\rho}. \end{aligned}$$

4) Now we write  $\delta'_n = \alpha \Delta_n^\varpi (1 + \Delta_n^\eta)$  and also

$$\theta_{t,s} = \mathbb{E}_t \left( \int_t^{t+s} \overline{F}_r(\delta_n) dr \right), \quad \theta'_{t,s} = \mathbb{E}_t \left( \int_t^{t+s} \overline{F}_r(\delta'_n) dr \right),$$

and the following two counting process

$$N_t^n = \sum_{s \leq t} 1_{\{|\Delta X_s| > \delta_n\}}, \quad N_t'^n = \sum_{s \leq t} 1_{\{|\Delta X_s| > \delta'_n\}}.$$

Their predictable compensators are  $\int_0^t \overline{F}_r(\delta_n) dr$  and  $\int_0^t \overline{F}_r(\delta'_n) dr$ , whereas both  $\overline{F}_r(\delta_n)$  and  $\overline{F}_r(\delta'_n)$  are smaller than  $K/\Delta_n^{\varpi\beta}$ . Hence (58) gives

$$(70) \quad \begin{cases} |\mathbb{P}_t(N_{t+s}^n - N_t^n = 1) - \theta_{t,s}| + \mathbb{P}_t(N_{t+s}^n - N_t^n \geq 2) \leq K \Delta_n^{2(1-\varpi\beta)}, \\ |\mathbb{P}_t(N_{t+s}'^n - N_t'^n = 1) - \theta'_{t,s}| \leq K \Delta_n^{2(1-\varpi\beta)}. \end{cases}$$

Since  $N^n - N'^n$  is non-decreasing, we have

$$\begin{aligned} \mathbb{P}_t(N_{t+s}^n - N_t^n = 1, N_{t+s}'^n - N_t'^n = 0) &= \mathbb{P}_t(N_{t+s}^n - N_t^n = 1) \\ &\quad - \mathbb{P}_t(N_{t+s}'^n - N_t'^n = 1) + \mathbb{P}_t(N_{t+s}^n - N_t^n \geq 2, N_{t+s}'^n - N_t'^n = 1). \end{aligned}$$

Then (70) yields

$$(71) \quad |\mathbb{P}_t(N_{t+s}^n - N_t^n = 1, N_{t+s}'^n - N_t'^n = 0) - (\theta_{t,s} - \theta'_{t,s})| \leq K \Delta_n^{2(1-\varpi\beta)}.$$

Moreover (45) clearly implies  $\theta_{t,s} - \theta'_{t,s} \leq K \Delta_n^{1-\varpi\beta} (\Delta_n^\eta + \Delta_n^{\varpi(\gamma \wedge (\beta - \beta'))}) \leq K \Delta_n^{1-\varpi\beta+\rho}$ . We then deduce from (71) that

$$(72) \quad \mathbb{P}_t(N_{t+s}^n - N_t^n = 1, N_{t+s}'^n - N_t'^n = 0) \leq K \Delta_n^{1-\varpi\beta+\rho}.$$

5) If  $N_{t+s}^n - N_t^n = N_{t+s}'^n - N_t'^n = 1$  and  $|X_{t+s} - X_t| \leq \delta_n$ , then necessarily  $|X(\delta_n)'_{t+s} - X(\delta_n)'_t| > \alpha \Delta_n^{\varpi+\eta}$ . Hence

$$\begin{aligned} \mathbb{P}_t(N_{t+s}^n - N_t^n = 1, |X_{t+s} - X_t| \leq \alpha \Delta_n^\varpi) &\leq \mathbb{P}_t(N_{t+s}^n - N_t^n = 1, N_{t+s}'^n - N_t'^n = 0) \\ &\quad + \mathbb{P}_t(N_{t+s}^n - N_t^n = 1, |X(\delta_n)'_{t+s} - X(\delta_n)'_t| > \alpha \Delta_n^{\varpi+\eta}). \end{aligned}$$

Then if we apply (59) with  $p$  large enough and  $\delta = \delta_n$  and  $\zeta = \Delta_n^\varpi$ , and (72), we deduce that as soon as  $\Delta_n$  is small enough,

$$(73) \quad \mathbb{P}_t(N_{t+s}^n - N_t^n = 1, |X_{t+s} - X_t| \leq \alpha \Delta_n^\varpi) \leq K \Delta_n^{1-\varpi\beta+\rho}.$$

Finally  $X_{t+s} - X_t = X(\delta_n)'_{t+s} - X(\delta_n)'_t$  on the set  $\{N_{t+s}^n - N_t^n = 0\}$ , so

$$\begin{aligned} \mathbb{P}_t(|X_{t+s} - X_t| > \alpha \Delta_n^\varpi) &= \mathbb{P}_t(N_{t+s}^n - N_t^n = 1) \\ &\quad - \mathbb{P}_t(N_{t+s}^n - N_t^n = 1, |X_{t+s} - X_t| \leq \alpha \Delta_n^\varpi) \\ &\quad + \mathbb{P}_t(N_{t+s}^n - N_t^n = 0, |X(\delta_n)'_{t+s} - X(\delta_n)'_t| > \alpha \Delta_n^\varpi) \\ &\quad + \mathbb{P}_{i-1}^n(N_{t+s}^n - N_t^n \geq 2, |X_{t+s} - X_t| > \alpha \Delta_n^\varpi). \end{aligned}$$

Then if we combine (69), (70) and (73), we get that for all  $\Delta_n$  small enough, we readily obtain (62). We also trivially deduce (64) from (45) and (62),

6) Finally, a close look at the previous argument shows that (62) also holds with  $\alpha \Delta_n^\varpi(1 + \Delta_n^\eta)$  and  $\theta_{t,s}i'$  in place of  $\alpha \Delta_n^\varpi$  and  $\theta_{t,s}$ . Therefore (63) follows, upon using the property  $\theta_{t,s} - \theta'_{t,s} \leq K \Delta_n^{1-\varpi\beta+\rho}$  proved above.  $\square$

**Lemma 7.** *Under the assumption and with the notation of Lemma 6, and if  $M$  is a bounded martingale, we have (with  $K$  depending also on  $M$ , recall  $s \leq \Delta_n$ ):*

$$(74) \quad \left| \mathbb{E}_t((M_{t+s} - M_t)1_{\{|X_{t+s} - X_t| > \alpha \Delta_n^\varpi\}}) \right| \leq K \Delta_n^{1-\varpi\beta+\rho} \\ + K \Delta_n^{1-(\varpi+\eta)\beta} \mathbb{E}_t(|M_{t+s} - M_t|) + K \Delta_n^{(1-(\varpi+\eta)\beta)/2} \sqrt{\mathbb{E}_t(|M_{t+s} - M_t|^2)}.$$

*Proof.* 1) There exist  $C^2$  functions  $f_n$  such that

$$(75) \quad \begin{cases} 1_{\{|x| > \alpha \Delta_n^\varpi(1+2\Delta_n^\eta/3)\}} \leq f_n(x) \leq 1_{\{|x| > \alpha \Delta_n^\varpi(1+\Delta_n^\eta/3)\}} \\ |f'_n(x)| \leq \frac{K}{\Delta_n^{\varpi+\eta}}, \quad |f''_n(x)| \leq \frac{K}{\Delta_n^{2(\varpi+\eta)}}. \end{cases}$$

With  $\widehat{X}' = X - B - X^c$ , and since  $M$  is bounded, we have

$$(76) \quad \begin{aligned} &\left| \mathbb{P}_t((M_{t+s} - M_t)1_{\{|X_{t+s} - X_t| > \alpha \Delta_n^\varpi\}}) - \mathbb{E}_t((M_{t+s} - M_t) f_n(\widehat{X}'_{t+s} - \widehat{X}'_t)) \right| \\ &\leq K \mathbb{P}_t(\alpha \Delta_n^\varpi < |X_{t+s} - X_t| \leq \alpha \Delta_n^\varpi(1 + \Delta_n^\eta)) \\ &\quad + K \mathbb{E}_t(|f_n(X_{t+s} - X_t) - f_n(\widehat{X}'_{t+s} - \widehat{X}'_t)|). \end{aligned}$$

Now we have

$$|f_n(x+y) - f_n(x)| \leq 1_{\{|y| > \alpha \Delta_n^{\varpi+\eta}/3\}} + \frac{K}{\Delta_n^{\varpi+\eta}} |y| 1_{\{\alpha \Delta_n^\varpi < |x+y| \leq \alpha \Delta_n^\varpi(1+\Delta_n^\eta)\}}.$$

If we apply this with  $x = \widehat{X}'_{t+s} - \widehat{X}'_t$  and  $y = (B + X^c)_{t+s} - (B + X^c)_t$ , plus (57-(b)) for  $p$  large enough and Bienaymé-Tchebycheff inequality and  $1 - 2\varpi - 2\eta > 0$ , plus



(63) and (57-(b)) again and Hölder's inequality, we obtain that the right side of (76) is smaller than  $K\Delta_n^{1-\varpi\beta+\rho}$ . Therefore it remains to prove that

$$(77) \quad \left| \mathbb{E}_t((M_{t+s} - M_t) f_n(\widehat{X}'_{t+s} - \widehat{X}'_t)) \right| \leq K\Delta_n^{1-\varpi\beta+\rho} \\ + K\Delta_n^{1-(\varpi+\eta)\beta} \mathbb{E}_t(|M_{t+s} - M_t|) + K\Delta_n^{(1-(\varpi+\eta)\beta)/2} \sqrt{\mathbb{E}_t(|M_{t+s} - M_t|^2)}.$$

2) According to Theorem III.4.20 of Jacod and Shiryaev (2003), we can "project" the martingale  $M$  onto the random measure  $\mu$ , which amounts to decomposing it as the sum of two martingales  $M = M' + M''$  with the following properties:

$$(78) \quad \left\{ \begin{array}{l} \bullet M' = \delta \star (\mu - \nu) \text{ for some predictable function } \delta \\ \bullet \sum_{s \leq t} \phi(s, \Delta X_s) \Delta M_s'' \text{ is a martingale as soon as} \\ \quad \phi \text{ is a predictable function satisfying } |\phi(\omega, s, x)| \leq K(1 \wedge |x| \\ \bullet \text{ the difference of the predictable quadratic variations} \\ \quad \langle M, M \rangle - \langle M', M' \rangle \text{ is non-decreasing.} \end{array} \right.$$

Note that we may even choose  $\delta$  bounded, because  $M$  is bounded (the fact that all the above processes are martingales and not only local martingales comes from the boundedness of  $M$  and of  $\int (x^2 \wedge 1) F_t(dx)$ ). We will also use the following consequence of the third property above:

$$(79) \quad \mathbb{E}_t \left( \int_t^{t+s} du \int F_{t+u}(x) \delta(t+u, x)^2 \right) = \mathbb{E}_t(\langle M', M' \rangle_{t+s} - \langle M', M' \rangle_t) \\ \leq \mathbb{E}_t((M_{t+s} - M_t)^2).$$

With  $t$  being fixed below, for simplicity we write  $Y_r = \widehat{X}'_{t+r} - \widehat{X}'_t$ . Since  $M$  is a bounded martingale and  $Y$  a semimartingale with vanishing continuous martingale part, and  $f_n(Y)$  is bounded, we deduce from Itô's formula and the properties (78) that the product  $(M_{t+r} - M_t) f_n(Y_r)$  is the sum of a martingale plus the process  $\int_0^r \gamma_u^n du$ , where

$$\gamma_u^n = \int F_{t+u}(dx) \left( (M_{t+u} - M_t) g_n(Y_u, x) + \delta(t+u, x) h_n(Y_u, x) \right), \\ \text{where } h_n(y, x) = f_n(y+x) - f_n(y), \quad g_n(y, x) = h_n(x, y) - f'_n(y)x 1_{\{|x| \leq 1\}}.$$

An easy computation allows to deduce of (75) that

$$|h_n(y, x)| \leq K \frac{|x| \wedge 1}{\Delta_n^{\varpi+\eta}} \\ |g_n(y, x)| \leq 1_{\{|x| > \alpha \Delta_n^{\varpi+\eta}\}} \\ + K 1_{\{\alpha \Delta_n^{\varpi} < |y| \leq \alpha \Delta_n^{\varpi} (1 + \Delta_n^\eta)\}} \left( \frac{x^2}{\Delta_n^{2\varpi+2\eta}} 1_{\{|x| \leq \alpha \Delta_n^{\varpi+\eta}\}} + \frac{|x| \wedge 1}{\Delta_n^{\varpi+\eta}} 1_{\{|x| > \alpha \Delta_n^{\varpi+\eta}\}} \right).$$

Now, we apply the first estimate of (45) with  $x = \alpha \Delta_n^{\varpi+\eta}$ , and the second and third ones with  $u = \alpha \Delta_n^{\varpi+\eta}$ , plus Cauchy-Schwarz inequality, to get for any  $\varepsilon > 0$  (recall that  $\delta$  is bounded):

$$|\gamma_u^n| \leq K \Delta_n^{-(\varpi+\eta)\beta} |M_{t+u} - M_t| + K \Delta_n^{-(\varpi+\eta)\beta/2} \left( \int F_{t+u}(x) \delta(t+u, x)^2 \right)^{1/2} \\ + K_\varepsilon |M_{t+u} - M_t| \Delta_n^{-(\beta+\varepsilon)(\varpi+\eta)} 1_{\{\alpha \Delta_n^{\varpi} < |Y_u| \leq \alpha \Delta_n^{\varpi} (1 + \Delta_n^\eta)\}}.$$

Since  $\eta < 1/2 - \varpi$  we have  $\beta(\varpi + \eta) < 1$  and thus  $(\beta + \varepsilon)(\varpi + \eta) = 1$  for a suitable  $\varepsilon > 0$ . Moreover  $\mathbb{E}_t(|Z_u|) \leq \mathbb{E}_t(|Z_s|)$  if  $u \leq s$  because  $Z$  is a martingale. Therefore, since  $M$  is bounded and  $s \leq \Delta_n$ , we get

$$\begin{aligned} & \left| \mathbb{E}_t((M_{t+s} - M_t) f_n(\widehat{X}'_{t+s} - \widehat{X}'_t)) \right| = \left| \mathbb{E}_t \left( \int_0^s \gamma_r^n dr \right) \right| \leq \int_0^s \mathbb{E}_t(|\gamma_r^n|) dr \\ & \leq K \Delta_n^{1-(\varpi+\eta)\beta} \mathbb{E}_t(|M_{t+s} - M_t|) \\ & \quad + K \Delta_n^{(1-(\varpi+\eta)\beta)/2} \left( \mathbb{E}_t \left( \int_t^{t+s} du \int F_{t+u}(x) \delta(t+u, x)^2 \right) \right)^{1/2} \\ & \quad + K \Delta_n^{-1} \int_0^{\Delta_n} \mathbb{P}_t(\alpha \Delta_n^\varpi < |Y_r| \leq \alpha \Delta_n^\varpi (1 + \Delta_n \eta)) dr. \end{aligned}$$

where we have used Cauchy-Schwarz inequality. By (63) for the process  $\widehat{X}'$  instead of  $X$  and by (79), we readily deduce (77).  $\square$

**8.3. Some auxiliary limit theorems.** Below, recall the process  $\bar{A}$  of (11). We still assume Assumptions 5 and 6, and also  $\varpi \in (0, \frac{1}{2})$  and  $\alpha > 0$ .

**Lemma 8.** *Let  $\rho' = \frac{1}{2} \wedge (\varpi(\beta - (\beta - \gamma) \vee \beta'))$ . Then for all  $t > 0$  the sequence*

$$(80) \quad \left( \Delta_n^{-\rho'} \left| \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \Delta_n^{\varpi\beta} \mathbb{E}_{i-1}^n \left( \int_{(i-1)\Delta_n}^{i\Delta_n} \bar{F}_t(\alpha \Delta_n^\varpi) dt \right) - \frac{\bar{A}_t}{\alpha^\beta} \right| \right)_{n \geq 1}$$

is tight.

*Proof.* Let  $\theta_i^n = \int_{(i-1)\Delta_n}^{i\Delta_n} \bar{F}_t(\alpha \Delta_n^\varpi) dt$  and  $\eta_i^n = \int_{(i-1)\Delta_n}^{i\Delta_n} A_t dt$ . We deduce from (45) that

$$\left| \Delta_n^{\varpi\beta} \theta_i^n - \frac{1}{\alpha^\beta} \eta_i^n \right| \leq K \Delta_n^{1+\varpi(\beta - (\beta - \gamma) \vee \beta')}.$$

Then obviously

$$\mathbb{E} \left( \Delta_n^{-\rho'} \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \mathbb{E}_{i-1}^n \left( \left| \Delta_n^{\varpi\beta} \theta_i^n - \frac{1}{\alpha^\beta} \eta_i^n \right| \right) \right) \leq Kt,$$

and since  $A_t$  is bounded we have  $\left| \bar{A}_t - \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \eta_i^n \right| \leq Kt\Delta_n$ , whereas  $\rho' < 1$ . It thus remains to prove that

$$(81) \quad \text{the sequence } \left( \Delta_n^{-\rho'} \left| \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (\eta_i^n - \mathbb{E}_{i-1}^n(\eta_i^n)) \right| \right)_{n \geq 1} \text{ is tight.}$$

Since  $\zeta_i^n = \Delta_n^{-\rho'} (\eta_i^n - \mathbb{E}_{i-1}^n(\eta_i^n))$  is a martingale increment, for (81) it is enough to check that  $a_n(t) = \mathbb{E} \left( \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (\zeta_i^n)^2 \right)$  is bounded. However, since  $A_t$  is bounded, we have  $|\zeta_i^n|^2 \leq K \Delta_n^{2-2\rho'}$ , so  $a_n(t) \leq Kt \Delta_n^{1-2\rho'} \leq K$  because  $\rho' \leq 1/2$ .  $\square$

**Lemma 9.** a) Let

$$(82) \quad \chi' = (\varpi\gamma) \wedge \frac{1 - \varpi\beta}{3} \wedge \frac{\varpi(\beta - \beta')}{1 + \beta'} \wedge \frac{1 - 2\varpi}{2}.$$

Then for all  $\varepsilon > 0$  and all  $t > 0$  the sequence of variables

$$(83) \quad \left( \Delta_n^{\varepsilon - \chi'} \left| \Delta_n^{\varpi\beta} \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \mathbb{P}_{i-1}^n (|\Delta_i^n X| > \alpha \Delta_n^{\varpi}) - \frac{\bar{A}_t}{\alpha^\beta} \right| \right)_{n \geq 1}$$

is tight, and in particular we have

$$(84) \quad \Delta_n^{\varpi\beta} \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \mathbb{P}_{i-1}^n (|\Delta_i^n X| > \alpha \Delta_n^{\varpi}) \xrightarrow{\mathbb{P}} \frac{\bar{A}_t}{\alpha^\beta}$$

b) If further  $\beta' < \frac{\beta}{2+\beta}$  and  $\gamma > \frac{\beta}{2}$  and  $\varpi < \frac{1}{2+\beta} \wedge \frac{2}{5\beta}$ , and if  $M$  is a bounded continuous martingale, we also have

$$(85) \quad \Delta_n^{-\varpi\beta/2} \left| \Delta_n^{\varpi\beta} \sum_{i=1}^{\lfloor s/\Delta_n \rfloor} \mathbb{P}_{i-1}^n (|\Delta_i^n X| > \alpha \Delta_n^{\varpi}) - \frac{\bar{A}_s}{\alpha^\beta} \right| \xrightarrow{\mathbb{P}} 0.$$

$$(86) \quad \Delta_n^{\varpi\beta/2} \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \left| \mathbb{E}_{i-1}^n \left( \Delta_i^n M 1_{\{|\Delta_i^n X| > \alpha \Delta_n^{\varpi}\}} \right) \right| \xrightarrow{\mathbb{P}} 0.$$

*Proof.* a) Let  $\eta \in (0, \frac{1}{2} - \varpi)$  and  $\rho$  given by (61) and  $\rho'$  like in Lemma 8. From (62) and (80) we deduce the tightness of the sequence (83), provided we substitute  $\chi' - \varepsilon$  with  $\rho' \wedge \rho$ . It is thus enough to show that we can choose  $\eta \leq \frac{1-2\varpi}{2}$  such that  $\rho' \wedge \rho = \chi' - \varepsilon$ , and this is achieved by taking  $\eta = \frac{1-\varpi\beta}{3} \wedge \frac{\varpi(\beta-\beta')}{1+\beta'} \wedge \frac{1-2\varpi}{2} - \varepsilon$ , as a simple computation shows.

b) In view of (a), (85) follows from the property  $\chi' > \varpi\beta/2$ , which is an easy consequence of the assumptions.

It remains to prove (86). By (74), the left side of this expression is smaller than

$$Kt\Delta_n^{\rho-\varpi\beta/2} + K\Delta_n^{1-\eta\beta-\varpi\beta/2} \sum_{i=1}^{\lfloor t/\Delta_n \rfloor + 1} \mathbb{E}_{i-1}^n (|\Delta_i^n M|) \\ + K\Delta_n^{(1-\eta\beta)/2} \sum_{i=1}^{\lfloor t/\Delta_n \rfloor + 1} \sqrt{\mathbb{E}_{i-1}^n (|\Delta_i^n M|)}.$$

By the Cauchy-Schwarz inequality this is smaller than

$$K(t + \sqrt{t}) \left( \Delta_n^{\rho-\varpi\beta/2} + \Delta_n^{1/2-\eta\beta-\varpi\beta/2} \left( \sum_{i=1}^{\lfloor t/\Delta_n \rfloor + 1} \mathbb{E}_{i-1}^n (|\Delta_i^n M|^2) \right) \right)^{1/2}.$$

A well known property of martingales yields

$$\mathbb{E} \left( \sum_{i=1}^{\lfloor t/\Delta_n \rfloor + 1} \mathbb{E}_{i-1}^n (|\Delta_i^n M|^2) \right) = \mathbb{E} ((M_{\Delta_n(\lfloor t/\Delta_n \rfloor + 1)} - M_0)^2),$$

which is bounded (in  $n$ ). Therefore we deduce that (86) holds, provided we have  $\rho > \varpi\beta/2$  and also  $1 - 2\eta\beta > \varpi\beta$ . The first condition has already been checked, and the second is satisfied because  $\eta \leq (1 - \varpi\beta)/3$ . This ends the proof.  $\square$

8.4. *Law of large numbers and central limit theorems.* Here again, we assume Assumptions 5 and 6, and we have  $\alpha > 0$  and  $\varpi \in (0, 1/2)$ .

**Proposition 1.** *If  $\chi$  is given by (18), for each  $t > 0$  and each  $\varepsilon > 0$  the sequence*

$$(87) \quad \left( \Delta_n^{\varepsilon - \chi} \left| \Delta_n^{\varpi\beta} U(\varpi, \alpha)_t^n - \frac{\bar{A}_t}{\alpha^\beta} \right| \right)_{n \geq 1}$$

*is tight, and in particular*

$$(88) \quad \Delta_n^{\varpi\beta} U(\varpi, \alpha)_t^n \xrightarrow{\mathbb{P}} \frac{\bar{A}_t}{\alpha^\beta}.$$

*Proof.* Set

$$(89) \quad \zeta_i^n = \Delta_n^{\varpi\beta/2} \left( 1_{\{|\Delta_i^n X| > \alpha \Delta_n^\varpi\}} - \mathbb{P}_{i-1}^n(|\Delta_i^n X| > \alpha \Delta_n^\varpi) \right).$$

By virtue of Lemma 9, and since  $\chi = \chi' \wedge (\varpi\beta/2)$ , it suffices to prove that the sequence  $\sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \zeta_i^n$  is tight. Since the  $\zeta_i^n$ 's are martingale increments, it is enough to show that the sequence  $a_n(t) = \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \mathbb{E}((\zeta_i^n)^2)$  is bounded. But (64) yields  $\mathbb{E}((\zeta_i^n)^2) \leq K\Delta_n$ , which in turn yields  $a_n(t) \leq Kt$ .  $\square$

**Proposition 2.** *Let  $\alpha' > \alpha$ . If we have  $\beta' < \frac{\beta}{2+\beta}$  and  $\gamma > \frac{\beta}{2}$  and  $\varpi < \frac{1}{2+\beta} \wedge \frac{2}{5\beta}$ , the pair of processes*

$$(90) \quad \Delta_n^{-\varpi\beta/2} \left( \Delta_n^{\varpi\beta} U(\varpi, \alpha)_t^n - \frac{\bar{A}_t}{\alpha^\beta}, \Delta_n^{\varpi\beta} U(\varpi, \alpha')_t^n - \frac{\bar{A}_t}{\alpha'^\beta} \right)$$

*converges stably in law to a continuous Gaussian martingale  $(\bar{W}, \bar{W}')$  independent of  $\mathcal{F}$ , with*

$$(91) \quad \mathbb{E}(\bar{W}_t^2) = \frac{\bar{A}_t}{\alpha^\beta} \quad \mathbb{E}(\bar{W}'_t^2) = \frac{\bar{A}_t}{\alpha'^\beta} \quad \mathbb{E}(\bar{W}_t \bar{W}'_t) = \frac{\bar{A}_t}{\alpha'^\beta}.$$

*Proof.* Define  $\zeta_i^n$  by (89), and associate  $\zeta_i^{n'}$  with  $\alpha'$  in the same way. The variables  $\zeta_i^n$  and  $\zeta_i^{n'}$  are martingale increments and smaller than  $K\Delta_n^{\varpi\beta/2}$ , so in view of (85) it is enough, by using a criterion for stable convergence of triangular arrays found in Jacod and Shiryaev (2003) (see Theorem IX.7.28) to prove the following

$$(92) \quad \begin{cases} \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \mathbb{E}_{i-1}^n((\zeta_i^n)^2) \xrightarrow{\mathbb{P}} \frac{\bar{A}_t}{\alpha^\beta}, \\ \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \mathbb{E}_{i-1}^{n'}((\zeta_i^{n'})^2) \xrightarrow{\mathbb{P}} \frac{\bar{A}_t}{\alpha'^\beta}, \\ \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \mathbb{E}_{i-1}^n(\zeta_i^n \zeta_i^{n'}) \xrightarrow{\mathbb{P}} \frac{\bar{A}_t}{\alpha'^\beta}. \end{cases}$$

$$(93) \quad \begin{cases} \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \mathbb{E}_{i-1}^n(\zeta_i^n \Delta_i^n M) \xrightarrow{\mathbb{P}} 0, \\ \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} \mathbb{E}_{i-1}^{n'}(\zeta_i^{n'} \Delta_i^n M) \xrightarrow{\mathbb{P}} 0, \end{cases}$$

where  $M$  is any bounded martingale.

Recalling  $\alpha < \alpha'$  and also  $\mathbb{P}_{i-1}^n(|\Delta_i^n X| > \alpha \Delta_n^\varpi) \leq K\Delta_n^{1-\varpi\beta}$  by (64), we deduce (92) from (84). As for (93), and since  $\mathbb{E}_{i-1}^n(\Delta_i^n M) = 0$ , it readily follows from (86).  $\square$

8.5. *Proof of the main theorems.*

*Proof of Theorems 1 and 2.* As said before, we can assume Assumptions 5 and 6. We will write  $\eta = \chi - \varepsilon$  if we want to prove the first theorem, in which case we choose  $\varepsilon$  small enough to have  $\eta > 0$ , and  $\eta = \varpi\beta/2$  if we want to prove the second one. Then we set

$$V_n = \Delta_n^{-\eta} \left( \Delta_n^{\varpi\beta} U(\varpi, \alpha)_t^n - \bar{A}_t / \alpha^\beta \right), \quad V'_n = \Delta_n^{-\eta} \left( \Delta_n^{\varpi\beta} U(\varpi, \alpha')_t^n - \bar{A}_t / \alpha'^\beta \right),$$

so that on the set  $\{\bar{A}_t > 0\}$  we have

$$(94) \quad \Delta_n^{-\eta} \left( \hat{\beta}_n(t, \varpi, \alpha, \alpha') - \beta \right) = \frac{\Delta_n^{-\eta}}{\log(\alpha'/\alpha)} \log \frac{1 + \alpha^\beta \Delta_n^\eta V_n / \bar{A}_t}{1 + \alpha'^\beta \Delta_n^\eta V'_n / \bar{A}_t}.$$

In all cases, the two sequences  $(V_n)$  and  $(V'_n)$  are tight, so the right side is equivalent (in probability) to the variable  $(\alpha^\beta V_n - \alpha'^\beta V'_n) / (\bar{A}_t \log(\alpha'/\alpha))$  on the set  $\{\bar{A}_t > 0\}$ . At this stage, Theorems 1 and 2 readily follows from Lemma 9-(a) and Proposition 2, respectively.  $\square$

*Proof of Theorem 3.* This is a trivial consequence of the properties of the stable convergence en law, and of Theorem 2.  $\square$

*Proof of Theorems 4 and 5.* In the situation of these two theorems we have  $\bar{A}_t = At > 0$ , so we can apply Theorems 1 and 3 with  $\{\bar{A}_t > 0\} = \Omega$ . Indeed we have Assumptions 1 and 2 with  $\beta' = 0$  and  $\gamma$  arbitrary large. So the only differences are on the conditions on  $\varpi$ , namely  $\varpi < 1/\beta$  for the consistency and  $\varpi < 2/(3\beta)$  for the CLT in the first theorem, and  $\varpi < 1/(2 + \beta)$  for the CLT in the second one, instead of  $\varpi < 1/2$  in Theorem 1 and  $\varpi < \frac{1}{2+\beta} \wedge \frac{2}{5\beta}$  in Theorem 3.

In fact the improvement lies in Lemma 6. When  $X = Y$  we have the trivial estimate, coming from (22) and from the scaling property of  $Y$ :

$$(95) \quad \mathbb{P}_t(|X_{t+s} - X_t| > x) = \mathbb{P}(|X_s| > x) = \frac{As}{x^\beta} + O(s^2/x^{2\beta}),$$

regardless of the value of  $x > 0$ . So we have all claims of Lemma 6 with any  $\eta \in (0, 1 - \varpi\beta]$  and with  $\rho = 1 - \varpi\beta$ , and thus (a) of Lemma 9 holds for some  $\chi > 0$  as soon as  $\varpi < 1/\beta$ , whereas (b) of that lemma holds if further  $\rho > \varpi\beta/2$ , that is when  $\varpi < 2/(3\beta)$ .

When  $X_t = \sigma W_t + Y_t$  we obtain for any  $\eta \in (0, 1/2 - \varpi)$  and  $s \leq \Delta_n$

$$\mathbb{P}_t(|X_{t+s} - X_t| > \alpha \Delta_n^\varpi) = \mathbb{P}(|X_s| > \alpha \Delta_n^\varpi) = \frac{A \Delta_n^{1-\varpi\beta}}{\alpha^\beta} + O(\Delta_n^{1-1\varpi\beta+\eta}),$$

(by applying the estimate (95) for  $Y$  and the fact that  $\mathbb{E}(|X_s^c|^p) = K_p s^{p/2}$  and Bienaymé-Tchebycheff inequality). Then we have all claims of Lemma 6 with any  $\eta \in (0, (\frac{1}{2} - \varpi) \wedge (1 - \varpi\beta))$  and with  $\rho = \eta$ , and thus (b) of Lemma 9 holds as soon as we can find  $\eta$  as above, with  $\rho = \eta > \frac{\varpi\beta}{2}$ . This is possible, provided  $\varpi < \frac{1}{2+\beta} \wedge \frac{2}{3\beta}$ . The rest of the proofs of Theorems 1 and 3 goes through, and we are finished.  $\square$

## REFERENCES

- AÏT-SAHALIA, Y. and JACOD, J. (2006). Testing for jumps in a discretely observed process. *Annals of Statistics, forthcoming* .
- AÏT-SAHALIA, Y. and JACOD, J. (2008). Fisher's information for discretely sampled Lévy processes. *Econometrica, forthcoming* .
- BLUMENTHAL, R. and GETTOOR, R. (1961). Sample functions of stochastic processes with stationary independent increments. *Journal of Mathematics and Mechanics* **10** 493–516.
- CONT, R. and MANCINI, C. (2007). Nonparametric tests for probing the nature of asset price processes. Tech. rep., Università di Firenze.
- DUMOUCHEL, W. H. (1983). Estimating the stable index  $\alpha$  in order to measure tail thickness: A critique. *The Annals of Statistics* **11** 1019–1031.
- HILL, B. M. (1975). A simple general approach to inference about the tail of a distribution. *The Annals of Statistics* **3** 1163–1174.
- JACOD, J. and SHIRYAEV, A. N. (2003). *Limit Theorems for Stochastic Processes*. 2nd ed. Springer-Verlag, New York.
- WOERNER, J. H. (2006). Analyzing the fine structure of continuous-time stochastic processes. Tech. rep., University of Göttingen.
- ZOLOTAREV, V. M. (1986). *One-dimensional Stable Distributions*, vol. 65 of *Translations of Mathematical Monographs*. American Mathematical Society.

DEPARTMENT OF ECONOMICS  
PRINCETON UNIVERSITY AND NBER  
PRINCETON, NJ 08544-1021  
E-MAIL: YACINE@PRINCETON.EDU

INSTITUT DE MATHÉMATIQUES DE JUSSIEU  
CNRS UMR 7586  
UNIVERSITÉ P. ET M. CURIE (PARIS-6)  
75252 PARIS CÉDEX 05  
E-MAIL: JJ@CCR.JUSSIEU.FR



本文献由“学霸图书馆-文献云下载”收集自网络，仅供学习交流使用。

学霸图书馆（www.xuebalib.com）是一个“整合众多图书馆数据库资源，提供一站式文献检索和下载服务”的24小时在线不限IP图书馆。

图书馆致力于便利、促进学习与科研，提供最强文献下载服务。

#### 图书馆导航：

[图书馆首页](#)   [文献云下载](#)   [图书馆入口](#)   [外文数据库大全](#)   [疑难文献辅助工具](#)