TESTING FOR JUMPS IN A DISCRETELY OBSERVED PROCESS

By Yacine Aït-Sahalia[1] and Jean Jacod

Princeton University and UPMC (Université Paris-6)

We propose a new test to determine whether jumps are present in asset returns or other discretelly sampled processses. As the sampling interval tends to 0, our test statistic converges to 1 if there are jumps, and to another deterministic and known value (such as 2) if there are no jumps. The test is valid for all Itô semimartingales, depends neither on the law of the process nor on the coefficients of the equation which it solves, does not require a preliminary estimation of these coefficients, and when there are jumps the test is applicable whether jumps have finite or infinite activity and for an arbitrary Blumenthal-Getoor index. We finally implement the test on simulations and asset returns data.

1. Introduction. The problem of deciding whether the continuous time process which models an economic or financial time series should have continuous paths or exhibit jumps is becoming an increasingly important issue, in view of the high frequency observations that are now widely available. In the case where a large jump occurs, a simple glance at the dataset might be sufficient to decide this issue. But such large jumps are usually infrequent, may not belong to the model itself, can be considered as breakdowns in the homogeneity of the model, or may be dealt with separately using other methods such as risk management.

On the other hand, a visual inspection of most such time series in practice does not provide a clear evidence for either the presence or the absence of small or medium sized jumps. Since small frequent jumps should definitely be incorporated into the model, and since models with and without jumps do have quite different mathematical properties and financial consequences (for option hedging, portfolio optimization, etc.), it is important to have statistical methods that can shed some light on the issue.

Determining whether a process has jumps has been considered by a number of authors. Aït-Sahalia (2002) relies on restrictions on the transition function of the process that are compatible with continuity of the process, or lack thereof, to derive a test for the presence of jumps at any observable frequency. Using high frequency data, Carr and Wu (2003) exploit the differential behavior of short dated options

[1]Supported in part by NSF Grants SBR-0350772 and DMS-0532370.

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

Key words and phrases. Jumps, test, discrete sampling, high frequency.

to test for the presence of jumps. Multipower variations can separate the continuous part of the quadratic variation, see Barndorff-Nielsen and Shephard (2004) or Barndorff-Nielsen and Shephard (2006). Andersen et al. (2003) and Huang and Tauchen (2006) study financial datasets using multipower variations, in order to assess the proportion of quadratic variation attributable to jumps. Jiang and Oomen (2005) construct a test motivated by the hedging error of a variance swap replication strategy. Other methods have been introduced as well, see e.g. Lee and Mykland (2005), and the literature about evaluating the volatility when there are jumps (see e.g., Aït-Sahalia (2004), Aït-Sahalia and Jacod (2006), Mancini (2001), Mancini (2004) and Woerner (2006b)) can also be viewed as an indirect way of checking for jumps. Woerner (2006a) proposes estimators of the Blumenthal-Getoor index or the Hurst exponent of a stochastic process based on high frequency data, and those can be used to detect jumps and their intensity.

These are two closely related but different issues: one is to decide whether jumps are present or not, another one is to determine the impact of jumps on the overall variability of the observed process. So far, most of the literature has concentrated on the second issue, and it usually assumes a special structure on the process, especially regarding its jump part if one is present, such as being a compound Poisson, or sometimes a Lévy process. We plan to have a systematic look at the question of measuring the impact of jumps in a future paper, using a variety of distances and doing so under weak assumptions, and also taking into account the usually non-negligible market microstructure noise.

In this paper, we concentrate on the first problem, namely whether there are jumps or not, while pretending that the underlying process is perfectly observed at n discrete times, and n is large. We introduce a direct and very simple test which gives a solution to this problem, irrespectively of the precise structure of the process within the very large class of Itô semimartingales.

More specifically, a process $X = (X_t)$ on a given time interval $[0, T]$ is observed at times $i\Delta_n$ for $\Delta_n = T/n$. We propose an easy-to-compute family of test statistics, say S_n, which converge as $\Delta_n \to 0$ to 1 if there are jumps, and to another deterministic and known value (such as 2) if there are no jumps. This holds as soon as the process X is an Itô semimartingale, and it depends neither on the law of the process nor on the coefficients of the equation which it solves, and it does not require any preliminary estimation of these coefficients. We provide a central limit theorem for S_n, under both alternatives (jumps and no jumps): again this does not require any a priori knowledge of the coefficients of the model, and when there are jumps it is applicable whether the jumps have finite or infinite activity, and for an arbitrary Blumenthal-Getoor index. Hence we can construct tests with a given level of significance asymptotically and which are fully non-parametric.

The paper is organized as follows. Sections 2 and 3 describe our setup and the statistical problem. We provide Central Limit Theorems for our proposed statistics in Section 4, and use them to construct the actual tests in Section 5. We present

the results of Monte Carlo simulations in Section 6, and in Section 7 examine the empirical distribution of the test statistic over all 2005 transactions of the Dow Jones stocks. Proofs are in Section 8.

2. Setting and assumptions. The underlying process X which we observe at discrete times is a 1-dimensional process which we specify below. Observe that taking a 1-dimensional process is not a restriction in our context since, if it were multi-dimensional, a jump would necessarily be a jump for at least one of its components, so any test for jumps can be performed separately on each of its components.

As already mentioned, we do not want to make any specific model assumption on X, such as assuming some parametric family of models. We do need, however, a mild structural assumption which is satisfied in all continuous time models used in finance, at least as long as one wants to rule out arbitrage opportunities. In any case, in the absence of some kind of assumption, anything can happen from a continuous process which is a linear interpolation between the observations, to a pure jump process which is constant between successive observation times.

Our structural assumption is that X is an Itô semimartingale on some filtered space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, which means it can be written as

$$\begin{aligned} X_t \;=\; & X_0 + \int_0^t b_s ds + \int_0^t \sigma_s \, dW_s + \int_0^t \int_E \kappa \circ \delta(s,x)(\mu - \nu)(ds,dx) \\ & + \int_0^t \int_E \kappa' \circ \delta(s,x)\mu(ds,dx). \end{aligned} \tag{1}$$

where W and μ are a Wiener process and a Poisson random measure on $[0,1] \times E$ with (E,E) an auxiliary measurable space on the space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ and the predictable compensator (or intensity measure) of μ is $\nu(ds,dx) = ds \otimes \lambda(dx)$ for some given finite or σ-finite measure on (E,E). Moreover κ is a continuous function with compact support and $\kappa(x) = x$ on a neighborhood of 0, and $\kappa'(x) = x - \kappa(x)$.

Above, the function κ is arbitrary, a change of this function amounts to a change in the drift coefficient b_t. Also, it is always possible to take (E, E, λ) to be $\mathbb{R}$ equipped with Lebesgue measure.

Of course the coefficients $b_t(\omega)$, $\sigma_t(\omega)$ and $\delta(\omega, t, x)$ should be such that the various integrals in (1) make sense (see for example Jacod and Shiryaev (2003) for a precise definition of the last two integrals in (1)), and in particular b_t and σ_t are optional processes and δ is a predictable function. However we need a bit more than just the minimal integrability assumptions, and our additional requirements are expressed as follows:

Assumption 1. *(a) All paths $t \mapsto \sigma_t(\omega)$ are right-continuous with left limits.*

(b) All paths $t \mapsto b_t(\omega)$ are locally bounded.

(c) All paths $t \mapsto \sup_{x\in E} \frac{|\delta(\omega,t,x)|}{\gamma(x)}$ are locally bounded, where γ is a (non-random) nonnegative function which satisfies $\int_E (\gamma(x)^2 \wedge 1)\, \lambda(dx) < \infty$.

(d) We have $\int_0^t |\sigma_s|\, ds > 0$ a.s. for all $t > 0$.

Apart from (a), these conditions are simply some local boundedness conditions on the coefficients, which are satisfied in all concrete models, plus the non-degeneracy condition (d) which says that almost surely the continuous martingale part of X is not identically 0 on any interval $[0,t]$ with $t>0$. Most results of this paper are true without (d), but not all, and in any case this condition is satisfied in all applications we have in mind. For example if we consider a d-dimensional equation

$$dY_t = f(Y_{t-})dZ_t \tag{2}$$

where Z is a multi-dimensional Lévy process with Gaussian components, and f is a continuous, locally Lipschitz function with at most linear growth, then any of the components of Y satisfies (a,b,c) of Assumption 1. The same holds for more general equations driven by a Wiener process and a Poisson random measure.

For some of the results we will derive, we will need more than Assumption 1. Below, W and μ are as in (1), and W' is another Wiener process independent of (W,μ):

Assumption 2. *We have Assumption 1, and the process σ has the form*

$$\begin{aligned}\sigma_t \;=\; \sigma_0 &+ \int_0^t \widetilde{b}_s\,ds + \int_0^t \widetilde{\sigma}_s\,dW_s + \int_0^t \widetilde{\sigma}'_s\,dW'_s + \int_0^t\int_E \kappa\circ\widetilde{\delta}(s,x)(\mu-\nu)(ds,dx)\\ &+ \int_0^t\int_E \kappa'\circ\widetilde{\delta}(s,x)\mu(ds,dx),\end{aligned} \tag{3}$$

and further:

(a) All paths $t\mapsto\widetilde{b}_t(\omega)$ are locally bounded.

(b) All paths $t\mapsto b_t(\omega)$, $t\mapsto\widetilde{\sigma}_t(\omega)$, $t\mapsto\widetilde{\sigma}'_t(\omega)$ are right-continuous with left limits.

(c) All paths $t\mapsto\delta(\omega,t,x)$ and $t\mapsto\widetilde{\delta}(\omega,t,x)$ are left-continuous with right limits, and further $\widetilde{\delta}$ satisfies the same condition than δ in (Assumption 1-(c)).

This hypothesis is satisfied for instance if X is a component of a process Y like in (2) as soon as f is C^2, and it also accommodates virtually all models for stochastic volatility, including those with jumps, and allows for correlation between the volatility and the asset price processes.

3. The statistical problem.

3.1. *Preliminaries.* Our process X is discretely observed over a given time interval $[0,T]$, and it is convenient in practice to allow T to vary. So we suppose that X can be observed at times $i\Delta_n$ for all $i=0,1,\cdots$, and we take into account only those observation times $i\Delta_n$ smaller than or equal to T, whether T is a multiple of Δ_n or not. Moreover the testing procedures given below are asymptotic, in the sense that we can specify the level or the power function asymptotically as $n\to\infty$ and $\Delta_n\to 0$.

Recall that we want to decide whether there are jumps or not, for the process (1); equivalently, we want to decide whether the coefficient δ is identically 0 or not. A few remarks can be stated right away, which describe some of the difficulties or peculiarities of this statistical problem:

1. It is a non-parametric problem: we do not wish to specify the coefficients b, σ, δ.

2. It is an asymptotic problem, which only makes sense for high frequency data.

3. When "n is infinite", that is in the ideal although unrealistic situation of a complete observation of the path of X over $[0,T]$, we can of course tell whether our particular path has jumps of not. However, when the measure λ is finite there is a positive probability that the path $X(\omega)$ has no jump on $[0,T]$, although the model itself may allow for jumps.

4. In the realistic case $n < \infty$ we cannot do better than in the "completely observed case". That is, we can hopefully infer something about the jumps which actually occurred for our particular observed path, but nothing about those which belong to the model but did not occur on the observed path.

5. Rather than "testing for jumps", there are cases in practice where one wants to estimate in some sense the part of the variability of the process which is due to the jumps, compared to the part due to the continuous component. This is what most authors have studied so far, in some special cases at least, and we will take a systematic look at this question in our further paper in preparation on the topic, using essentially the same tools as in the present paper. However, under the null hypothesis that jumps are present, it is not clear how one should go about specifying the proportion of quadratic variation attributable to jumps without already assuming not only the type but also the "quantity" of jumps. The test we propose below does not have this problem.

6. Coming back to testing, an important property of our test statistics is that they should be scale-invariant (invariant if X is multiplied by an arbitrary constant). It would also be desirable for the limiting behavior of the statistic to be independent of the dynamics of the process. We will see that our test has all these features.

In view of comments 3 and 4 above, the problem which we really try to solve in this paper is to decide, on the basis of the observations $X_{i\Delta_n}$ which belong to the time interval $[0,T]$, in which of the following two complementary sets the path which we have discretely observed falls:

$$\begin{cases} \Omega_T^j = \{\omega : \ s \mapsto X_s(\omega) \text{ is discontinuous on } \ [0,T]\} \\ \Omega_T^c = \{\omega : \ s \mapsto X_s(\omega) \text{ is continuous on } \ [0,T]\}. \end{cases} \tag{4}$$

If we decide on Ω_T^j then we also implicitly decide that the model has jumps, whereas if we decide on Ω_T^c it does not mean that the model is continuous, even on the interval $[0,T]$ (of course, in both cases we can say nothing about what happens after T!).

3.2. *Measuring the variability of X.* Let us now introduce a number of processes which all measure some kind of variability of X, or perhaps its continuous and jump components separately, and depend on the whole (unobserved) path of X:

$$A(p)_t = \int_0^t |\sigma_s|^p \, ds, \qquad B(p)_t = \sum_{s \le t} |\Delta X_s|^p, \tag{5}$$

where p is a positive number and

$$\Delta X_s = X_s - X_{s-} \tag{6}$$

denotes the jumps of the X process. Of course, $\Delta X_s = 0$ for all s when X is continuous, and for all s outside a (random) countable set in all cases.

Note that $A(p)$ is finite-valued for all $p > 0$ (under Assumption 1), whereas $B(p)$ is finite-valued if $p \ge 2$ but often not when $p < 2$. Also, recall that $[X,X] = A(2) + B(2)$.

We have $\Omega_T^j = \{B(p)_T > 0\}$ for any $p > 0$, so in a sense our problem "simply" amounts to determine whether $B(p)_T > 0$ for our particular observed path, and with any pre-specified p. Moreover a reasonable measure of the relative variability, or variance, due to the jumps is $B(2)_T/[X,X]_T$, and this is the measure used for example by Huang and Tauchen (2006); other measures of this variability could be $B(p)_T/A(p)_T$ or $B(p)_T/[X,X]_T^{p/2}$ for other values of p (the power $p/2$ in the denominator is to ensure the scale-invariance).

In any event, everything boils down to estimating, on the basis of the actual observations, the quantity $B(p)_T$ in (5), and the difficulty of this estimation depends on the value of p. Let

$$\Delta_i^n X = X_{i\Delta_n} - X_{(i-1)\Delta_n}, \tag{7}$$

denote the observed discrete increments of X (all of them, not just those due to jumps, unlike (6)) and define for $p > 0$ the estimator

$$\widehat{B}(p, \Delta_n)_t \;:=\; \sum_{i=1}^{[t/\Delta_n]} |\Delta_i^n X|^p. \tag{8}$$

For $r \in (0, \infty)$, let

$$m_r = \mathbb{E}(|U|^r) = \pi^{-1/2} 2^{r/2} \, \Gamma\left(\frac{r+1}{2}\right) \tag{9}$$

denote the rth absolute moment of a variable $U \sim N(0,1)$. We have the following convergences in probability, locally uniform in t:

$$
\begin{cases}
p > 2, \quad X \text{ a semimartingale} & \Rightarrow \quad \widehat{B}(p,\Delta_n)_t \xrightarrow{\mathbb{P}} B(p)_t \\
p = 2, \quad X \text{ a semimartingale} & \Rightarrow \quad \widehat{B}(p,\Delta_n)_t \xrightarrow{\mathbb{P}} [X,X]_t \\
p < 2, \quad \text{Assumption 1 holds} & \Rightarrow \quad \frac{\Delta_n^{1-p/2}}{m_p}\, \widehat{B}(p,\Delta_n)_t \xrightarrow{\mathbb{P}} A(p)_t \\
\text{Assumption 1 and } X \text{ is continuous} & \Rightarrow \quad \frac{\Delta_n^{1-p/2}}{m_p}\, \widehat{B}(p,\Delta_n)_t \xrightarrow{\mathbb{P}} A(p)_t.
\end{cases} \tag{10}
$$

These properties are known: for $p = 2$ this is the classical convergence of the realized quadratic variation, for $p > 2$ this is due to Lepingle (1976), and for the other cases one may see Jacod (2006) for example.

The intuition for the behavior of $\widehat{B}(p,\Delta_n)_t$ is as follows. Suppose that X can jump. Among the increments of X, there are those that contain a large jump and those that do not. While the increments containing large jumps are much less frequent than those that contain only a Brownian contribution (in the case where λ is finite, otherwise a Brownian and many small jumps), they are so much bigger than the rest that when jumps occur, their contribution to $B(p)$ for $p > 2$ overwhelms everything else. This is because high powers ($p > 2$) magnify the large increments at the expense of the small ones. Then the sum behaves like the sum coming from the jumps only: this is the first result in (10). When p is small ($p < 2$), on the other hand, the magnification of the large increments by the power is not strong enough to overcome the fact that there are many more small increments than large ones. Then the behavior of the sum is driven by the summation of all these small increments: this is the third result in (10). When $p = 2$, we are in the situation where these two effects (magnification of the relatively few large increments vs. summation of many small increments) are of the same magnitude: this is the second result in (10). When X is continuous, we are only summing small increments and we get for all values of p the same behavior as the one in the third result where the summation of the small increments dominates the sum: this is the fourth result in (10).

3.3. *The test statistics.* Based on this intuition, upon examining (10), we see that when $p > 2$ the limit of $\widehat{B}(p,\Delta_n)$ does not depend on the sequence Δ_n going to 0, and it is strictly positive if X has jumps between 0 and t. On the other hand when X is continuous on $[0,t]$ then $\widehat{B}(p,\Delta_n)$ converges again to a limit not depending on Δ_n, but only after a normalization which does depend on Δ_n.

These considerations lead us to compare $\widehat{B}(p,\Delta_n)$ on two different Δ_n-scales. Specifically, we choose an integer k and compare $\widehat{B}(p,\Delta_n)$ with $\widehat{B}(p,k\Delta_n)$, the latter obtained by considering only the increments of X between successive multiples of $k\Delta_n$. Then we set

$$
\widehat{S}(p,k,\Delta_n)_t \;=\; \frac{\widehat{B}(p,k\Delta_n)_t}{\widehat{B}(p,\Delta_n)_t} \tag{11}
$$

as our (family of) test statistics.

In view of the first and fourth limits in (10), we readily get:

Theorem 1. *Under Assumption 1, let $t > 0$, $p > 2$ and $k \geq 2$. Then the variables $\widehat{S}(p,k,\Delta_n)_t$ converge in probability to the variable $S(p,k)_t$ defined by*

$$S(p,k)_t = \begin{cases} 1 & \text{on the set } \Omega_t^j \\ k^{p/2-1} & \text{on the set } \Omega_t^c. \end{cases} \tag{12}$$

Therefore our test statistics will converge to 1 in the presence of jumps and, with the selection of $p = 4$ and $k = 2$, to 2 in the absence of jumps. The following corollary is immediate:

Corollary 1. *Under Assumption 1, let $t > 0$, $p > 2$ and $k \geq 2$. The decision rule defined by:*

$$\Im(n) = \Im(n,t,p,k) = \begin{cases} X \text{ is discontinuous on } [0,t] & \text{if } \widehat{S}(p,k,\Delta_n)_t < a \\ X \text{ is continuous on } [0,t] & \text{if } \widehat{S}(p,k,\Delta_n)_t \geq a \end{cases}$$

is consistent, in the sense that the probability that it gives the wrong answer tends to 0 as $\Delta_n \to 0$, for any choice of a in the interval $(1, k^{p/2} - 1)$.

4. Central limit theorems.

4.1. *CLT for power variations.* The previous corollary provides the first step towards constructing a test for the presence or absence of jumps, but it is hardly enough. To construct tests, we need to derive the rates of convergence and the asymptotic variances when X jumps and when it is continuous.

We start with the following general theorem, to be proved in Section 8 (as above, k is an integer larger than 1). The asymptotic variances in the theorem involve the process $A(p)_t$ already defined in 5 as well as the more complex process

$$D(p)_t = \sum_{s \leq t} |\Delta X_s|^p \, (\sigma_{s-}^2 + \sigma_s^2) \tag{13}$$

for $p > 0$. Like $B(p)_t$, and $D(p)_t$ is finite-valued if $p \geq 2$ but often not when $p < 2$.

Theorem 2. *a) Under Assumption 1, let $p > 3$. The pair of variables*

$$\Delta_n^{-1/2} \quad \Big(\widehat{B}(p,\Delta_n)_t - B(p)_t, \quad \widehat{B}(p,k\Delta_n)_t - B(p)_t \Big)$$

converges stably in law to a two-dimensional variable of the form $(Z(p)_t, Z(p)_t + Z'(p,k)_t)$, where both $Z(p)_t$ and $Z'(p,k)_t$ are defined on an extension $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, (\widetilde{\mathcal{F}}_t)_{t\geq 0}, \widetilde{\mathbb{P}})$ of the original filtered space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ and conditionally on $\mathcal{F}$ are centered, with $Z'(p)_t$ having the following conditional variance:

$$\widetilde{\mathbb{E}}(Z'(p,k)_t^2 \mid \mathcal{F}) \;=\; \frac{k-1}{2} \; p^2 \; D(2p-2)_t. \tag{14}$$

Moreover if the processes σ and X have no common jumps, the variables $Z'(p)_t$ are $\mathcal{F}$-conditionally Gaussian.

b) Under Assumption 2, assume in addition that X is continuous, and let $p \geq 2$. The pair of variables

$$\Delta_n^{-1/2} \quad \Big(\Delta_n^{1-p/2}\widehat{B}(p,\Delta_n)_t - m_p A(p)_t, \quad \Delta_n^{1-p/2}\widehat{B}(p,k\Delta_n)_t - k^{p/2-1}m_p A(p)_t\Big)$$

converges stably in law to a two-dimensional variable $(Y(p)_t, Y'(p,k)_t)$ defined on an extension $(\widetilde{\Omega},\widetilde{\mathcal{F}},(\widetilde{\mathcal{F}}_t)_{t\geq 0},\widetilde{\mathbb{P}})$ of the original filtered space $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\geq 0},\mathbb{P})$ and which conditionally on $\mathcal{F}$ is a centered Gaussian variable with variance-covariance given by

$$\begin{cases} \widetilde{\mathbb{E}}(Y(p)_t^2 \mid \mathcal{F}) = (m_{2p} - m_p^2)A(2p)_t \\ \widetilde{\mathbb{E}}(Y'(p,k)_t^2 \mid \mathcal{F}) = k^{p-1}(m_{2p} - m_p^2)A(2p)_t \\ \widetilde{\mathbb{E}}(Y(p)_t Y'(p,k)_t \mid \mathcal{F}) = (m_{k,p} - k^{p/2}m_p^2)A(2p)_t \end{cases} \tag{15}$$

and where

$$m_{k,p} = \mathbb{E}(|U|^p|U + \sqrt{k-1}\ V|^p) \tag{16}$$

for U, V independent $N(0,1)$ variables.

The "stable convergence in law" mentioned above is a mode of convergence introduced by Rényi (1963), which is slightly stronger than the mere convergence in law, and its important feature for us is that if V_n is any sequence of variable converging in probability to a limit V on the space $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\geq 0},\mathbb{P})$, whereas the variables V_n' converge stably in law to V', then the pair (V_n, V_n') converges stably in law again to the pair (V,V').

The reader will observe the restriction $p > 3$ in (a) above: the CLT simply does not hold if $p \leq 3$, when there are jumps. On the other hand, (b) holds also if $p < 2$, under the additional assumption that σ does not vanish, but we do not need this improvement here. Observe also that (a) requires Assumption 1 only, whereas (b) requires the stronger assumption Assumption 2.

4.2. *CLT for the non-standardized statistics.* From this theorem, we deduce a CLT for our statistics $\widehat{S}(p,k,\Delta_n)_t$:

Theorem 3. *a) Under Assumption 1, let $p > 3$ and $t > 0$. Then $\Delta_n^{-1/2}$ $(\widehat{S}(p,n,\Delta_n)_t - 1)$ converge stably in law, in restriction to the set Ω_t^j of (12), to a variable $S(p,k)_t^j$ which, conditionally on $\mathcal{F}$, is centered with variance*

$$\widetilde{\mathbb{E}}\Big((S(p,k)_t^j)^2 \;\Big|\; \mathcal{F}\Big) \;=\; \frac{(k-1)p^2}{2}\ \frac{D(2p-2)_t}{B(p)_t^2}. \tag{17}$$

Moreover if the processes σ and X have no common jumps, the variable $S(p,k)_t^j$ is $\mathcal{F}$-conditionally Gaussian.

b) Under Assumption 2, assume in addition that X is continuous, and let $p \geq 2$ and $t > 0$. Then $\Delta_n^{-1/2}$ $(\widehat{S}(p,k,\Delta_n)_t - k^{p/2-1})$ converge stably in law to a variable $S(p,k)^c_t$ which, conditionally on $\mathcal{F}$, is centered normal with variance

$$\widetilde{\mathbb{E}}\left((S(p,k)^c_t)^2 \;\middle|\; \mathcal{F}\right) \;=\; M(p,k)\,\frac{A(2p)_t}{A(p)_t^2}, \tag{18}$$

where

$$M(p,k) = \frac{1}{m_p^2}\,\left(k^{p-2}(1+k)m_{2p} + k^{p-2}(k-1)m_p^2 - 2k^{p/2-1}m_{k,p}\right). \tag{19}$$

Note that (a) and (b) are not contradictory, since $\Omega_t^j = \emptyset$ when X is continuous. It is also worth noticing that the conditional variances (17) and (18), although of course random, are more or less behaving in time like $1/t$.

4.3. *Consistent estimators of the asymptotic variances.* To evaluate the level of tests based on the statistic $\widehat{S}(p,k,\Delta_n)_t$, we need consistent estimators for the asymptotic variances obtained in Theorem 3. That is, we will need to estimate $D(p)$ when $p \geq 2$ and when there are jumps, and also for $A(p)$ when $p \geq 2$ and X is continuous.

To estimate $A(p)$, we can use a realized truncated pth variation: for any constants $\alpha > 0$ and $\varpi \in (0, \frac{1}{2})$, we have from Jacod (2006) that, if Assumption 1 holds and either $p = 2$, or $p > 2$ and X is continuous, then

$$\widehat{A}(p,\Delta_n)_t \;:=\; \frac{\Delta_n^{1-p/2}}{m_p}\;\sum_{i=1}^{[t/\Delta_n]} |\Delta_i^n X|^p\;1_{\{|\Delta_i^n X| \leq \alpha\Delta_n^{\varpi}\}} \;\xrightarrow{\mathbb{P}}\; A(p)_t. \tag{20}$$

Alternatively, we can use the multipower variations of Barndorff-Nielsen and Shephard (2004). For any $r \in (0,\infty)$ and any integer $q \geq 1$, we have from Barndorff-Nielsen et al. (2006a) that if Assumption 1 holds and X is continuous

$$\tilde{A}(r,q,\Delta_n)_t \;:=\; \frac{\Delta_n^{1-qr/2}}{m_r^q}\;\sum_{i=1}^{[t/\Delta_n]-q+1}\;\prod_{j=1}^{q} |\Delta_{i+j-1}^n X|^r \;\xrightarrow{\mathbb{P}}\; A(qr)_t \tag{21}$$

(when $q = 1$, we have $\tilde{A}(r,1,\Delta_n) = (\Delta_n^{1-r/2}/m_r)\,\widehat{B}(r,\Delta_n)$).

It turns out that we also need some results concerning the behavior of $\widehat{A}(p,\Delta_n)_t$ or $\tilde{A}(r,q,\Delta_n)$ when X is discontinuous. These results will be stated below, together with the consistency of the estimators of $D(p)_t$ which we presently describe. Estimators for $D(d)_t$ are a bit more difficult to construct because we need to evaluate σ_{s-}^2 and σ_s^2 when s is a jump time, and this involves a kind of non-parametric estimation. A possibility, among many others, is as follows: take any sequence k_n of integers satisfying

$$k_n \to \infty, \qquad k_n\Delta_n \to 0, \tag{22}$$

and then let $I_{n,t}(i) = \{j \in N : \ j \neq i : 1 \leq j \leq [t/\Delta_n], |i-j| \leq k_n\}$ define a local window in time of length $k_n\Delta_n$ around time $i\Delta_n$, and

$$\widehat{D}(p,\Delta_n)_t = \frac{1}{k_n\Delta_n}\sum_{i=1}^{[t/\Delta_n]} |\Delta_i^n X|^p \sum_{j\in I_{n,t}(i)} (\Delta_j^n X)^2 1_{\{|\Delta_j^n X|\leq \alpha\Delta_n^{\varpi}\}}, \tag{23}$$

where $\alpha > 0$ and $\varpi \in (0, 1/2)$.

The following theorem establishes the consistency of these estimators:

Theorem 4. *a) Under Assumption 1, we have:*

$$p \geq 2,\ t > 0,\ \alpha > 0,\ \frac{1}{2} - \frac{1}{p} < \varpi < \frac{1}{2} \quad \Rightarrow \quad \limsup_n \frac{\Delta_n\, \widehat{A}(2p,\Delta_n)_t}{\widehat{A}(p,\Delta_n)_t^2} \ \xrightarrow{\mathbb{P}}\ 0, \tag{24}$$

$$r \in (0,2),\ q \in \mathbb{N} \quad \Rightarrow \quad \tilde{A}(r,q,\Delta_n)_t \ \xrightarrow{\mathbb{P}}\ A(qr)_t, \quad \textit{locally uniformly in } t. \tag{25}$$

b) Under Assumption 2, if $\alpha > 0$, $\varpi \in (0,1/2)$ and $p > 2$, we have

$$\widehat{D}(p,\Delta_n)_t \ \xrightarrow{\mathbb{P}}\ D(p)_t, \qquad \forall\ t \geq 0. \tag{26}$$

If further X is continuous, then

$$\Delta_n^{1-p/2}\, \widehat{D}(p,\Delta_n)_t \ \xrightarrow{\mathbb{P}}\ m_p\, A(p+2)_t, \quad \textit{locally uniformly in } t. \tag{27}$$

Note that (26) and (27) are not contradictory, because $D(p) = 0$ when X is continuous. In fact we have more than (24), namely the convergence (20) when there are jumps, but only under some restrictions on the jumps and also some restrictions on ϖ connected with the structure of the jumps and the value $p > 2$; but these refinements are useless here. Of course (25) reduces to (21) when X is continuous.

4.4. *CLT for the standardized statistics.* Using Theorem 4, we can immediately construct consistent estimators of the asymptotic variances of the test statistics established in (17) and (18), respectively. We deduce from (10), (25) and (26) the following CLT for the standardized test statistics (recall $\int_0^t |\sigma_s|^p ds > 0$ a.s., by (d) of Assumption 1):

Theorem 5. *a) Under Assumption 1, let $p > 3$ and $t > 0$. With*

$$\widehat{V}_n^j = \frac{\Delta_n\,(k-1)\,p^2\,\widehat{D}(2p-2,\Delta_n)_t}{2\widehat{B}(p,\Delta_n)_t^2}, \tag{28}$$

the variables $(\widehat{V}_n^j)^{-1/2}\ \Big(\widehat{S}(p,k,\Delta_n)_t - 1\Big)$ converge stably in law, in restriction to the set Ω_t^j of (4), to a variable which, conditionally on $\mathcal{F}$, is centered with variance 1, and which is $N(0,1)$ if in addition the processes σ and X have no common jumps.

b) Under Assumption 2, assume in addition that X is continuous, and let $p \geq 2$ and $t > 0$. The variables $(\widehat{V}_n^c)^{-1/2}\left(\widehat{S}(p,k,\Delta_n)_t - k^{p/2-1}\right)$ converge stably in law to a variable which, conditionally on $\mathcal{F}$, is $N(0,1)$, where $\widehat{V}_n^c$ is based on truncations:

$$\widehat{V}_n^c \;=\; \frac{\Delta_n\, M(p,k)\, \widehat{A}(2p,\Delta_n)_t}{\widehat{A}(p,\Delta_n)_t^2}, \tag{29}$$

or is replaced by the multipower estimator:

$$\tilde{V}_n^c \;=\; \frac{\Delta_n\, M(p,k)\, \tilde{A}(\frac{p}{[p]+1}, 2[p]+2, \Delta_n)_t}{\tilde{A}(\frac{p}{[p]+1}, [p]+1, \Delta_n)_t^2}. \tag{30}$$

In (30) we have chosen $r = p/([p]+1)$ and respectively $q = 2[p]+2$ and $q = [p]+1$. Any other choice with $r \in (0,2)$ and respectively $q = 2p/r$ and $q = p/r$ would do as well.

5. Testing for jumps. We now use the preceding results to construct actual tests, either for the null hypothesis that there are no jumps, or for the null hypothesis that jumps are present.

5.1. *When there are no jumps under the null hypothesis.* In a first case, we set the null hypothesis to be "no jump". We choose an integer $k \geq 2$ and a real $p > 3$, and associate the critical (rejection) region of the form

$$C_{n,t}^c = \{\widehat{S}(p,k,\Delta_n)_t < c_n^c\} \tag{31}$$

for some sequence $c_n^c > 0$, possibly $c_n^c = c^c$ for all n, and possibly even a random sequence.

The customary way for defining the asymptotic level of this test is as follows: if $\alpha_{n,t}^c(b,\sigma,\delta) = \mathbb{P}(C_{n,t}^c)$, a notation which emphasizes the dependency on the coefficients (b,σ,δ) of (1), one should take the supremum over all triples (b,σ,δ) in the null hypothesis (that is with $\delta \equiv 0$) of $\limsup_n \alpha_t^n(b,\sigma,\delta)$. However, as mentioned at the end of Subsection 3.1, we only observe a particular path $s \mapsto X_s(\omega)$ over $[0,t]$, and even only at times $i\Delta_n$! so there is obviously no way of statistically separating the genuine null hypothesis from the case where there are jumps, but none occurred in the interval $[0,t]$. Therefore, recalling the sets Ω_t^c and Ω_t^j of (4), it is natural to take the following as our definition of the asymptotic level:

$$\alpha = \sup_{b,\sigma,\delta}\ \limsup_n\ \mathbb{P}(C_{n,t}^c \mid \Omega_t^c), \tag{32}$$

with the convention that $\mathbb{P}(. \mid \Omega_t^c) = 0$ if $\mathbb{P}(\Omega_t^c) = 0$. Observe that we take first the lim sup and next the supremum; should we proceed the other way around, we would

find $\alpha = 1$. In a similar way, the power function for the coefficients triple (b, c, δ) is the following conditional probability:

$$\beta^c_{n,t}(b, \sigma, \delta) = \mathbb{P}(C^c_{n,t} \mid \Omega^j_t). \tag{33}$$

The right side above makes sense as soon as the function δ in restriction to $\Omega \times [0, t] \times E$ is not $\mathbb{P}(d\omega)\ ds\ \lambda(dx)$-almost everywhere vanishing, since $\mathbb{P}(\Omega^j_t) > 0$ in that case, whereas otherwise $\mathbb{P}(\Omega^j_t > 0) = 0$ and our process is continuous on $[0, t]$ (since our setting is non-homogeneous in time, a test based on observations on $[0, t]$ can of course say nothing about what happens after time t).

For $\alpha \in (0, 1)$, denote by z_α the α-quantile of $N(0, 1)$, that is $\mathbb{P}(U > z_\alpha) = \alpha$ where U is $N(0, 1)$. We have:

Theorem 6. *Under Assumption 2, let $t > 0$, choose a real $p > 3$ and an integer $k \geq 2$, and let*

$$c^c_n = k^{p/2-1} - z_\alpha \sqrt{\widehat{V}^c_n}, \tag{34}$$

where $\widehat{V}^c_n$ is given by (29) with $\alpha > 0$ and $1/2 - 1/p < \varpi < 1/2$, or is replaced by $\tilde{V}^c_n$ of (30). Then:

a) The asymptotic level (32) of the critical region defined by (31) for testing the null hypothesis "no jump" equals α.

b) The power function (37) satisfies $\beta^c_{n,t}(b, \sigma, \delta) \to 1$ for all coefficients (b, σ, δ) such that $\mathbb{P}(\Omega^j_t) > 0$ (and of course satisfying Assumption 2).

5.2. *When jumps are present under the null hypothesis.* In a second case, we set the null hypothesis to be that "there are jumps", in which case the critical (rejection) region of the form

$$C^j_{n,t} = \{\widehat{S}(p, k, \Delta_n)_t > c^j_n\} \tag{35}$$

for some sequence $c^j_n > 0$. As in (32), the asymptotic level is

$$\alpha' = \sup_{b,\sigma,\delta} \limsup_n \ \mathbb{P}(C^j_{n,t} \mid \Omega^j_t), \tag{36}$$

with the convention that $\mathbb{P}(. \mid \Omega^j_t) = 0$ if $\mathbb{P}(\Omega^j_t) = 0$, and the power function for the coefficients triple (b, σ, δ) is

$$\beta^j_{n,t}(b, \sigma, \delta) = \mathbb{P}(C^j_{n,t} \mid \Omega^c_t). \tag{37}$$

The right-hand side above is simply $\mathbb{P}(C^j_{n,t})$ when $\delta = 0$, and it makes sense as soon as $\mathbb{P}(\Omega^c_t) > 0$, whereas when $\mathbb{P}(\Omega^c_t) = 0$ then the null hypothesis is of course satisfied.

Theorem 7. *Under Assumption 1, let $t > 0$ and choose a real $p > 3$ and an integer $k \geq 2$. Let $\alpha \in (0,1)$.*

a) Letting

$$c_n^j = 1 + \sqrt{\widehat{V}_n^j/\alpha}, \tag{38}$$

where $\widehat{V}_n^j$ is given by (28), the asymptotic level (36) of the critical region defined by (31) for testing the null hypothesis "there are jumps" is smaller than α.

b) Suppose that we restrict our attention to models in which the processes X and σ have no common jumps. If

$$c_n^j = 1 + z_\alpha \sqrt{\widehat{V}_n^j}, \tag{39}$$

then the asymptotic level (36) of the critical region defined by (31) for testing the null hypothesis "there are jumps" is equal to α.

c) In all cases the power function (37) satisfies $\beta_{n,t}^j(b,\sigma,\delta) \to 1$ for all coefficients (b,σ,δ) such that $\mathbb{P}(\Omega_t^c) > 0$.

Finally, let us calculate more explicitly the critical regions for a specific choice of p. There is no special reason for choosing two different values for p in the case of the two previous null hypotheses. Since we must have $p > 3$, a rather natural choice seems to be $p = 4$. Also, in (17) and (18) we see that the variances are increasing with k, so it is probably wise to take $k = 2$ (although when $k > 2$ we have to separate the two points 1 and k, which are further apart than 1 and 2). With such choices, we obtain that $M(4,2) = 160/3$ (or more generally $M(4,k) = 16k(2k^2 - k - 1)/3$) and the cut-off point in (34) becomes

$$c_n^c = 2 - z_\alpha \sqrt{\frac{160\ \Delta_n\ \widehat{A}(8,\Delta_n)_t}{3\,\widehat{A}(4,\Delta_n)_t^2}}, \quad \text{or} \quad c_n^c = 2 - z_\alpha \sqrt{\frac{160\ \Delta_n\ \tilde{A}(4/5,10,\Delta_n)_t}{3\,\tilde{A}(4/5,5,\Delta_n)_t^2}}. \tag{40}$$

In a similar way, (39) becomes

$$c_n^j = 1 + z_\alpha \sqrt{\frac{8\Delta_n\ \widehat{D}(6,\Delta_n)_t}{\widehat{B}(4,\Delta_n)_t^2}}. \tag{41}$$

The α-quantiles of $N(0,1)$ at the 10% and 5% level are $z_{0.1} = 1.28$ and $z_{0.05} = 1.64$ respectively.

5.3. *The effect of microstructure noise.* As already said before, the observations of the process X are blurred with a (small) noise, which messes things up when data are recorded with high frequency.

We do not intend to provide a deep study of this topic here, but merely to establish some basic facts from the point of view of the consistency of our statistics (rates of convergence are more difficult to obtain, and strongly depend on the structure of

the noise). We assume that each observation is affected by an additive noise, that is instead of $X_{i\Delta_n}$ we really observe $Y_{i\Delta_n} = X_{i\Delta_n} + \varepsilon_i$, and the ε_i are supposed to be i.i.d. with $E(\varepsilon_i^2)$ and $E(\varepsilon_i^4)$ finite. Then, instead of $\widehat{B}(4,\Delta_n)_t$ (we consider the case $p=4$ here), we actually observe (for $k=1$ and also $k \geq 2$):

$$\begin{aligned}\widehat{B}'(4,k\Delta_n)_t &= \sum_{i=1}^{[t/k\Delta_n]} \left(\Delta^n_{k(i-1)}X + \varepsilon_{ki} - \varepsilon_{k(i-1)}\right)^4 \\ &= \widehat{B}(4,k\Delta_n)_t + 2\sum_{i=1}^{[t/k\Delta_n]} (\Delta^n_{k(i-1)}X)^2(\varepsilon_{ki} - \varepsilon_{k(i-1)})^2 \\ &+ \sum_{i=1}^{[t/k\Delta_n]} (\varepsilon_{ki} - \varepsilon_{k(i-1)})^4.\end{aligned}$$

The second term above behaves like $4E(\varepsilon_i^2)\widehat{B}(2,k\Delta_n)_t$, and the third one like $t\left(2E(\varepsilon_i^4) + 6E(\varepsilon_i^2)^2\right)/(k\Delta_n)$. It follows that the statistics $\widehat{S}'(4,k,\Delta_n)_t = \widehat{B}'(4,k\Delta_n)_t/\widehat{B}'(4,\Delta_n)_t$ that we actually observe instead of (11) has the following behavior, as $\Delta_n \to 0$:

$$\widehat{S}'(4,k,\Delta_n)_t \xrightarrow{\mathbb{P}} \frac{1}{k}. \tag{42}$$

The relevance of this limit will become clear when we apply the test to real data in Section 7 below. Finally, note that when Δ_n is moderately small, things may be different: if $E(\varepsilon_i^2)$ and $E(\varepsilon_i^4)$ are very small, $\widehat{S}'(4,k,\Delta_n)_t$ will be close to 1 on Ω_t^j and to k on Ω_t^c.

6. Simulation results. Throughout the simulations, we use $p=4$. We calibrate the 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 $23,400$ seconds. The simulations contain no microstructure noise.

Under the null of no jumps, the performance of the test in simulations is reported in Table 1. The distribution of the test statistic in simulations with constant σ is close to the theoretical normal limit centered at k in simulations: histograms for the test statistic and the corresponding asymptotic distribution are shown in Figure 1 for the non-standardized (top panel) and standardized test statistics (lower panel) and the cases $k=2$ (left panel) and $k=3$ (right panel.) Large values of k lead to a decrease in the effective sample size employed to estimate the numerator $\widehat{B}(p,k\Delta_n)_t$ of the test statistic, which is inefficient. There is a trade-off between using different values of k: lower values ($k=2$) lead to a smaller variance, as can be seen from (19), but on the other hand larger values such as $k=3$ better separate the test statistic under the null ($\widehat{S}(p,k,\Delta_n)_t \to k$) from its value under the alternative ($\widehat{S}(p,k,\Delta_n)_t \to 1$), as seen in Theorem 1.

Δ_n	n	k	Mean Value of $\widehat{S}(4,k,\Delta_n)$ Asymptotic	Simulations	Rejection Rate in Simulations $\alpha = 10\%$	$\alpha = 5\%$
1 sec	23,400	2	2	1.999	0.098	0.047
1 sec	23,400	3	3	2.998	0.098	0.044
1 sec	23,400	4	4	3.997	0.097	0.045
5 sec	4,680	2	2	1.999	0.104	0.050
5 sec	4,680	3	3	2.999	0.095	0.048
5 sec	4,680	4	4	3.996	0.097	0.042
10 sec	2,340	2	2	2.000	0.091	0.046
15 sec	1,560	2	2	2.002	0.086	0.038
30 sec	780	2	2	2.003	0.088	0.038

Table 1
Level of the test under the null hypothesis of no jumps.

Note: This table reports the results of 5,000 simulations of the test statistic under the null hypothesis of no jumps. the data generating process is $dX_t/X_t = \sigma dW_t$, with $\sigma = 0.4$. The test statistic is standardized with the estimator of $\widehat{V}_n^c$ given in (29) with $\alpha = 2$ and $\varpi = 0.47$ in (20).

Under the null that jumps are present, the test statistic is now centered around the predicted value of 1. We report in Table 2 the performance of the test statistic when jumps are compound Poisson. Figure 2 plots the distribution of the test statistic for different values of the jump arrival rate, λ. An interesting phenomenon happens when multiple jumps can occur with high probability (λ high). While the likelihood that jumps will take place in successive observations is small, such paths will happen over a large number of simulations. When that situation happens, we may see in $\widehat{B}(4, 2\Delta_n)$ the two jumps either compensate each other, if they are of opposite signs, or cumulate into a single larger jump if they are of the same sign. On the other hand, no such compensation or cumulation takes place in $\widehat{B}(4, \Delta_n)$. As a result, their ratio $\widehat{S}(4, 2, \Delta_n)$ can exhibit a small number of outliers.

Similar results for the case where jumps are generated by a Cauchy process are reported in Table 3 and in Figure 3 for the non-standardized (top panel) and standardized test statistics (bottom panel), for different values of the Cauchy scale parameter θ. The contrast between the top and bottom panels in Figure 3 illustrates the role of the standardization in making the test statistic asymptotically normal: recall from Theorem 3-(a) that the non-standardized statistic $\widehat{S}(4, 2, \Delta_n)$ is in general non-Normal unconditionally, whereas from Theorem 5-(a) the standardized statistic $(\widehat{V}_n^j)^{-1/2} \left(\widehat{S}(4, 2, \Delta_n) - 1\right)$ is asymptotically $N(0, 1)$ unconditionally.

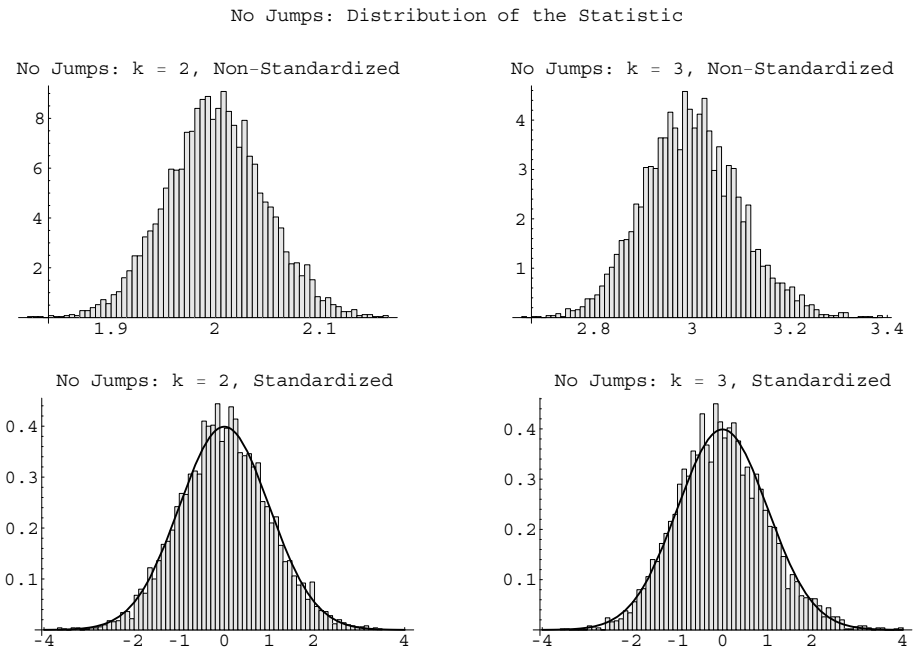

FIG. 1. *Monte Carlo and theoretical asymptotic distributions of the non-standardized (top row) and standardized (bottom row) test statistics* $S(4, k, \Delta_n)$ *for* $k = 2, 3$ *and* $\Delta_n = 1$ *second under the null hypothesis of no jumps. In the standardized case, the solid curve is the* $N(0, 1)$ *density.*

In Figure 4, we illustrate what happens when there are paths containing either no jumps or jumps that are too tiny to be detected as jumps. In the Poisson case (left panel), we set λ to correspond to 1 jump per day on average, but keep all paths, including those where no jump took place. Such paths occur with positive probability, unlike the Cauchy case. But since we condition on having jumps, we removed those paths in Figure 2 to compute the distribution of the statistic under the null that jumps are present. When those paths are kept, we obtain a clear bimodal result, with the paths where no jumps occurred grouped around 2 and those where (at least) 1 jump occurred grouped around 1.

In the Cauchy case (right panel of Figure 4), when a small value of θ is selected, although each path has infinitely many jumps, a sizeable proportion of paths has no jump large enough to make the path look different from that of a pure Brownian motion at our observation frequency. As a result, we get a bimodal distribution with a second mode at the (continuous) value of 2.

7. Empirical application. We now conduct the test for each of the 30 Dow Jones Industrial Average (DJIA) stocks 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

Δ_n	n	k	Mean value of $\widehat{S}(4,k,\Delta_n)$ Asymptotic	Simulations	Rejection Rate in Simulations $\alpha = 10\%$	$\alpha = 5\%$
				$\lambda = 1$ jump per day		
1 sec	23,400	2	1	1.000	0.101	0.050
4 sec	5,850	2	1	1.001	0.105	0.057
15 sec	1,560	2	1	1.002	0.113	0.059
				$\lambda = 5$ jumps per day		
1 sec	23,400	2	1	1.000	0.108	0.055
4 sec	5,850	2	1	1.004	0.112	0.064
15 sec	1,560	2	1	1.015	0.127	0.076
				$\lambda = 10$ jumps per day		
1 sec	23,400	2	1	1.003	0.107	0.058
4 sec	5,850	2	1	1.007	0.122	0.064
15 sec	1,560	2	1	1.027	0.159	0.093

Table 2

Compound Poisson jumps: Level of the test under the null hypothesis that jumps are present.

Note: This table reports the results of 5,000 simulations of the test statistic under the null hypothesis that jumps are present. The model under the null is $dX_t/X_t = \sigma dW_t + J_t dN_t$, where J is the product of a uniformly distributed variable on $[-2,-1]\cup[1,2]$ times a constant J_S and N is a Poisson process with intensity λ. The total variance of the increments, $\sigma^2 + (7/3)J_S^2\lambda$ is held constant at 0.4^2. As a result, jumps that are more frequent tend to be smaller in size. In the simulations, 25% of the total variance is due to the Brownian motion and 75% to the jumps. Since the test is conditional on a path containing jumps, paths that do not contain any jump are excluded from the simulated sample and replaced by new simulations. Thus, in sample, the number of jumps is slightly higher than specified by the value of λ in the table. The test statistic is standardized with the estimator of $\widehat{V}_n^j$ given in (28) with $k_n = [50\Delta_n^{-1/4}]$ in (23).

calendar time every 5 second. Each day and stock is treated on its own. We use filters to eliminate clear data errors (price set to zero, etc.) as is standard in the empirical market microstructure literature.

We plot in Figure 5 a histogram showing the empirical distribution of the statistic computed on these data. As expected, we see evidence of market microstructure noise in the form of density mass below 1 (the limit is 1/2, see Section 5.3, if the noise is i.i.d., but may be different with other kinds of noise). The striking feature of the results is that most of the observed values are around 1, providing evidence

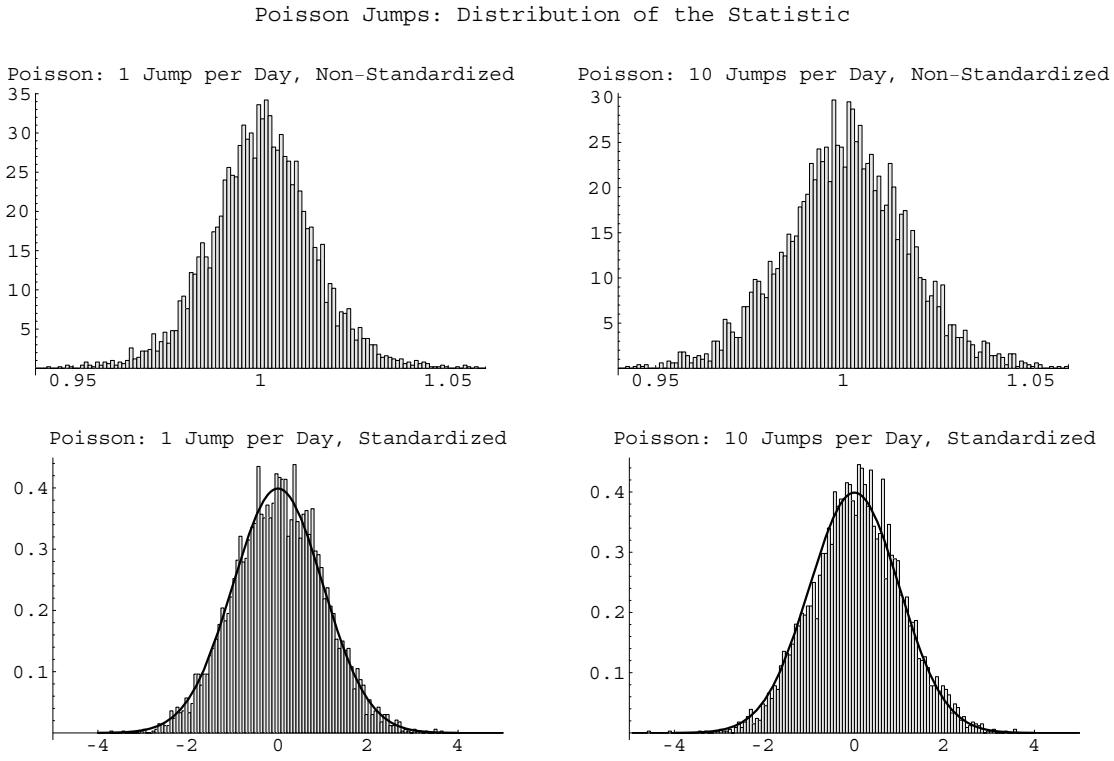

FIG. 2. *Monte Carlo and theoretical $N(0,1)$ asymptotic distributions of the standardized test statistic $(\widehat{V}_n^j)^{-1/2}\left(\widehat{S}(4,2,\Delta_n)-1\right)$ for $\Delta_n = 1$ second under the null hypothesis that jumps are present, using the same data generating processs with Poisson jumps as in Table 2.*

for the presence of jumps, with only few observations around 2, the expected limit in the continuous case. As we sample less frequently, the distribution spreads out, consistently with both the asymptotic theory and the simulations above.

8. Proofs.

8.1. *Preliminaries.* We use the shorthand notation $E_{i-1}^n(Y)$ for $E(Y \mid \mathcal{F}_{(i-1)\Delta_n})$, and we set

$$\delta_i^n = \sigma_{(i-1)\Delta_n}\Delta_i^n W, \qquad \theta_i^n = |\Delta_i^n X| 1_{\{|\Delta_i^n X| \le \alpha\Delta_n^{\varpi}\}}$$

for a given pair $\alpha > 0$ and $\varpi \in (0, \frac{1}{2})$. We will consider strengthened versions of Assumption 1 and Assumption 2:

Assumption 3. *We have Assumption 1, and $|b_t|+|\sigma_t| \le K$ and $|\delta(t,x)| \le \gamma(x) \le K$ for some constant K.*

Δ_n	n	k	Mean value of $\widehat{S}(4,k,\Delta_n)$ Asymptotic	Simulations	Rejection Rate in Simulations $\alpha = 10\%$	$\alpha = 5\%$
				Jump size: $\theta = 10$		
1 sec	23,400	2	1	1.002	0.110	0.058
5 sec	4,680	2	1	1.010	0.123	0.067
15 sec	1,560	2	1	1.027	0.142	0.083
				Jump size: $\theta = 50$		
1 sec	23,400	2	1	1.000	0.098	0.052
5 sec	4,680	2	1	1.001	0.096	0.056
15 sec	1,560	2	1	1.002	0.110	0.071

Table 3

Cauchy jumps: Level of the test under the null hypothesis that jumps are present.

Note: This table reports the results of 5,000 simulations of the test statistic under the null hypothesis that jumps are present. The model under the null is $dX_t/X_t = \sigma dW_t + \theta dY_t$, where Y is a Cauchy process standardized to have characteristic function $\mathbb{E}(\exp(iuY_t)) = \exp(-t|u|/2)$. The value of σ is 0.2. Given σ, the parameter θ measures the size of the jumps relative to the volatility. The test statistic is standardized with the estimator of $\widehat{V}_n^j$ given in (28) with $k_n = [50\Delta_n^{-1/4}]$ in (23).

Assumption 4. *We have Assumption 2, Assumption 3 and* $|\widetilde{b}_t| + |\widetilde{\sigma}_t| + |\widetilde{\sigma}'_t| \leq K$ *and* $|\widetilde{\delta}(t,x)| \leq \gamma(x)$ *for some constant* K.

A localization procedure shows that for proving Theorems 2 and 4 we can replace everywhere Assumption 1 or Assumption 2 by Assumptions 3 or 4, respectively: this localization procedure is described in details in Jacod (2006) and works similarly here, so we omit it.

Now we state a number of consequences of this strengthened assumptions, to be used a number of times in the following proofs. Below. K denotes a constant which may depend on the coefficients (b,c,δ) and $(\widetilde{b},\widetilde{\sigma},\widetilde{\sigma}',\widetilde{\delta}')$ and which changes from line to line. We write it K_a if it depends on an additional parameter a. We also write X'_t and X''_t for the sum of the first three terms, resp. of the last two terms, in (1). First, for all $q \geq 2$ we have the following classical estimates (proved e.g. in Jacod (2006)):

$$\begin{array}{l} \text{Assumption 3} \Rightarrow \mathbb{E}^n_{i-1}(|\Delta^n_i X'|^q + |\delta^n_i|^q) \leq K_q \Delta_n^{q/2},\ \mathbb{E}^n_{i-1}(|\Delta^n_i X''|^q) \leq K_q \Delta_n \\ \text{Assumption 4} \Rightarrow \mathbb{E}^n_{i-1}(|\Delta^n_i \sigma|^q) \leq K_q \Delta_n,\ \mathbb{E}^n_{i-1}(|\Delta^n_i X' - \delta^n_i|^q) \leq K \Delta_n^{1+q/2}. \end{array} \tag{43}$$

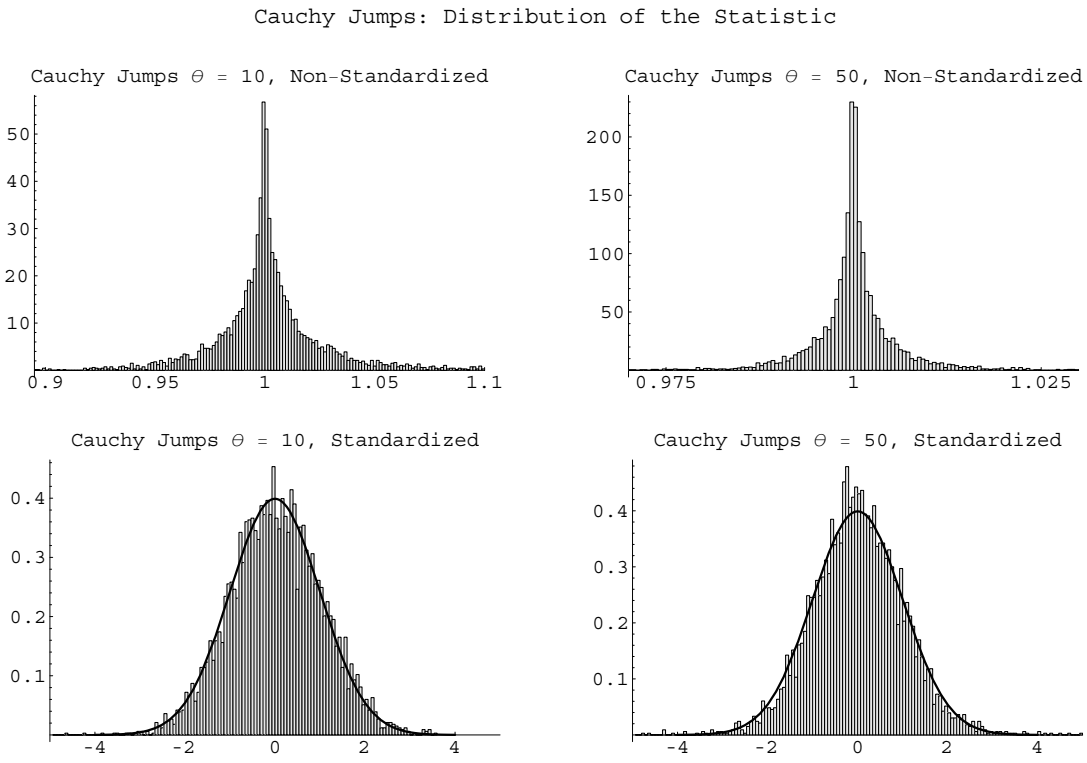

FIG. 3. *Monte Carlo distribution of the non-standardized test statistic* $S(4,2,\Delta_n)$ *for* $\Delta_n = 1$ *second under the null hypothesis that jumps are present, using the same data generating process with Cauchy jumps as in Table 3. The asymptotic distribution of the non-standardized statistic is non-normal as expected (see upper panel). However, the asymptotic distribution of the standardized test statistic is* $N(0,1)$ *(the solid curve in the lower panel).*

Next, the proof of Lemma 5.12 of Jacod (2006) applied with ε being the supremum of the bounded function γ shows that under Assumption 3,

$$\mathbb{E}^n_{i-1}(|\Delta^n_i X''|^2 \bigwedge \eta^2) \leq K\Delta_n \left(\frac{\eta^2 + \Delta_n}{\theta^2} + \Gamma(\theta)\right) \tag{44}$$

for all $\eta > 0$ and $\theta \in (0,1)$, and where $\Gamma(\theta) \to 0$ as $\theta \to 0$. Combining this with the second line in (43) yields that under Assumption 4 we also have

$$\mathbb{E}^n_{i-1}(|\Delta^n_i X - \delta^n_i|^2 \bigwedge \eta^2) \leq K\Delta_n \left(\frac{\eta^2 + \Delta_n}{\theta^2} + \Gamma(\theta)\right) \tag{45}$$

Finally let us also mention two elementary (although a bit tedious to show) inequalities, valid for all $x, y \in \mathbb{R}$ and $0 < \varepsilon < 1 < A$ and for $p \geq 2$ and $0 < r < 2$:

$$\left||x+y|^p 1_{\{|x+y|\leq\varepsilon\}} - |x|^p\right| \leq K_p\left(|x|^p 1_{\{|x|>\varepsilon/2\}} + \varepsilon^{p-2}(y^2 \wedge \varepsilon^2) + |x|^{p-1}(|y| \wedge \varepsilon)\right) \tag{46}$$

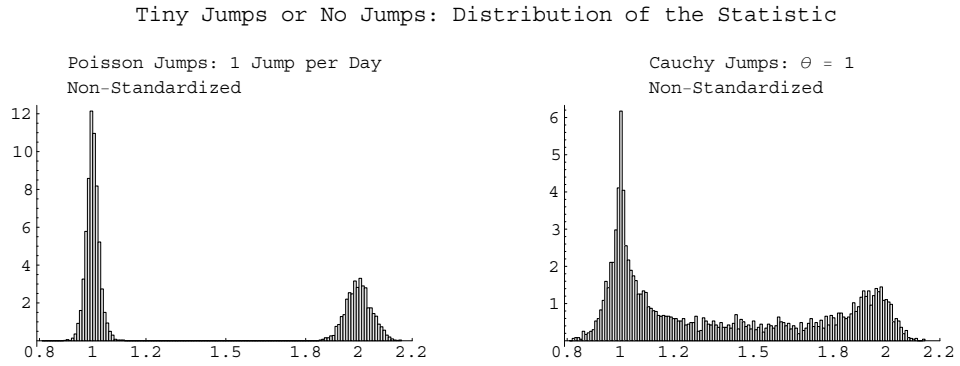

FIG. 4. *Monte Carlo distribution of the non-standardized test statistic* $S(4,2,\Delta_n)$ *for* $\Delta_n = 1$ *second, computed using a data generating process with either one Poisson jump per day on average including paths that contain no jumps (left panel) or tiny Cauchy jumps* ($\theta = 1$, *right panel).*

$$\left||x+y|^r - |x|^r\right| \ \leq \ K_r\left(\varepsilon^r + A\varepsilon + A^{r-2}(x^2+y^2) + A^r\varepsilon^{-2}(y^2 \wedge 1)\right). \tag{47}$$

8.2. *Proof of Theorem 2.* For proving Theorem 2 we need to exhibit the limits $Z(p)$, $Z'(p,k)$, $Y(p)$ and $Y'(p,k)$ and it takes some additional notation to do so. We consider an auxiliary space $(\Omega', \mathcal{F}', \mathbb{P}')$ which supports a number of variables and processes:

- four sequences (U_q), (U'_q), $(\overline{U}_q)$, $\overline{U}'_q)$ of $N(0,1)$ variables;
- a sequence (κ_q) of uniform variables on $[0,1]$;
- a sequence (L_q) of uniform variables on the finite set $\{0,1,\cdots,k-1\}$ ($k \geq 2$ is the integer showing up in the theorem).
- two standard Wiener processes $\overline{W}$ and $\overline{W}'$;

and all these processes or variables are mutually independent. Then we put

$$\widetilde{\Omega} = \Omega \times \Omega', \qquad \widetilde{\mathcal{F}} = \mathcal{F} \otimes \mathcal{F}', \qquad \widetilde{\mathbb{P}} = \mathbb{P} \otimes \mathbb{P}'.$$

and we extend the variables X_t, b_t, ... defined on Ω and $\overline{W}$, U_q,... defined on Ω' to the product $\widetilde{\Omega}$ in the obvious way, without changing the notation. We write $\widetilde{\mathbb{E}}$ for the expectation w.r.t. $\widetilde{\mathbb{P}}$. Finally, denote by $(T_n)_{n\geq 1}$ an enumeration of the jump times of X which are stopping times, and let $(\widetilde{\mathcal{F}}_t)$ be the smallest (right-continuous) filtration of $\widetilde{\mathcal{F}}$ containing the filtration $(\mathcal{F}_t)$ and w.r.t. which $\overline{W}$ and $\overline{W}'$ are adapted and such that U_n, U'_n, $\overline{U}_n$, $\overline{U}'_n$, κ_n and L_n are $\widetilde{\mathcal{F}}_{T_n}$-measurable for all n. We hence get an extension $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, (\widetilde{\mathcal{F}}_t)_{t\geq 0}, \widetilde{\mathbb{P}})$ of the original space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$. Obviously, W, W', $\overline{W}$, $\overline{W}'$ are Wiener processes and μ is a Poisson measure with compensator ν on $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, (\widetilde{\mathcal{F}}_t)_{t\geq 0}, \widetilde{\mathbb{P}})$.

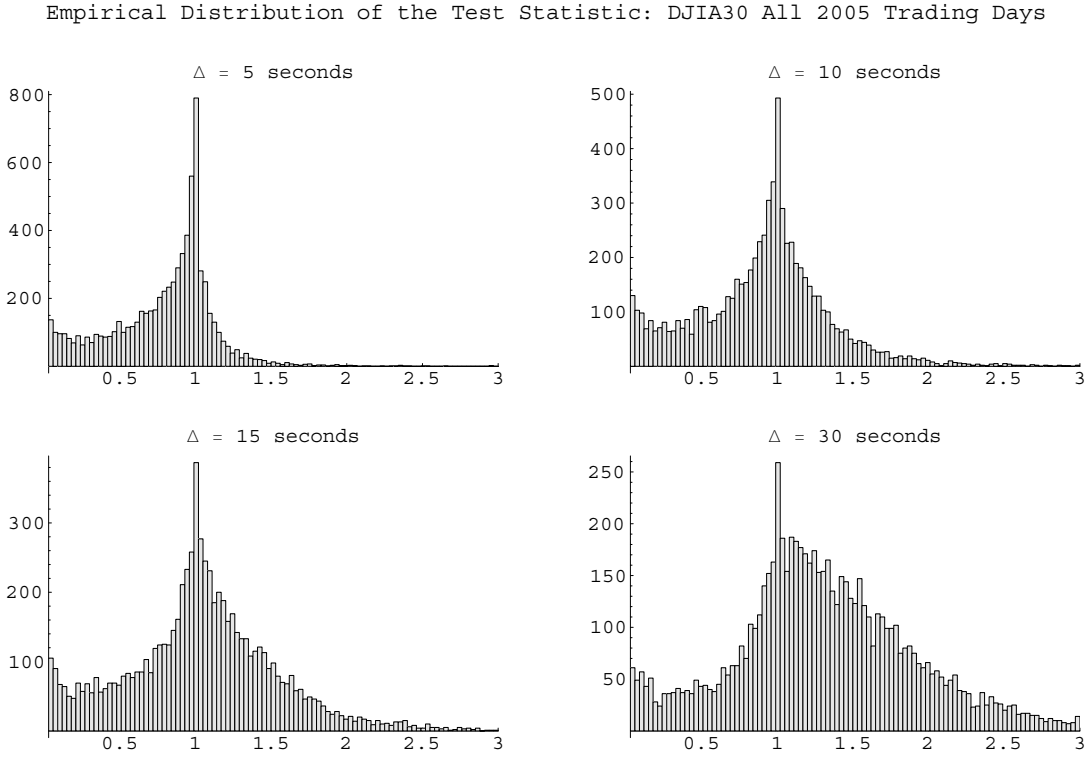

FIG. 5. *Empirical distribution of the non-standardized statistic* $\widehat{S}(4,2,\Delta_n)$ *for different values of the sampling interval* Δ_n. *Each sample point is computed using all the transactions for one of the 30 DJIA stocks observed over one trading day in 2005. This produces* $7,560$ *realizations of the statistic.*

Now we exhibit the limits. If (h_s) and (h'_s) are adapted processes with right-continuous or left-continuous paths, defined on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, we set

$$Y(h,h')_t = \int_0^t h_s \, dBW_s + \int_0^t h'_s \, d\overline{W}'_s. \tag{48}$$

This defines a local martingale on the extension, which conditionally on the σ-field $\mathcal{F}$ is a centered Gaussian martingale. Furthermore if (k_s, k'_s) is another pair of processes we have

$$\widetilde{\mathbb{E}}(Y(h,h')_t \, Y(k,k')_t \mid \mathcal{F}) = \int_0^t (h_s k_s + h'_s k'_s) ds. \tag{49}$$

In a similar way, with any function g on R which is locally bounded and with $g(x)/x \to 0$ as $x \to 0$, we associate the following two processes

$$Z(g)_t = \sum_{q:\ T_q \leq t} g(\Delta X_{T_q}) \Big(\sqrt{\kappa_q} \, U_q \, \sigma_{T_q -} + \sqrt{1-\kappa_q} \, U'_q \, \sigma_{T_q} \Big), \tag{50}$$

$$Z'(g)_t = \sum_{q:\ T_q \le t} g(\Delta X_{T_q})\Big(\sqrt{L_q}\ \overline{U}_q\ \sigma_{T_q-} + \sqrt{k-1-L_q}\ \overline{U}'_q\ \sigma_{T_q}\Big). \tag{51}$$

That $Z(g)$ is well defined is proved in Jacod (2006), and the proof that $Z'(g)$ is also well defined is exactly the same. Again, conditionally on $\mathcal{F}$, these two processes are independent martingales with mean 0 and

$$\left.\begin{array}{l} \widetilde{\mathbb{E}}(Z(g)_t^2 \mid \mathcal{F}) = \frac{1}{2}\ \sum_{s\le t} g(\Delta X_s)^2(\sigma_{s-}^2 + \sigma_s^2) \\ \widetilde{\mathbb{E}}(Z'(g)_t^2 \mid \mathcal{F}) = \frac{k-1}{2}\ \sum_{s\le t} g(\Delta X_s)^2(\sigma_{s-}^2 + \sigma_s^2) \end{array}\right\} \tag{52}$$

Moreover, conditionally on $\mathcal{F}$, $Y(h,h')$ and $(Z(g), Z'(g))$ are independent, and the latter is also a Gaussian martingale as soon as the processes X and σ have no common jumps.

At this stage we can proceed to proving Theorem 2-(a). As a matter of fact it is clearer to prove a more general result: with k an integer strictly bigger that 1 and with any function f on R, we consider the two processes

$$\left.\begin{array}{l} V^n(f)_t = \sum_{i=1}^{[t/\Delta_n]} f(\Delta_i^n X) \\ \overline{V}^n(f)_t = \sum_{i=1}^{[t/k\Delta_n]} f(\Delta^n_{ki-k+1}X + \Delta^n_{ki-k+2}X + \cdots + \Delta^n_{ki}X). \end{array}\right\} \tag{53}$$

We also write $V(f)_t = \sum_{s\le t} f(\Delta X_s)$, which is well defined as soon as $f(x)/x^2$ is bounded on a neighborhood of 0.

Theorem 8. *Under Assumption 1, let f be C^2 with $f(0) = f'(0) = f''(0) = 0$ (f' and f'' are the first two derivatives). The pair of processes*

$$\Big(\frac{1}{\sqrt{\Delta_n}}\ (V^n(f)_t - V(f)_{\Delta_n[t/\Delta_n]}), \frac{1}{\sqrt{\Delta_n}}\ (\overline{V}^n(f)_t - V(f)_{k\Delta_n[t/\Delta_n]})\Big)$$

converges stably in law, on the product $\mathbb{D}(\mathbb{R}_+,\mathbb{R})\times\mathbb{D}(\mathbb{R}_+,\mathbb{R})$ of the Skorokhod spaces, to the process $(Z(f'), Z(f') + Z'(f'))$.

We have the (stable) convergence in law of the above processes, as elements of the product functional space $D(R_+,R)^2$, but usually not as elements of the space $D(R_+,R^2)$ with the (2-dimensional) Skorokhod topology, because a jump of X entails a jump for both components above, but "with a probability close to j/k" the times at which these two component jump differ by an amount $j\Delta_n$, for $j = 1,\cdots,k-1$.

Proof. As said in the beginning of this Section, we can assume Assumption 3. The proof is an extension of the proof of Theorem 2.12-(i) of Jacod (2006). We start the proof under the additional assumption that f vanishes on $[-2k\varepsilon, 2k\varepsilon]$ for some $\varepsilon > 0$. Let S_q be the successive jump times of the Poisson process $\mu([0,t]\times\{x : \gamma(x) > \varepsilon\})$. Let $\Omega_n(t,\varepsilon)$ be the set of all ω such that each interval $[0,t]\cap(i\Delta_n,(i+k)\Delta_n]$ contains at most one $S_q(\omega)$, and $S_1(\omega) > k\Delta_n$, and $|X_{(i+1)\Delta_n}(\omega)| - X_{i\Delta_n}(\omega)| \le 2\varepsilon$ for all

$i \le t/\Delta_n$ and $|X_{(i+1)k\Delta_n}(\omega)| - X_{ik\Delta_n}(\omega)| \le 2\varepsilon$ for all $i \le t/k\Delta_n$. Next, on the set $\{(ik+j)\Delta_n < S_q \le (ik+j+1)\Delta_n\}$ for $i \ge 1$ and $0 \le j < k$, we set

$$\begin{cases} L(n,q) = j \\ K(n,q) = \frac{S_p}{\Delta_n} - (ik+j) \\ \alpha_-(n,q) = \frac{1}{\sqrt{\Delta_n}}(W_{S_q} - W_{(ik+j)\Delta_n}), \quad \alpha_+(n,q) = \frac{1}{\sqrt{\Delta_n}}(W_{(ik+j+1)\Delta_n} - W_{S_q}) \\ \beta_-(n,q) = \frac{1}{\sqrt{\Delta_n}}(W_{(ik+j)\Delta_n} - W_{ik\Delta_n}), \ \beta_+(n,q) = \frac{1}{\sqrt{\Delta_n}}(W_{(i+1)k\Delta_n} - W_{(ik+j+1)\Delta_n}) \\ R_-(n,q) = X(\varepsilon)_{(ik+j)\Delta_n} - X(\varepsilon)_{ik\Delta_n}, \quad R'_+(n,q) = X(\varepsilon)_{(i+1)k\Delta_n} - X(\varepsilon)_{(ik+j+1)\Delta_n} \\ R^n_q = \Delta^n_{(ik+j+1)\Delta_n} X(\varepsilon), \qquad R'^n_q = R^n_q + R_-(n,q) + R_+(n,q) \\ R_q = \sqrt{\kappa_q}\ U_q\ \sigma_{S_q-} + \sqrt{1-\kappa_q}\ U'_q\ \sigma_{S_q} \\ R_-(q) = \sqrt{L_q}\ \overline{U}_q\ \sigma_{S_q-}, \quad R_+(q) = \sqrt{k-1-L_q}\ \overline{U}'_q\ \sigma_{S_q} \end{cases}$$

We can extend the proof of Lemma 6.2 of Jacod and Protter (1998) in this more complicated context, and with $\overset{\mathcal{L}-(s)}{\longrightarrow}$ denoting the stable convergence in law, in order to obtain that

$$\Big(L(n,q), K(n,q), \alpha_-(n,q), \alpha_+(n,q), \beta_-(n,q), \beta_+(n,q)\Big)_{q\ge1} \overset{\mathcal{L}-(s)}{\longrightarrow} \Big(L_q, \kappa_q, \sqrt{\kappa_q}\ U_q, \sqrt{1-\kappa_q}\ U'_q, \sqrt{L_q}\ \overline{U}_q, \sqrt{k-L_q-1}\ \overline{U}'_q\Big)_{q\ge1}.$$

From this we deduce, as in Lemma 5.9 of Jacod (2006), that

(54)
$$\Big(R^n_q/\sqrt{\Delta_n}, R_-(n,q)/\sqrt{\Delta_n}, R_+(n,q)/\sqrt{\Delta_n}\Big)_{q\ge1} \overset{\mathcal{L}-(s)}{\longrightarrow} \Big(R_q, R_-(q), R_+(q)\Big)_{q\ge1}.$$

Now, since $f(x) = 0$ if $|x| \le 2k\varepsilon$, on the set $\Omega_n(t,\varepsilon)$ and for all $s \le t$ we have

$$\begin{aligned} V^n(f)_s - V(f)_{\Delta_n[s/\Delta_n]} &= \sum_{q:\ S_q \le \Delta_n[s/\Delta_n]} \big(f(\Delta X_{S_q} + R^n_q) - f(\Delta X_{S_q})\big) \\ &= \sum_{q:\ S_q \le \Delta_n[s/\Delta_n]} f'(\Delta X_{S_q} + \widetilde{R}^n_q)\ R^n_q \end{aligned}$$

where $\widetilde{R}^n_p$ is between ΔX_{S_q} and $\Delta X_{S_q} + R^n_p$. In a similar way, we get

$$\overline{V}^n(f)_s - V(f)_{k\Delta_n[s/k\Delta_n]} = \sum_{q:\ S_q \le k\Delta_n[s/k\Delta_n]} f'(\Delta X_{S_q} + \widetilde{R}'^n_q)\ R'^n_q,$$

with $\widetilde{R}'^n_q$ between ΔX_{S_q} and $\Delta X_{S_q} + R'^n_-$. Since R^n_p, $R_-(n,p)$ and $R_+(n,p)$ go to 0, we deduce that $\widetilde{R}^n_p$ and $\widetilde{R}'^n_p$ also go to 0 (for each ω), whereas $\Omega_n(t,\varepsilon) \to \Omega$. Since f' is continuous, we readily deduce the result of the theorem from (54).

Now we turn to the general case, where f does not necessarily vanish around 0. With ψ_ρ a C^2 function equal to 1 on $[-\rho,\rho]$ and to 0 outside $[-2\rho,2\rho]$ and with $f_\rho = f\psi_\rho$, we have for all $\eta > 0$:

$$\lim_{\rho\to 0} \limsup_n \mathbb{P}\Big(\sup_{s\le t} \big|(V^n(f_\rho)_s - V(f_\rho)_{\Delta_n[s/\Delta_n]}\big| / \sqrt{\Delta_n} > \eta) = 0$$

which is (5.48) in Jacod (2006). We obviously have the same for $\overline{V}^n$. Since we have the stable convergence in law when we take the functions $f - f_\rho$, and since obviously $Z(f'_\rho)$ and $Z'(f'_\rho)$ converge locally uniformly in time to $Z(f')$ and $Z'(f')$ respectively, we readily deduce the stable convergence in law for the function f. □

We now prove part (a) of the theorem.

Proof of Theorem 2-(a) under Assumption 3. If we apply the previous theorem with the function $f(x) = |x|^p$ (recall $p > 3$), we obtain the stable convergence in law of the processes

$$\Big(\frac{1}{\sqrt{\Delta_n}}\,(\widehat{B}(p,\Delta_n)_t - B(p)_{\Delta_n[t/\Delta_n]}), \frac{1}{\sqrt{\Delta_n}}\,(\widehat{B}(p,k\Delta_n)_t - B(p)_{k\Delta_n[t/k\Delta_n]})\Big)$$

towards $(Z(f')_t, Z(f')_t + Z'(f')_t)$, and $f'(x) = p|x|^{p-1}\mathrm{sign}(x)$. Hence if $Z'(p,k)_t = Z'(f')_t$ the formula (14) follows from (52).

Hence it remains to prove that $(B(p)_t - B(p)_{k\Delta_n[t/k\Delta_n]})/\sqrt{\Delta_n} \xrightarrow{\mathbb{P}} 0$ for all integers k. Since $|\delta(\omega,t,x)| \le \gamma(x) \le K$ we have

$$\begin{aligned}
\mathbb{E}(B(p)_t - B(p)_{k\Delta_n[t/k\Delta_n]}) &= \mathbb{E}\left(\sum_{k\Delta_n[t/k\Delta_n]<s\le t} |\Delta X_s|^p\right)\\
&= \mathbb{E}\left(\int_{k\Delta_n[t/k\Delta_n]}^t ds \int_E |\delta(s,x)|^q \lambda(dx)\right)\\
&\le \int_{k\Delta_n[t/k\Delta_n]}^t ds \int_E \gamma(x)^q \lambda(dx) \;\le\; Kk\Delta_n
\end{aligned}$$

where $K = \int \gamma(x)^p \lambda(dx)$ is finite (recall $p > 3$), and the result follows. □

For Theorem 2-(b) it is also convenient to prove a more general result. Let $g = (g_j)_{1\le j\le 2}$ be an R^2-valued C^2 function on R^k with second partial derivatives having polynomial growth, and which is globally even (that is $g(-x) = g(x)$) and set

$$V^n(g)_t = \sum_{i=1}^{[t/k\Delta_n]} g(\Delta^n_{ki-k+1}X/\sqrt{\Delta_n}, \Delta^n_{ki-k+2}X/\sqrt{\Delta_n}, \cdots, \Delta^n_{ki}X/\sqrt{\Delta_n}). \tag{55}$$

We also denote by ρ_y the normal law $N(0,y)$, and by $\rho_y^{k\otimes}$ its k-fold tensor product, and by $\rho_y^{k\otimes}(f)$ the integral of any function f w.r.t. it. With the above assumptions we then have the following (a d-dimensional version is also available):

Lemma 1. *Under Assumption 2, assume in addition that X is continuous. Then*

$$\Delta_n V^n(g)_t \xrightarrow{\mathbb{P}} V(g)_t := \frac{1}{k}\int_0^t \rho^{k\otimes}_{\sigma_u}(g). \tag{56}$$

Furthermore the processes $\frac{1}{\sqrt{\Delta_n}}(\Delta_n V^n(g) - V(g))$ converge stably in law to the 2-dimensional process $(Y(h,0), Y(h',h''))$ (recall (48), where the processes (h_t), $h'_t)$ and (h''_t) is such that the matrix $\begin{pmatrix} h_t & 0 \\ h'_t & h''_t \end{pmatrix}$ is a square-root of the matrix $\Theta_t = (\Theta^{ij}_t)_{1\le i,j\le 2}$ given by

$$\Theta^{ij}_t = \frac{1}{k}\left(\rho^{k\otimes}_{\sigma_t}(g_i g_j) - \rho^{k\otimes}_{\sigma_t}(g_i)\rho^{k\otimes}_{\sigma_t}(g_j\right). \tag{57}$$

Proof. The proof of this lemma is similar to the uni-power case in Barndorff-Nielsen et al. (2006a). We indicate the few changes that are necessary. Everywhere the sums over i from 1 to $[nt]$ are replaced by sums from 1 to $[t/k\Delta_n]$. In (4.1) of that paper we replace $\beta^n_i = \sqrt{n}\,\sigma_{(i-1)/n}\,\Delta^n_i W$ by the collection $\beta^n_i(j) = \frac{1}{\sqrt{\Delta_n}}\,\sigma_{k(i-1)\Delta_n}\,\Delta^n_{k(i-1)+j}W$ for $j = 1,\cdots,k$. Then the same proof as for Proposition 4.1 of that paper shows that if

$$\widetilde{V}^n(g) = \sum_{i=1}^{[t/k\Delta_n]} g(\beta^n_i(1), \beta^n_i(2), \cdots, \beta^n_i(k)),$$

we have (56) and the stable convergence in law with $\widetilde{V}'^n(g)$ instead of $V'^n(g)$, and the same process $V'(g)$. Next, similar to Theorem 5.6 of that paper, and with $\zeta^n_i = g(\Delta^n_{ki-k+1}X/\sqrt{\Delta_n}, \cdots, \Delta^n_{ki}X/\sqrt{\Delta_n})$, the processes

$$\sqrt{\Delta_n}\sum_{i=1}^{[t/k\Delta_n]}\left(\zeta^n_i - \mathbb{E}(\zeta^n_i \mid \mathcal{F}_{k(i-1)\Delta_n})\right)$$

converge stably in law to $(Y(h,0), Y(h',h''))$, and this easily yields (56) for $V'^n(g)$.

For the stable convergence in law, it remains to prove that the array

$$\zeta'^n_i = \sqrt{\Delta_n}\left(\mathbb{E}(\zeta^n_i \mid \mathcal{F}_{k(i-1)\Delta_n}) - \frac{1}{k\Delta_n}\int_{k(i-1)\Delta_n}^{ki\Delta_n} \rho^{k\otimes}_{\sigma_u}(g)du\right)$$

satisfies $\sum_{i=1}^{[t/\Delta_n]} |\zeta'^n_i| \xrightarrow{\mathbb{P}} 0$. This is proved as in Barndorff-Nielsen et al. (2006a): the fact that g is a function of several variables makes no real difference, and since g here is C^2 we are in the case of Hypothesis (K) of that paper (and not in the more complicated case of Hypothesis (K')). □

Finally, we prove part (b) of the theorem.

Proof of Theorem 2-(b) under Assumption 4. We simply apply the previous theorem to the even function g with components

$$g_1(x_1,\cdots,x_k) = |x_1|^p + \cdots + |x_k|^p, \qquad g_2(x_1,\cdots,x_k) = |x_1 + \cdots + x_k|^p$$

(recall that $p \geq 2$, hence g is C^2). The matrix Θ_t of (57) is then

$$\Sigma_t^{11} = (m_{2p} - m_p^2)|\sigma_t|^{2p}, \quad \Sigma_t^{12} = (m_{k,p} - k^{p/2} m_p^2)|\sigma_t|^{2p}, \quad \Sigma_t^{22} = k^{p-1}\,(m_{2p} - m_p^2)|\sigma_t|^{2p}.$$

Then a version of the triple (h, h', h'') is given by $h_t = \alpha|\sigma_t|^p$ and $h'_t = \alpha'|\sigma_t|^p$ and $h''_t = \alpha''|\sigma_t|^p$, where

$$\alpha = \sqrt{m_{2r} - m_r^2}, \quad \alpha' = \frac{1}{\alpha}\,(m_{k,r} - k^{r/2} m_r^2), \quad \alpha'' = \sqrt{k^{r-1}(m_{2r} - m_r^2) - \alpha\alpha'}.$$

Hence the result obtains, with $Y(p)_t = Y(h, 0)_t$ and $Y'(p, k)_t = Y(h', h'')_t$, upon noting that (15) follows from (49). □

8.3. *Proof of Theorem 3.*

Proof. a) Write $U_n = (\Delta_n)^{-1/2}\,(\widehat{B}(p, \Delta_n)_t - B(p)_t)$ and $V_n = (\Delta_n)^{-1/2}\,(\widehat{B}(p, k\Delta_n)_t - B(p)_t)$. Then

$$\widehat{S}(p, k, \Delta_n)_t - 1 = \frac{\widehat{B}(p, k\Delta_n)_t^n}{\widehat{B}(p, \Delta_n)_t} - 1 = (\Delta_n)^{1/2}\,\frac{V_n - U_n}{\widehat{B}(p, \Delta_n)_t}.$$

Theorem 3 yields that under Assumption 1 and if $p > 3$, then $V_n - U_n$ converges stably in law to $Z'(p, k)_t$, and the result follows from (14).

b) Write $U'_n = (\Delta_n)^{-1/2}\,(\Delta_n^{1-p/2}\widehat{B}(p, \Delta_n) - m_p A(p))$ and $V'_n = (\Delta_n)^{-1/2}\,(\Delta_n^{1-p/2}\widehat{B}(p, k\Delta_n) - k^{p/2-1} m_p A(p))$. Then

$$\widehat{S}(p, k, \Delta_n)_t - k^{p/2-1} = \frac{\widehat{B}(p, k\Delta_n)_t^n}{\widehat{B}(p, \Delta_n)_t} - k^{p/2-1} = (\Delta_n)^{1/2}\,\frac{V'_n - k^{p/2-1}U'_n}{\widehat{B}(p, \Delta_n)_t}.$$

Since $V'_n - k^{p/2-1}U'_n$ converges stably in law to $Y'(p, k)_t - k^{p/2-1}Y(p)_t$ under Assumption 2 and when $p \geq 2$ and X is continuous, the result easily follows from (15). □

8.4. *Proof of Theorem 4.*

Proof of (24) under Assumption 3. We fix $p \geq 2$, $t > 0$, $\varpi \in (1/2 - 1/p, 1/2)$ and $\alpha > 0$. Apply (46 to get, for any $B \geq 1$:

$$\left|\Delta_n^{1-p/2} \sum_{i=1}^{[t/\Delta_n]} |\Delta_i^n X|^p 1_{\{|\Delta_i^n X| \leq \sqrt{B\Delta_n}\}} - \Delta_n^{1-p/2} \sum_{i=1}^{[t/\Delta_n]} |\Delta_i^n X'|^p\right| \leq K_p(A_n + A'_n + A''_n), \tag{58}$$

where

$$\begin{cases} A_n = \Delta_n^{1-p/2} \sum_{i=1}^{[t/\Delta_n]} |\Delta_i^n X'|^p 1_{\{|\Delta_i^n X'| > \sqrt{B\Delta_n}/2\}} \\ A'_n = B^{p/2-1} \sum_{i=1}^{[t/\Delta_n]} (|\Delta_i^n X''|^2 \wedge B\Delta_n) \\ A''_n = \Delta_n^{1-p/2} \sum_{i=1}^{[t/\Delta_n]} |\Delta_i^n X'|^{p-1} (|\Delta_i^n X''| \wedge \sqrt{B\Delta_n}). \end{cases}$$

Then if we apply (43) and (44) with $\eta = \sqrt{B\Delta_n}$ and $\theta = \Delta_n^{1/4}$, and also Bienaymé-Chebyshev for A_n and Cauchy-Schwarz for A''_n, and with the notation $\varepsilon_n^2 = \Delta_n^{1/2} + \Gamma(\Delta_n^{1/4})$ (so $\varepsilon \to 0$), we readily obtain (recall $B \geq 1$):

$$\mathbb{E}(A_n) \leq \frac{K_p t}{B}, \qquad \mathbb{E}(A'_n) \leq K_p t B^{p/2} \varepsilon_n^2, \qquad \mathbb{E}(A''_n) \leq K_p t B \varepsilon_n.$$

On the other hand we can apply the last property in (10) to X', which satisfies Assumption 3 and is continuous, to get $\Delta_n^{1-p/2} \sum_{i=1}^{[t/\Delta_n]} |\Delta_i^n X'|^p \xrightarrow{\mathbb{P}} m_p A(p)_t$. Combining this with the above estimates and (58), and since $\varepsilon_n \to 0$, we obtain for some constant $K = K_p$ and all $\eta > 0$:

$$\mathbb{P}\left(\Delta_n^{1-p/2} \sum_{i=1}^{[t/\Delta_n]} |\Delta_i^n X|^p 1_{\{|\Delta_i^n X| \leq \sqrt{B\Delta_n}\}} < m_p A(p)_t - \frac{K}{B} - \eta \right) \ \to \ 0. \tag{59}$$

Now, for any $B \geq 1$ we have $\alpha \Delta_n^{\varpi} > \sqrt{B\Delta_n}$ for all n large enough because $\varpi < 1/2$. Therefore (59) remains valid if we replace the sets $\{|\Delta_i^n X| \leq \sqrt{B\Delta_n}\}$ by $\{|\Delta_i^n X| \leq \alpha \Delta_n^{\varpi}\}$. Since $A(p)_t > 0$ a.s. and B is arbitrarily large and η arbitrarily small, we deduce that $C_n = \sum_{i=1}^{[t/\Delta_n]} |\Delta_i^n X|^p 1_{\{|\Delta_i^n X| \leq \alpha \Delta_n^{\varpi}\}}$ satisfies

$$\mathbb{P}\left(\frac{1}{\Delta_n^{1-p/2} C_n} > \frac{2}{m_p A(p)_t} \right) \ = \ \mathbb{P}\left(\Delta_n^{1-p/2} C_n < \frac{m_p A(p)_t}{2} \right) \ \to \ 0. \tag{60}$$

At this stage, the proof of (24) is straightforward. Indeed, since $|\Delta_i^n X|^{2p} \leq \alpha^p \Delta_n^{p\varpi} |\Delta_i^n X|^p$ when $|\Delta_i^n X| \leq \alpha \Delta^{\varpi}$, one deduces from (20) that

$$\frac{\Delta_n \widehat{A}(2p, \Delta_n)_t}{\widehat{A}(p, \Delta_n)_t^2} \ \leq \ \frac{K \Delta_n^{p\varpi}}{C_n} \ = \ \frac{K \Delta_n^{p\varpi + 1 - p/2}}{\Delta_n^{1-p/2} C_n}.$$

Since $p\varpi + 1 - p/2 > 0$, the result readily follows from (60). □

Proof of (25) under Assumption 3. Barndorff-Nielsen et al. (2006b) proved a similar result under more restrictive conditions. Apply (44) with $\eta = \sqrt{\Delta_n}$ and $\varepsilon = \Delta_n^{1/4}$ to get

$$\mathbb{E}_{i-1}^n \left(\left| \Delta_i^n X'' \right|^2 \bigwedge \Delta_n \right) \leq \Delta_n \alpha_n \tag{61}$$

for a (deterministic) sequence α_n going to 0. Next, we apply (47) with $0 < r < 2$ and $x = \Delta_i^n X' / \sqrt{\Delta_n}$ and $y = \Delta_i^n X'' / \sqrt{\Delta_n}$ and we use (43) and (61) to get

$$\Delta_n^{-r/2} \mathbb{E}_{i-1}^n \left(\left| |\Delta_i^n X|^r - |\Delta_i^n X'|^r \right| \right) \leq K \left(\varepsilon^r + A\varepsilon + A^{r-2} + A^r \varepsilon^{r-2} \alpha_n \right).$$

Hence by taking first $n \to \infty$, then $\varepsilon' \to 0$, then $A \to \infty$, we deduce that for some deterministic sequence α'_n going to 0,

$$\Delta_n^{-r/2} \sup_i \mathbb{E}_{i-1}^n \left(\left| |\Delta_i^n X|^r - |\Delta_i^n X'|^r \right| \right) \ \leq \ \alpha'_n. \tag{62}$$

We know that the result holds when X is continuous: the processes $V'(r,q,\Delta_n)$ defined by (21) with X substituted with X' converge in probability, locally uniformly in time, to the process $A(qr)$ (this holds even when $r \geq 2$). Therefore it is enough to prove that

$$\mathbb{E}\left(\sup_{s\leq t} |V(q,r,\Delta_n)_s - V'(q,r,\Delta_n)_s|\right) \to 0. \tag{63}$$

The left side of (63) at time t is smaller than $\sum_{i=1}^{[t/\Delta_n]} E(\zeta_i^n)$, where

$$\zeta_i^n = \frac{\Delta_n^{1-qr/2}}{m_r^q}\left|\prod_{j=1}^{q}|\Delta_{i+j-1}^n X|^r - \prod_{j=1}^{q-1}|\Delta_{i+j-1}^n X'|^r\right| = \frac{1}{m_r^q}\sum_{l=1}^{q}\zeta_i^n(l),$$

$$\zeta_i^n(l) = \Delta_n^{1-rq/2}\prod_{j=1}^{l-1}|\Delta_{i+j-1}^n X|^r\ \left||\Delta_{i+l-1}^n X|^r - |\Delta_{i+l-1}^n X'|^r\right|\ \prod_{j=l+1}^{q}|\Delta_{i+j-1}^n X'|^r,$$

where an empty product is set to be 1. Taking successive conditional expectations, and using (43) and (62), we readily obtain that $E(\zeta_i^n(l)) \leq K\alpha'_n\Delta_n$. Then obviously $\sum_{i=1}^{[t/\Delta_n]} E(\zeta_i^n(l)) \leq Kt\alpha'_n \to 0$ for each $l = 1, \cdots, q$, hence (63). □

Lemma 2. *Under Assumption 4 and if*

$$\widehat{D}_t'^n := \frac{1}{k_n\Delta_n}\sum_{i=1}^{[t/\Delta_n]}|\Delta_i^n X|^p \sum_{j\in I_{n,t}(i)}(\delta_j^n)^2 \xrightarrow{\mathbb{P}} D(p)_t, \tag{64}$$

and also, when X is continuous,

$$\widehat{D}_t''^n := \frac{1}{k_n\Delta_n^{p/2}}\sum_{i=1}^{[t/\Delta_n]}|\delta_i^n|^p \sum_{j\in I_{n,t}(i)}(\delta_j^n)^2 \xrightarrow{\mathbb{P}} m_p\, A(p+2)_t, \tag{65}$$

then (26) holds, as well as (27) when X is continuous.

Proof. In (27) both members are increasing in t and the limit is continuous in t, so it is enough to prove the convergence separately for each t. To unify the proof of the two results, we write $u_n = 1$ and $D = D(p)$ if X jumps, and $u_n = \Delta_n^{1-p/2}$ and $D = m_p A(p+2)$ if X is continuous, in which case we also set

$$\widehat{D}_t'^n := \frac{u_n}{k_n\Delta_n}\sum_{i=1}^{[t/\Delta_n]}|\Delta_i^n X|^p \sum_{j\in I_{n,t}(i)}(\delta_j^n)^2.$$

With this notation, (26) and (27) amounts to $u_n\widehat{D}(p,\Delta_n)_t \xrightarrow{\mathbb{P}} D_t$, and in a first step we show that this is equivalent to $\widehat{D}_t'^n \xrightarrow{\mathbb{P}} D_t$.

Apply (46) with $p = 2$ and $x = \delta_i^n$ and $y = \Delta_i^n X - \delta_i^n$ and $\varepsilon = \alpha\Delta_n^\varpi$, and use Cauchy-Schwarz and Bienaymé-Chebyshev and (43), and also (45) with $\eta = \eta_n = \alpha\Delta_n^\varpi$ and $\varepsilon = \eta_n^{1/2}$, to get after some simple calculations that, with $K = K_\alpha$: :

$$\frac{1}{\Delta_n}\,\mathbb{E}\left(|(\theta_i^n)^2 - (\delta_i^n)^2| \mid \mathcal{F}_{(i-1)\Delta_n}\right)$$

$$\leq \widehat{\Gamma}_n := K\left(\Delta_n + \Delta_n^\varpi + \Gamma(\eta_n^{1/2}) + \Delta_n + \Delta_n^{\varpi/2} + \sqrt{\Gamma(\eta_n^{1/2})}\right) \;\to\; 0, \tag{66}$$

where the final convergence follows from $\eta_n \to 0$, hence $\Gamma(\eta_n^{1/2}) \to 0$ as well. Observe that $u_n\widehat{D}(p,\Delta_n)$ is the same as $\widehat{D}'^n$, with δ_j^n substituted with θ_i^n. Hence the difference $u_n\widehat{D}(p,\Delta_n)_t - \widehat{D}_t'^n$ is the sum of less than $2k_n[t/\Delta_n]$ terms (strictly less, because of the border effects at 0 and $[t/\Delta_n]$), each one being smaller than $\frac{u_n}{k_n\Delta_n}\,|\Delta_i^n X|^p|\,|(\theta_j^n)^2 - (\delta_j^n)^2|$, for some $i \neq j$. Then, by taking two successive conditional expectations and using (43) and (66), we see that the expectation of such a term is smaller than $K_p\Delta_n\widehat{\Gamma}_n/k_n$ in both cases (X continuous or not). Therefore $E(|u_n\widehat{D}(p,\Delta_n)_t - \widehat{D}_t'^n|) \leq K_p t\widehat{\Gamma}_n$. Since $\widehat{\Gamma}_n \to 0$, the claim of this step is complete.

So far we have proved that (64) implies (26). For proving that (65) implies (27) when X is continuous, it remains to show that in this case $E(|\widehat{D}_t'^n - \widehat{D}_t''^n|) \to 0$. Exactly as above, $\widehat{D}_t'^n - \widehat{D}_t''^n$ is the sum of less than $2k_n[t/\Delta_n]$ terms, each one being smaller than $\zeta_{i,j}^n = \frac{u_n}{k_n\Delta_n}\,\Big||\Delta_i^n X|^p| - |\delta_j^n|^p\Big|(\delta_j^n)^2$, for some $i \neq j$. Since $\Big||x+y|^p - |x|^p\Big| \leq K_p(|x|^{p-1}|y| + |y|^p)$, and since $X = X'$ because X is continuous, we deduce from (43) and by taking two successive conditional expectations that $E(\zeta_{i,j}^n) \leq K_p u_n \Delta_n^{1/2+p/2}/k_n$, hence $E(|\widehat{D}_t'^n - \widehat{D}_t''^n|) \leq K_p t\sqrt{\Delta_n}$, and we are done. □

Proof of (26) under Assumption 4. Step 1) For any $\rho \in (0,1)$ we set $\psi_\rho(x) = 1 \wedge (2 - |x|/\rho)^+$ (a continuous function with values in $[0,1]$, equal to 1 if $|x| \leq \rho$ and to 0 if $|x| \geq 2\rho$), and we introduce two increasing processes:

$$Y(\rho)_t^n = \frac{1}{k_n\Delta_n}\sum_{i=1}^{[t/\Delta_n]} \psi_\rho(\Delta_i^n X)|\Delta_i^n X|^p \sum_{j\in I_{n,t}(i)} (\delta_j^n)^2, \qquad Z(\rho)_t^n = \widehat{D}'^n_t - Y(\rho)'^n_t.$$

By the previous lemma we need to prove (64), and for this it is obviously enough to show the following three properties, for some suitable processes $Z(\rho)$:

$$\lim_{\rho\to 0}\ \limsup_n\ \mathbb{E}(Y(\rho)_t^n) \;=\; 0, \tag{67}$$

$$\rho \in (0,1),\ n \to \infty \quad\Rightarrow\quad Z(\rho)_t^n \xrightarrow{\mathbb{P}} Z(\rho)_t, \tag{68}$$

$$\rho \to 0 \quad\Rightarrow\quad Z(\rho)_t \xrightarrow{\mathbb{P}} D_t. \tag{69}$$

Step 2) By singling out the cases $2|x| > |y|$ and $2|x| \leq |y|$, we check that (recall $\rho < 1$):

$$\psi_\rho(x+y)|x+y|^p \leq K_p(|x|^p + (y^2 \wedge \rho^2)).$$

Using this with $x = \delta_i^n$ and $y = \Delta_i^n X - \delta_i^n$, plus (43) and (45) with $\eta = \rho$ and $\varepsilon = \sqrt{\rho}$, we obtain

$$\frac{1}{\Delta_n}\,\mathbb{E}_{i-1}^n(\psi_\rho(\Delta_i^n X)|\Delta_i^n X|^p) \leq \Gamma'_n(\rho) := K_p(\Delta_n^{p/2-1} + \rho + \Gamma(\sqrt{\rho}) + \Delta_n\rho^{-1}).$$

Now, $Y(\rho)_t^n$ is the sum of less than $2k_n[t/\Delta_n]$ terms, all smaller than $\frac{1}{k_n\Delta_n}\,|\Delta_i^n X|^p\,|\delta_j^n|^2$ for some $i \neq j$. By taking two successive conditional expectations, as in the previous lemma, and by using (43) and the above, we see that the expectation of such a term is smaller than $K_p\Delta_n\Gamma'_n(\rho)/k_n$. Thus $E(Y(\rho)_t^n) \leq Kt\gamma'_n(\rho)$, and since obviously $\lim_{\rho\to 0}\ \limsup_n\ \Gamma'_n(\rho) = 0$ we obtain (67).

Step 3) Now we define the process $Z(\rho)$. Let us call $T_q(\rho)$ for $q = 1, 2, \cdots$ the successive jump times of the Poisson process $\mu([0,t] \times \{x : \gamma(x) > \rho/2\})$, and set

$$Z(\rho)_t = \sum_{q:T_q(\rho)\leq t} |\Delta X_{T_q(\rho)}|^p\,(1 - \psi_\rho(\Delta X_{T_p(\rho)}))\ (\sigma^2_{T_q(\rho)-} + \sigma^2_{T_q(\rho)}).$$

For all $\omega \in \Omega$, $q \geq 1$, $\rho' \in (0,\rho)$ there is q' such that $T_q(\rho)(\omega) = T_{q'}(\rho')(\omega)$, whereas $1 - \psi_\rho$ increases to the indicator of $R\backslash\{0\}$. Thus $Z(\rho)_t(\omega) \uparrow D(p)_t(\omega)$, and we have (69).

Step 4) It remains to prove (68). Fix $\rho \in (0,1)$ and write $T_q = T_q(\rho)$. Recall that for u different from all T_q's, we have $|\Delta X_u| \leq \rho/2$. Hence, for each ω and each $t > 0$, we have the following properties for all n large enough: there is no T_q in $(0, k_n\Delta_n]$, nor in $(t-(k_n+1)\Delta_n, t]$; there is at most one T_q in an interval $((i-1)\Delta_n, i\Delta_n]$ with $i\Delta_n \leq t$, and if this is not the case we have $\psi_\rho(\Delta_i^n X) = 1$. Hence for n large enough we have

$$Z(\rho)_t = \sum_{q:\ k_n\Delta_n < T_q \leq t-(k_n+1)\Delta_n} \zeta_q^n,$$

where

$$\zeta_q^n = \frac{1}{k_n\Delta_n}\,|\Delta^n_{i(n,q)}X|^p(1 - \psi_\rho(\Delta^n_{i(n,q)}X)) \sum_{j\in I'(n,q)} (\delta_j^n)^2,$$

and $i(n,q) = \inf(i : i\Delta_n \geq T_q)$ and $I'(n,q) = \{j : j \neq i(n,q), |j - i(n,q)| \leq k_n\}$.

To get (68) it is enough that $\zeta_q^n \xrightarrow{\mathbb{P}} |\Delta X_{T_q}|^p\,(1-\psi_\rho(\Delta X_{T_q}))\ (\sigma^2_{T_q-} + \sigma^2_{T_q})$ for any q. Since $|\Delta^n_{i(n,q)}X|^p(1 - \psi_\rho(\Delta^n_{i(n,q)}X)) \to |\Delta X_{T_q}|^p\,(1 - \psi_\rho(\Delta X_{T_p}))$ pointwise, so it remains to prove that

$$(70) \qquad \frac{1}{k_n\Delta_n} \sum_{j\in I'_-(n,q)} (\delta_j^n)^2 \ \xrightarrow{\mathbb{P}}\ \sigma^2_{T_q-}, \quad \frac{1}{k_n\Delta_n} \sum_{j\in I'_+(n,q)} (\delta_j^n)^2 \ \xrightarrow{\mathbb{P}}\ \sigma^2_{T_q},$$

where $I'_-(n,q)$ and $I'_+(n,q)$ are the subsets of $I'(n,q)$ consisting in those j smaller, respectively bigger, than $i(n,q)$. We write

$$U^n_q \;=\; \frac{1}{k_n\Delta_n}\sum_{j\in I'_-(n,q)}(\Delta^n_j W)^2, \quad s^n_q = \inf_{u\in[T_q-k_n\Delta_n,T_q)}\sigma^2_u, \quad S^n_q = \sup_{u\in[T_q-k_n\Delta_n,T_q)}\sigma^2_u.$$

On the one hand, both s^n_q and S^n_q converge as $n\to\infty$ to $\sigma^2_{T_q-}$, because $k_n\Delta_n\to 0$. On the other hand the left side of the first expression in (69) is in between the two quantities $s^n_q U^n_q$ and $S^n_q U^n_q$. Moreover the variables $\Delta^n_i W$ are i.i.d. $N(0,\Delta_n)$, so U^n_q is distributed as $U'_n = \frac{1}{k_n}\sum_{i=1}^{k_n} V_i$ where the V_i's are i.i.d. $N(0,1)$, and by the usual law of large numbers we have $U'_n\to 1$ a.s., hence $U^n_q \xrightarrow{\mathbb{P}} 1$: these facts entails the first part of (70), and the second part is proved in the same way. □

Proof of (27) under Assumption 4. Step 1) We have to prove that when X is continuous, then (65) holds. We have

$$\Delta_n^{-p/2}\sum_{i=1}^{[t/\Delta_n]}|\delta^n_i|^{p+2} \;\xrightarrow{\mathbb{P}}\; m_{p+2}\,A(p+2)_t,$$

which can be deduced from (10) exactly as $\widehat{D}'^n_t - \widehat{D}''^n_t \xrightarrow{\mathbb{P}} 0$ in the end of the proof of Lemma 2. Therefore it is enough to prove that

$$Y^n_t := \frac{1}{k_n\Delta_n^{p/2}}\sum_{i=1}^{[t/\Delta_n]}\sum_{j\in I_{n,t}(i)}\zeta^n_{i,j} \;\xrightarrow{\mathbb{P}}\; 0, \quad \text{where} \quad \zeta^n_{i,j} = m_{p+2}|\delta^n_i|^p|\delta^n_j|^2 - m_p|\delta^n_i|^{p+2}.$$

Let $T(n,i) = (i-k_n-1)^+\Delta_n$. When $j>i$ we have (recall σ_t is bounded):

$$\begin{aligned}|\mathbb{E}^n_{i-1}(\zeta^n_{i,j})| \;&=\; \Big|\mathbb{E}^n_{i-1}\Big(m_{p+2}|\sigma_{(i-1)\Delta_n}|^p(\sigma^2_{(j-1)\Delta_n}-\sigma^2_{(i-1)\Delta_n})|\Delta^n_i W|^p|\Delta^n_j W|^2\\ &\qquad + |\sigma_{(i-1)\Delta_n}|^{p+2}(m_{p+2}|\Delta^n_i W|^p|\Delta^n_j W|^2 - m_p|\Delta^n_i W|^{p+2})\Big)\Big|\\ &\le K_p\Delta_n\mathbb{E}^n_{i-1}(|\sigma^2_{(j-1)\Delta_n}-\sigma^2_{(i-1)\Delta_n}|\;|\Delta^n_i W|^p|).\end{aligned}$$

When $i-k_n\le j<i$ we have

$$\begin{aligned}|\mathbb{E}(\zeta^n_{i,j}\mid\mathcal{F}_{(j-1)\Delta_n})| &= \Big|\mathbb{E}\Big(m_{p+2}|\sigma_{(j-1)\Delta_n}|^2(|\sigma_{(i-1)\Delta_n}|^p-|\sigma_{(j-1)\Delta_n}|^p)|\Delta^n_i W|^p|\Delta^n_j W|^2\\ &\qquad + |\sigma_{(j-1)\Delta_n}|^{p+2}(m_{p+2}|\Delta^n_i W|^p|\Delta^n_j W|^2 - m_p|\Delta^n_i W|^{p+2})\mid\mathcal{F}_{(j-1)\Delta_n}\Big)\Big|\\ &\le\; K_p\Delta_n^{p/2}\,\mathbb{E}\Big(\Big||\sigma_{(i-1)\Delta_n}|^p-|\sigma_{(j-1)\Delta_n}|^p\Big|\;|\Delta^n_j W|^2|\mid\mathcal{F}_{(j-1)\Delta_n}\Big).\end{aligned}$$

Therefore, since $\Big|\sigma_{(i-1)\Delta_n}|^q-|\sigma_{(j-1)\Delta_n}|^q\Big| \le K_q|\sigma_{(i-1)\Delta_n}-\sigma_{(j-1)\Delta_n}|$ for any $q\ge 1$, we deduce from (43) and Cauchy-Schwarz and the two estimates above that

$$j\in I_{n,t}(i) \quad\Rightarrow\quad |\mathbb{E}(\zeta^n_{i,j}\mid\mathcal{F}_{T(n,i)})|\le K_p\Delta_n^{3/2+p/2}. \tag{71}$$

Moreover, as a trivial consequence of (43) again, we get

$$j \in I_{n,t}(i) \quad \Rightarrow \quad \mathbb{E}(|\zeta^n_{i,j}|^2 \mid \mathcal{F}_{T(n,i)})| \leq K_p \Delta_n^{2+p}. \tag{72}$$

Now we set

$$\eta^n_i = \frac{1}{k_n \Delta_n^{p/2}} \sum_{j \in I_{n,t}} \zeta^n_{i,j}, \qquad Z_n = \sum_{i=1}^{[t/\Delta_n]} \mathbb{E}(\eta^n_i \mid \mathcal{F}_{T(n,i)}), \qquad Z'_n = Y^n_t - Z_n.$$

On the one hand (71) yields $|Z_n| \leq K_p t \sqrt{\Delta_n} \to 0$. On the other hand $Z'_n = \sum_{i=1}^{[t/\Delta_n]} (\eta^n_i - E(\eta^n_i \mid \mathcal{F}_{T(n,i)})$, whereas η^n_i is $\mathcal{F}_{T(n,i+2k_n+1)}$-measurable. Hence (72) gives

$$\begin{aligned}
\mathbb{E}(Z'^2_n) &= \sum_{i,i':1 \leq i,i' \leq [t/\Delta_n]} \mathbb{E}\Big(\Big(\eta^n_i - \mathbb{E}(\eta^n_i \mid \mathcal{F}_{T(n,i)}\Big)\Big(\eta^n_{i'} - \mathbb{E}(\eta^n_{i'} \mid \mathcal{F}_{T(n,i')}\Big)\Big) \\
&\leq \sum_{i,i':1 \leq i \leq [t/\Delta_n], |i-i'| \leq 2k_n+1} \mathbb{E}\Big(\Big(\eta^n_i - \mathbb{E}(\eta^n_i \mid \mathcal{F}_{T(n,i)}\Big)\Big(\eta^n_{i'} - \mathbb{E}(\eta^n_{i'} \mid \mathcal{F}_{T(n,i')}\Big)\Big) \\
&\leq \sum_{i,i':1 \leq i \leq [t/\Delta_n], |i-i'| \leq 2k_n+1} \mathbb{E}\Big(|\eta^n_i \ \eta^n_{i'}|\Big) \\
&\leq (4k_n+3)[t/\Delta_n] \ K_p \Delta_n^2 \ \leq \ K_p t (k_n \Delta_n),
\end{aligned}$$

which goes to 0 by virtue of (22). It then follows that $Y^n_t \xrightarrow{\mathbb{P}} 0$, and we are done. □

8.5. *Proof of Theorem 6.*

Proof. a) When X is continuous the variables $U_n = (\widehat{V}^c_n)^{-1/2} \ (\widehat{S}(p,k,\Delta_n)_t - k^{p/2-1})$ converge stably in law to $N(0,1)$ by Theorem 5-(b), for both choices of $\widehat{V}^c_n$. In particular if c^c_n is given by (34) we have that $\mathbb{P}(C^c_{n,t}) = \mathbb{P}(U_n < -z_\alpha) \to \alpha$, and we even have more, because of the stable convergence in law: namely, for any measurable subset B of Ω, then

$$\mathbb{P}(C^c_{n,t} \cap B) \ = \ \mathbb{P}(\{U_n < -z_\alpha\} \cap B) \ \to \ \alpha \, \mathbb{P}(B). \tag{73}$$

Now, this is not quite enough to prove (a), since it may happens that X is not continuous but the observed path is continuous on $[0,t]$. To deal with this case, we denote by X' the process given by (1) with the same b_t and σ_t as X, and with δ identically 0. Put a "prime" for the variables defined on the basis of X', writing for example $\widehat{V}'^c_n$ or $\widehat{S}(p,k,\Delta_n)'_t$ or $C'^c_{n,t}$. Then (73) applied with the continuous process X' shows that $\mathbb{P}(C'^c_{n,t} \cap B) \to \alpha \, \mathbb{P}(B)$ for any $B \in \mathcal{F}$. However, on the set Ω^c_t where X is continuous between 0 and t we have $\widehat{V}'^c_n = \widehat{V}^c_n$ and $\widehat{S}(p,k,\Delta_n)'_t = \widehat{S}(p,k,\Delta_n)_t$, hence $C'^c_{n,t} \cap \Omega^c_t = C^c_{n,t} \cap \Omega^c_t$. Therefore $\mathbb{P}(C^c_{n,t} \cap \Omega^c_t) \to \alpha \, \mathbb{P}(\Omega^c_t)$, and (i) follows.

b) Finally assume $\mathbb{P}(\Omega^j_t) > 0$. Then Theorem 1 yields that $\widehat{S}(p,k,\Delta_n)_t \xrightarrow{\mathbb{P}} 1$ on the set Ω^j_t. On the other hand if we use the version (29) for $\widehat{V}^c_n$ we deduce from

(24) that $c_n^c \xrightarrow{\mathbb{P}} k^{p/2-1} > 1$, whereas if we use the version (30) we have $\widehat{V}_n^c/\Delta_n \xrightarrow{\mathbb{P}} M(p,k)\, A(2p)_t/A(p)_t^2$, hence again $c_n^c \xrightarrow{\mathbb{P}} k^{p/2-1} > 1$: so the result is obvious. □

8.6. *Proof of Theorem 7.*

Proof. Relative to the proof of the previous theorem, we need to make a few changes. First we replace (a) of Theorem 6 by two statements (a) and (b) here: the case (b) corresponds to the situation where the limit in Theorem 5-(a) is normal, and so this is similar to Theorem 6-(a); the case (a) here corresponds to a non-normal limit with variance 1, and for this limit we cannot evaluate exactly the quantiles and we rely upon Chebyshev's inequality: this is why we only give a bound on the level in this situation. Apart from these changes, the proof for the level is the same.

For (c) we suppose that (b, σ, δ) are such that $\mathbb{P}(\Omega_t^c) > 0$, and we denote by X' the process with the same (b, σ) as X and with $\delta = 0$. Then similar to the previous proof, we see that on the set Ω_t^c we have $\widehat{S}(p, k, \Delta_n)_t \xrightarrow{\mathbb{P}} k^{p/2-1}$ and $\widehat{V}_n^j \xrightarrow{\mathbb{P}} 0$ (use (22) and (27) for the latter), hence the result. □

REFERENCES

Aït-Sahalia, Y. (2002). Telling from discrete data whether the underlying continuous-time model is a diffusion. *Journal of Finance* **57** 2075–2112.

Aït-Sahalia, Y. (2004). Disentangling diffusion from jumps. *Journal of Financial Economics* **74** 487–528.

Aït-Sahalia, Y. and Jacod, J. (2006). Volatility estimators for discretely sampled Lévy processes. *Annals of Statistics, forthcoming* .

Andersen, T., Bollerslev, T. and Diebold, F. X. (2003). Some like it smooth, and some like it rough. Tech. rep., Northwestern University.

Barndorff-Nielsen, O., Graversen, S., Jacod, J., Podolskij, M. and Shephard, N. (2006a). A central limit theorem for realised bipower variations of continuous semimartingales. In *From Stochastic Calculus to Mathematical Finance, The Shiryaev Festschrift* (Y. Kabanov, R. Liptser and J. Stoyanov, eds.). Springer Verlag, Berlin, 33–69.

Barndorff-Nielsen, O. E. and Shephard, N. (2004). Power and bipower variation with stochastic volatility and jumps (with discussion). *Journal of Financial Econometrics* **2** 1–48.

Barndorff-Nielsen, O. E. and Shephard, N. (2006). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics* **4** 1–30.

Barndorff-Nielsen, O. E., Shephard, N. and Winkel, M. (2006b). Limit theorems for multipower variation in the presence of jumps. *Stochastic Processes and Applications* **116** 796–806.

Carr, P. and Wu, L. (2003). What type of process underlies options? A simple robust test. *Journal of Finance* **58** 2581–2610.

Huang, X. and Tauchen, G. (2006). The relative contribution of jumps to total price variance. *Journal of Financial Econometrics* **4** 456–499.

Jacod, J. (2006). Asymptotic properties of realized power variations and related functionals of semimartingales. Tech. rep., Université de Paris VI.

Jacod, J. and Protter, P. (1998). Asymptotic error distributions for the Euler method for stochastic differential equations. *Annals of Probability* **26** 267–307.

Jacod, J. and Shiryaev, A. N. (2003). *Limit Theorems for Stochastic Processes*. 2nd ed. Springer-Verlag, New York.

Jiang, G. J. and Oomen, R. C. (2005). A new test for jumps in asset prices. Tech. rep., The University of Warwick, Warwick Business School.

Lee, S. and Mykland, P. A. (2005). Jumps in financial markets: a new nonparametric test and jump clustering. Tech. rep., University of Chicago.

Lepingle, D. (1976). La variation d'ordre p des semi-martingales. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **36** 295–316.

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

Mancini, C. (2004). Estimating the integrated volatility in stochastic volatility models with Lévy type jumps. Tech. rep., Universita di Firenze.

Rényi, A. (1963). On stable sequences of events. *Sankyā Series A* **25** 293–302.

Woerner, J. H. (2006a). Analyzing the fine structure of continuous-time stochastic processes. Tech. rep., University of Göttingen.

Woerner, J. H. (2006b). Power and multipower variation: Inference for high frequency data. In *Stochastic Finance* (A. Shiryaev, M. do Rosário Grosshino, P. Oliviera and M. Esquivel, eds.). Springer Verlag, Berlin, 264 – 276.

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