Journal of Econometrics 168 (2012) 207–222

Contents lists available at SciVerse ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Testing for jumps in noisy high frequency data[☆]

Yacine Aït-Sahalia[a,b,∗], Jean Jacod[c], Jia Li[a]

[a] Department of Economics, Bendheim Center for Finance, Princeton University, Princeton, NJ 08544-1021, United States

[b] NBER, Princeton, NJ 08544-1021, United States

[c] Institut de Mathématiques, de Jussieu, CNRS – UMR 7586, Université Pierre et Marie Curie, Paris 6, 75 013 Paris, France

ARTICLE INFO

Article history:
Received 13 April 2010
Received in revised form
17 November 2011
Accepted 15 December 2011
Available online 22 December 2011

JEL classification:
C22

Keywords:
Semimartingale
Testing for jumps
High frequency data
Market microstructure noise
Pre-averaging

ABSTRACT

This paper proposes a robustification of the test statistic of Aït-Sahalia and Jacod (2009b) for the presence of market microstructure noise in high frequency data, based on the pre-averaging method of Jacod et al. (2010). We show that the robustified statistic restores the test's discriminating power between jumps and no jumps despite the presence of market microstructure noise in the data.

1. Introduction

The recent availability of observations on financial returns at increasingly higher frequencies has prompted the development of methodologies designed to test the specification of suitable models for these data. Motivated both by mathematical tractability and the need to avoid introducing arbitrage opportunities in the model, semimartingales are often employed.

We focus here on testing for the presence of jumps in a discretely observed semimartingale, which has been among the first issue to be considered in the literature. Existing tests for jumps include Aït-Sahalia (2002) (based on the transition function of the process), Carr and Wu (2003) (based on short dated options), Barndorff-Nielsen and Shephard (2004); Huang and Tauchen (2005) and Andersen et al. (2007) (based on bipower variations), Jiang and Oomen (2008) (based on a swap variance), Lee and Mykland (2008) and Lee and Hannig (2010) (based on detecting large increments) and Aït-Sahalia and Jacod (2009b) (based on power variations sampled at different frequencies).

When implemented on high frequency data, as most of them are designed to be, these tests are confronted by the presence of market microstructure noise. Furthermore, that measurement error tends to grow in proportion of the observed increments of the process as the sampling frequency increases, which distinguishes this problem from the classical measurement error problem in statistics.

This issue has received a fair amount of attention in the recent literature, but focused on the base case of quadratic variation estimation. Considering only methods that are robust to the simplest forms of market microstructure noise, there are currently four main approaches to quadratic variation estimation: maximum likelihood estimation (Aït-Sahalia et al., 2005; Xiu, 2010), linear combination of realized volatilities obtained by subsampling (Zhang et al., 2005; Zhang, 2006), linear combination of autocovariances (Barndorff-Nielsen et al., 2008) and pre-averaging (Jacod et al., 2009, 2010). The simplest forms of noise include additive errors and rounding, and combinations thereof. Robust estimators are available as long as the noise is sufficiently "smooth"; a pure rounding error is not. Attempting to generalize the type of noise allowed to an "unsmooth" setting raises a different set of issues that are beyond the scope of this paper (see Li and Mykland (2007) for a discussion).

All the tests developed so far for jumps assume away the presence of noise in high frequency data. In this paper, we examine

☆ Financial support from the NSF under grant SES-0850533 is gratefully acknowledged. We are also grateful to the Editor, Associate Editor and two referees for comments and suggestions.

∗ Corresponding author at: Department of Economics, Bendheim Center for Finance, Princeton University, Princeton, NJ 08544-1021, United States. Tel.: +1 609 258 4015; fax: +1 609 258 0719.

E-mail addresses: yacine@princeton.edu (Y. Aït-Sahalia), jean.jacod@upmc.fr (J. Jacod), jiali@princeton.edu (J. Li).

0304-4076/$ – see front matter
doi:10.1016/j.jeconom.2011.12.004

the possibility of robustifying one of these tests for jumps, that of Aït-Sahalia and Jacod (2009b), using the pre-averaging method. The test, whose asymptotic properties were originally derived without allowing for the possibility of noise, is based on comparing variations of power greater than 2, at two different frequencies, and taking their ratio. If jumps are present, the two variations converge asymptotically as $\Delta_n \to 0$ to the same limit, which is simply the sum of the pth power of the jumps recorded between 0 and T; as a result their ratio converges to 1. On the other hand, if no jumps are present, the sum of the pth power of the jumps recorded between 0 and T is zero, and both variations then converge to 0. They do so at a rate that depends on the sampling interval Δ_n and so the ratio will pick up the difference between the two sampling frequencies: if the two sampling intervals are Δ_n and $k\Delta_n$, then the limit of the ratio will be $k^{p/2-1}$. Therefore, without noise, the test statistic has two sharply distinct limits depending upon whether jumps are present or not.

In the presence of noise, on the other hand, the theoretical limits of the statistic become respectively $1/k$ and $1/k^{1/2}$ in the two polar cases of additive noise and noise due to rounding error, and become so irrespectively of the presence or absence of jumps (see Aït-Sahalia and Jacod (2009b)). As a result, when either type of noise dominates, the test statistic loses its intended effectiveness at discriminating between presence and absence of jumps.

In this paper, we construct a robustified version of the statistic, using the pre-averaging approach, and show that this robustification restores the ability of the test to discriminate between jumps and no jumps, despite the presence of the noise. The results are nonparametric in nature, are valid for almost unrestricted semimartingales, and allow for a symmetric treatment of the two null hypotheses specifying either presence or absence of jumps.

The paper is organized as follows. Section 2 presents the model's setting and assumptions. Section 3 presents the test statistic, studies its properties when noise is taken into account, describes its robustification by pre-averaging and derives its asymptotic properties after robustification. Sections 4 and 5 report the results of simulations and of an empirical application to high frequency stock returns data. Section 6 concludes, while proofs are in the Appendix.

2. The model

2.1. The underlying process

We consider a one-dimensional underlying process $X = (X_t)_{t\geq 0}$, sampled at regularly spaced discrete times $i\Delta_n$ over a fixed time interval $[0, T]$, with a time lag which asymptotically goes to 0. In typical financial econometrics applications, X represents the logarithm of an asset price. The basic assumption is that X is an *Itô semimartingale* on a filtered space $(\Omega^{(0)}, \mathcal{F}^{(0)}, (\mathcal{F}_t^{(0)}), \mathbb{P}^{(0)})$, which means that it can be written as

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + (\delta 1_{\{|\delta|\leq 1\}}) \star (\underline{\mu} - \underline{\nu})_t + (\delta 1_{\{|\delta|>1\}}) \star \underline{\mu}_t, \tag{2.1}$$

where W is a Brownian motion, $\underline{\mu}$ is a Poisson random measure on $\mathbb{R}_+ \times E$ and its compensator is $\underline{\nu}(dt, dz) = dt \otimes \lambda(dz)$ where $(E, \mathcal{E})$ is an auxiliary space and λ is a σ-finite measure (all these are defined on the filtered space above and we refer for example to Jacod and Shiryaev (2003) for all unexplained terms). We further assume:

Assumption 1. (a) The process (b_t) is optional and locally bounded;

(b) The process (σ_t) is càdlàg (i.e., right-continuous with left limits) and adapted;

(c) The function δ is predictable, and there is a bounded function γ in $\mathbb{L}^2(E, \mathcal{E}, \lambda)$ such that the process $\sup_{z\in E}(|\delta(\omega^{(0)}, t, z)| \wedge 1)/\gamma(z)$ is locally bounded;

(d) We have almost surely $\int_0^t \sigma_s^2\, ds > 0$ for all $t > 0$.

In particular, when X is continuous, it has the form

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s. \tag{2.2}$$

In this case, we will sometimes need a stronger assumption putting some further structure on the stochastic volatility process, namely:

Assumption 2. We have Assumption 1 and σ_t is also an Itô semimartingale which can be written as

$$\sigma_t = \sigma_0 + \int_0^t \tilde{b}_s ds + \int_0^t \tilde{\sigma}_s dW_s + M_t + \sum_{s\leq t} \Delta\sigma_s\, 1_{\{|\Delta\sigma_s|>v\}}, \tag{2.3}$$

where M is a local martingale orthogonal to W and with bounded jumps and $\langle M, M\rangle_t = \int_0^t a_s ds$, and the compensator of $\sum_{s\leq t} 1_{\{|\Delta\sigma_s|>v\}}$ is $\int_0^t a'_s ds$, and where $\tilde{b}_t$, a_t, and a'_t are optional locally bounded processes, whereas the adapted processes b_t and $\tilde{\sigma}_t$ are left-continuous with right limits.

Overall, these assumptions are standard and fairly unrestrictive. They do not significantly restrict the essential aspects of the process, allowing for stochastic volatility, jumps of finite or infinite activity, all manners of dependence between the characteristics of the process, etc. Of course, they do exclude some examples such as fractional Brownian motion or models without a continuous martingale part, given (d) in Assumption 1. We need the latter requirement to avoid degenerate limiting theorems under the null hypothesis where no jumps are present.

2.2. The noise

The main purpose of this paper is to test for the presence of jumps when the process X is observed with an error: instead of X_t we now observe

$$Z_t = X_t + \epsilon_t. \tag{2.4}$$

Of course, the observation error ϵ_t comes into the picture only at those observation times $t = i\Delta_n$, but it is convenient to have it defined for all t. We assume that the observation error is, conditionally on the process X, mean zero and mutually independent. Note however that the ϵ_t's are not necessarily unconditionally independent (the independence is only conditional on X). The assumption we will make on the noise term allows for an additive error of the white noise type, but also for noise involving rounding since the assumptions allow the noise ϵ_t to depend on X_t, or in fact even on the whole past of X up to time t.

Mathematically speaking, this can be formalized as follows: for each $t \geq 0$, we have a transition probability $Q_t(\omega^{(0)}, dz)$ from $(\Omega^{(0)}, \mathcal{F}_t^{(0)})$ into $\mathbb{R}$. The space $\Omega^{(1)} = \mathbb{R}^{[0,\infty)}$ is endowed with the product Borel σ-field $\mathcal{F}^{(1)}$ and the "canonical process" $(\epsilon_t : t \geq 0)$ and the probability $\mathbb{Q}(\omega^{(0)}, d\omega^{(1)})$ which is the product $\otimes_{t\geq 0} Q_t(\omega^{(0)}, \cdot)$. We introduce the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ and the filtration $(\mathcal{G}_t)$ as follows:

$$\left.\begin{array}{l} \Omega = \Omega^{(0)} \times \Omega^{(1)}, \qquad \mathcal{F} = \mathcal{F}^{(0)} \otimes \mathcal{F}^{(1)}, \\ \mathcal{F}_t = \mathcal{F}_t^{(0)} \otimes \sigma(\epsilon_s : s \in [0, t)), \\ \quad \mathcal{G}_t = \mathcal{F}^{(0)} \otimes \sigma(\epsilon_s : s \in [0, t)), \\ \mathbb{P}(d\omega^{(0)}, d\omega^{(1)}) = \mathbb{P}^{(0)}(d\omega^{(0)})\mathbb{Q}(\omega^{(0)}, d\omega^{(1)}). \end{array}\right\} \tag{2.5}$$

In terms of notation, $\omega^{(0)}$ stands for the process X while $\omega^{(1)}$ stands for the noise ϵ, and similarly for other quantities with superscripts (0) or (1) respectively. Any variable or process which is defined on either $\Omega^{(0)}$ or $\Omega^{(1)}$ can be considered in the usual way as a variable or a process on Ω. Note that X is still a semimartingale, with the same decomposition (2.1), on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, and W and μ are a Wiener process and a Poisson random measure on this extended space as well. We make the following assumption on the noise, which matches that employed in the development of pre-averaging estimators by (Jacod et al., 2009, 2010):

Assumption 3. For each $q > 0$ there is a sequence of $(\mathcal{F}_t^{(0)})$-stopping times $(T_{q,n})_{n\geq 1}$ increasing to ∞, such that $\int Q_t(\omega^{(0)}, dz)\, |z|^q \leq n$ whenever $t < T_{q,n}(\omega^{(0)})$. We write

$$\beta(q)_t(\omega^{(0)}) = \int Q_t(\omega^{(0)}, dz)\, z^q, \qquad \alpha_t = \sqrt{\beta(2)_t}, \tag{2.6}$$

and we assume that the processes α and $\beta(3)$ are càdlàg, and that

$$\beta(1) \equiv 0. \tag{2.7}$$

We assume moments of all orders for the noise, although only moments up to $2p$ (where $p \geq 4$ is the power chosen below) should be finite; in practice, this is a mild restriction. The regularity properties of the paths of α and $\beta(3)$ are not needed all the time, but this is again a weak requirement. The really strong requirement in this assumption is (2.7), in conjunction with the conditional independence of the noise at different times. Note however that whereas the noise at different times is independent, conditionally on $\mathcal{F}$, it is *not* unconditionally independent.

2.3. The hypotheses to be tested

The problem we wish to solve is the same as in Aït-Sahalia and Jacod (2009b), namely to decide in which of the two complementary sets the observed path falls (the time horizon T is fixed):

$$\left.\begin{array}{l}\Omega_T^c = \{\omega^{(0)}: t \mapsto X_t\left(\omega^{(0)}\right) \text{ is continuous on } [0,T]\}\\ \Omega_T^j = \{\omega^{(0)}: t \mapsto X_t\left(\omega^{(0)}\right) \text{ is discontinuous on } [0,T]\}.\end{array}\right\} \tag{2.8}$$

We wish to do so in both cases where either Ω_T^c or Ω_T^j plays the role of the null hypothesis.

Recall that, when the null hypothesis is either $\Omega_0 = \Omega_T^c$ or $\Omega_0 = \Omega_T^j$, the asymptotic level of a sequence C_n of critical regions (C_n is the critical rejection region at stage n, which is measurable w.r.t. $\sigma(Z_{i\Delta_n}: i = 0, \ldots, [t/\Delta_n])$) is

$$\alpha = \sup\left\{\limsup_n \mathbb{P}(C_n \mid A): A \in \mathcal{F}, A \subset \Omega_0, \mathbb{P}(A) > 0\right\}. \tag{2.9}$$

As for the asymptotic power, when the alternative is the complement Ω_1 of Ω_0, it was defined as $\beta = \inf\{\liminf_n \mathbb{P}(C_n \mid A): A \in \mathcal{F}, A \subset \bar{\Omega}_0, \mathbb{P}(A) > 0\}$ in earlier papers such as for example Aït-Sahalia and Jacod (2009b), where for any set B, we denote by $\bar{B}$ its complement. However, the property $\beta = 1$ is in fact equivalent to saying

$$\mathbb{P}(\bar{C}_n \cap \Omega_1) = 0, \tag{2.10}$$

and a sequence of tests achieving this should perhaps be called *alternative-consistent*.

3. The test statistic

3.1. The case with no noise

We briefly recall the results of Aït-Sahalia and Jacod (2009b) and the construction of their base test statistic for jumps. For any process Y and any integer $i \geq 1$ and real $p > 0$ we write

$$\Delta_i^n Y = Y_{i\Delta_n} - Y_{(i-1)\Delta_n}, \qquad B(Y, p, \Delta_n)_t = \sum_{i=1}^{[t/\Delta_n]} |\Delta_i^n Y|^p. \tag{3.1}$$

(The notation $\Delta_i^n Y$ is not meant to denote an nth difference, but rather the ith increment among n such discrete increments of the process Y.) We take an integer $k \geq 2$ and consider the test statistic

$$S_J(p, k, \Delta_n)_n = \frac{B(X, p, k\Delta_n)_T}{B(X, p, \Delta_n)_T}. \tag{3.2}$$

When there is no noise, this is computable from the data, and the two tests (with the two possible null hypotheses Ω_T^c and Ω_T^j) are based upon the following asymptotic behavior, when $p > 2$:

$$S_J(p, k, \Delta_n)_n \xrightarrow{\mathbb{P}} \begin{cases} 1 & \text{on the set } \Omega_T^j \\ k^{p/2-1} & \text{on the set } \Omega_T^c. \end{cases} \tag{3.3}$$

Moreover a Central Limit Theorem allows one to specify, for any given $\alpha \in (0,1)$, a sequence $v_n \to 0$ (depending on the observations at stage n) such that the asymptotic level of the rejection region $C_n = \{S_J(p, k, \Delta_n)_n > 1 + v_n\}$ is α (and the asymptotic power is 1), when the null is Ω_T^j, and under Assumption 1. Analogously, under Assumption 2 and when the null is Ω_T^c we can proceed similarly for the rejection regions $C_n = \{S_J(p, k, \Delta_n)_n < k^{p/2-1} - v_n\}$.

In the presence of noise, the theoretical limits of the statistic $S_J(p, k, \Delta_n)_n$ become respectively $1/k$ and $1/k^{1/2}$ in the two polar cases of additive noise and noise due to rounding error on the whole set Ω (see Aït-Sahalia and Jacod (2009b)). As a result, when either type of noise dominates, the test statistic loses its intended effectiveness at discriminating between presence and absence of jumps.

3.2. Pre-averaging

These results suggest the need for a noise-robust test statistic. In order to construct such a robustified test statistic, we start with some notations in order to construct a pre-averaged version of the test statistic, using the tools introduced by (Jacod et al., 2009, 2010).

First, in order to define the pre-averaging window, we choose a sequence of integers k_n satisfying for some $\theta > 0$:

$$k_n\sqrt{\Delta_n} = \theta + o(\Delta_n^{1/4}). \tag{3.4}$$

Next, pre-averaging involves weighting the observations in the pre-averaging window, and for this purpose we use weight functions g on $\mathbb{R}$, satisfying

$$\left.\begin{array}{l} g \text{ is continuous, piecewise } C^1 \\ \qquad \text{with a piecewise Lipschitz derivative } g', \\ s \notin (0,1) \Rightarrow g(s) = 0, \qquad \displaystyle\int g(s)^2 ds > 0, \end{array}\right\} \tag{3.5}$$

and with which we associate the quantities (where $p \in (0, \infty)$ and $i \in \mathbb{Z}$):

$$\left.\begin{array}{ll} g_i^n = g(i/k_n), & g_i'^n = g_i^n - g_{i-1}^n, \\ \bar{g}(p) = \displaystyle\int |g(s)|^p\, ds, & \bar{g}'(p) = \displaystyle\int |g'(s)|^p\, ds. \end{array}\right\} \tag{3.6}$$

With any process $Y = (Y_t)_{t\geq 0}$ we associate the following random variables

$$\overline{Y}(g)^n_i = \sum_{j=1}^{k_n-1} g^n_j \Delta^n_{i+j} Y, \qquad \widehat{Y}(g)^n_i = \sum_{j=1}^{k_n} (g'^n_j \Delta^n_{i+j} Y)^2 \tag{3.7}$$

and processes

$$V(Y, g, q, r)^n_t = \sum_{i=0}^{[t/\Delta_n]-k_n} |\overline{Y}(g)^n_i|^q |\widehat{Y}(g)^n_i|^r \tag{3.8}$$

which – implicitly – depend on the two sequences Δ_n and k_n. When $Y = Z$ is the noisy process, the variable $\overline{Y}(g)^n_i$ plays the role of the simple increment $\Delta^n_i X$ when there is no noise; the variable $\widehat{Y}(g)^n_i$ is a kind of estimator of the local variance of the noise around time $i\Delta_n$; the process $V(Y, g, q, 0)^n_t$ is thus the counterpart of the process $B(Y, q, \Delta_n)_t$ in (3.1), whereas the processes $V(Y, g, q, r)^n_t$ for $r > 0$ serve to remove the bias involved by the use of the process with $r = 0$, and due to the noise.

Letting $p \geq 4$ be an *even integer*, we define $(\rho(p)_j)_{j=0,\ldots,p/2}$ as the unique numbers solving the following triangular system of linear equations:

$$\left.\begin{aligned} &\rho(p)_0 = 1, \\ &\sum_{l=0}^{j} 2^l m_{2j-2l} C^{p-2j}_{p-2l} \rho(p)_l = 0, \quad j = 1, 2, \ldots, p/2, \end{aligned}\right\} \tag{3.9}$$

where m_r denotes the rth absolute moment of the law $\mathcal{N}(0, 1)$. These could be explicitly computed, and for example when $p = 4$ (the case used in practice),

$$\rho(4)_0 = 1, \qquad \rho(4)_1 = -3, \qquad \rho(4)_2 = 0.75. \tag{3.10}$$

Finally, we also choose a sequence k_n satisfying (3.4), and we set for any process Y

$$\overline{V}(Y, g, p)^n_t = \sum_{l=0}^{p/2} \rho(p)_l V(Y, g, p-2l, l)^n_t. \tag{3.11}$$

These are going to be used as the robustified versions of the pth power variation estimators.

3.3. First order asymptotic properties

We are now ready to introduce our robustified test statistics for the presence or absence of jumps. For this, we fix *an even integer $p \geq 4$ and two weight functions g and h*. For simplicity we set

$$\gamma = \frac{\bar{g}(2)}{\bar{h}(2)}, \qquad \gamma' = \frac{\bar{g}(p)}{\bar{h}(p)}, \qquad \gamma'' = \frac{\gamma^{p/2}}{\gamma'}, \tag{3.12}$$

and we assume that $\gamma'' > 1$ (if γ'' were smaller than 1 one could always interchange g and h to get $\gamma'' > 1$, whereas if γ'' were equal to 1 the tests below would not separate our two hypotheses).

We then propose as our robustified test statistic

$$S_{RJ}(g, h, p)_n = \frac{\overline{V}(Z, g, p)^n_T}{\gamma' \, \overline{V}(Z, h, p)^n_T}. \tag{3.13}$$

In other words, up to the adjustment factor γ', the robustified test statistic is constructed by replacing the standard power variations B in (3.1) with their robustified versions $\overline{V}$ in (3.11).

We now derive the limiting behavior in probability of the test statistics $S_{RJ}(g, h; p)_n$ given above:

Theorem 1. *Under Assumptions 1 and 3, we have*

$$S_{RJ}(g, h; p)_n \xrightarrow{\mathbb{P}} \begin{cases} 1 & \text{on the set } \Omega^j_T \\ \gamma'' & \text{on the set } \Omega^c_T. \end{cases} \tag{3.14}$$

So we get an asymptotic behavior for the robustified test statistic $S_{RJ}(g, h; p)_n$ which, in the presence of noise, restores the discriminating ability of the statistic $S_J(p, k, \Delta_n)_n$ when there is no noise. For the appropriate selection of the weighting functions in pre-averaging, we can if desired obtain exactly the same limits for S_{RJ} as those of S_J (1 if there are jumps and $k^{p/2-1}$ if there are none). Indeed, if we take $h(s) = g(sk)$ for some $k > 1$ then $\bar{g}(q) = k\bar{h}(q)$ for any $q > 0$, so $\gamma'' = k^{p/2-1}$ and we retrieve exactly the same limits as given in (3.3), except that k does not need to be an integer here.

3.4. Second order properties

In order to use the statistic $S_{RJ}(g, h, p)_n$ in a test, we need a Central Limit Theorem associated with the convergence in (3.14), and there are of course two very distinct behaviors on the two sets Ω^c_T and Ω^j_T.

3.4.1. When the null hypothesis is absence of jumps

We start with what happens on the set Ω^c_T. We choose a sequence u_n as follows:

$$u_n = \alpha \, \Delta^{\varpi}_n, \quad \text{where } \alpha > 0, \ \frac{1}{12} < \varpi < \frac{1}{4}. \tag{3.15}$$

These will serve as truncation levels for estimating asymptotic variances. Next we introduce a number of constants, depending on the weight functions g and h. Although conceptually straightforward, these are quite complicated to write, although simple to compute numerically, and they will be motivated in the Appendix.

First we write, for any two functions ϕ and ψ and any integers $w \geq 1$ and $w' \in \{0, \ldots, 2w\}$:

$$\left.\begin{aligned} &a(\phi, \psi)_t = \int_{1\vee t}^{1+1\wedge t} \phi(u-1)\, \psi(u-t)\, du \\ &a'(\phi, \psi; w, w')_t = \sum_{r=0}^{[w'/2]} C^{2r}_{w'}\, m_{2r} m_{2w-2r}\, a(\phi, \phi)^{w-w'}_1 \\ &a(\phi, \psi)^{w'-2r}_t \left(a(\phi, \phi)_1\, a(\psi, \psi)_1 - a(\phi, \psi)^2_t\right)^r. \end{aligned}\right\} \tag{3.16}$$

These will be used when ϕ and ψ are either the weight functions g and h, or their derivatives; observe that $a(g, g)_1 = \bar{g}(2)$ and $a(g', g')_1 = \bar{g}'(2)$.

Finally we write for $w \in \mathbb{N}$:

$$\left.\begin{aligned} &A(g, h; w)_t = \sum_{l,l'\in\{0,\ldots,p/2\},\, l+l'\leq p-w} \\ &\qquad \sum_{w'=(2w-p+2l)^+}^{(2w)\wedge(p-2l')} \rho(p)_l\, \rho(p)_{l'}\, C^{2w-w'}_{p-2l}\, C^{w'}_{p-2l'} \\ &(2\bar{g}'(2))^l (2\bar{h}'(2))^{l'}\, a'(g, h; w, w')_t \\ &\qquad a'(g', h'; p-l-l'-w, p-2l'-w')_t \\ &A'(g, h; w) = \int_0^2 A(g, h; w)_t dt \\ &\qquad - 2m^2_p\, \bar{g}(2)^{p/2}\, \bar{h}(2)^{p/2}\, 1_{\{w=p\}}. \end{aligned}\right\} \tag{3.17}$$

We also complete the notation (3.8) with a truncated version:

$$V^*(Y, g, q, r)^n_t = \sum_{i=0}^{[t/\Delta_n]-k_n} |\overline{Y}(g)^n_i|^q\, 1_{\{|\overline{Y}(g)^n_i| \leq u_n\}}\, |\widehat{Y}(g)^n_i|^r. \tag{3.18}$$

We end this series of notations by setting, for any weight function ϕ:

$$M^*(g, h, \phi)^n_t = \Delta^{1-p/2}_n \sum_{w=0}^{p} \frac{\theta\, A'(g, h; w)}{m_{2w} 2^{p-w} \bar{\phi}(2)^w \bar{\phi}'(2)^{p-w}} \times \sum_{l=0}^{w} \rho(2w)_l\, V^*(Z, \phi, 2w-2l, p+l-w)^n_t. \tag{3.19}$$

Theorem 2. *Suppose that Assumptions 2 and 3 hold.*

(a) *The variables*

$$\frac{1}{\Delta_n^{1/4}}\left(S_{RJ}(g,h,p)_n-\gamma''\right)$$

converge stably in law, in restriction to the set Ω_T^c, toward a random variable defined on an extension of the probability space $(\Omega,\mathcal{F},\mathbb{P})$ and which, conditionally on $\mathcal{F}$, is a centered Gaussian variable, the variance of which we denote by $\Sigma_{RJ}^c(g,h,p,\theta)$ (an $\mathcal{F}$-measurable positive variable).

(b) *For any choice of the auxiliary weight function ϕ, the variables*

$$\Sigma_{RJ,n}^c=\frac{M^*(g,g,\phi;p)_T^n-2\gamma^{p/2}M^*(g,h,\phi,p)_T^n+\gamma^pM^*(h,h,\phi;p)_T^n}{\left(\Delta_n^{1-p/4}\,\overline{V}(Z,g,p)_T^n/\gamma''\right)^2} \tag{3.20}$$

converge in probability to the $\mathcal{F}$-conditional variance $\Sigma_{RJ}^c(g,h,p,\theta)$, in restriction to the set Ω_T^c.

3.4.2. When the null hypothesis is presence of jumps

We now turn to the behavior on the set Ω_T^j. For this, we choose another sequence k_n' of integers satisfying

$$k_n'/k_n\to\infty,\qquad k_n'\Delta_n\to 0. \tag{3.21}$$

Recall also the truncation level u_n of (3.15). We choose an arbitrary weight function ϕ (it may be g or h, or some other), and we consider the variables

$$\left.\begin{aligned}\eta(\phi,0)_i^n&=\frac{1}{k_n\,k_n'\,\Delta_n}\sum_{j=1}^{k_n'}\left((\overline{Z}(\phi)_{i+j}^n)^2-\frac{1}{2}\widehat{Z}(\phi)_{i+j}^n\right)\\&\quad\times 1_{\{|\overline{Z}(\phi)_{i+j}^n|\le u_n\}}\\ \eta(\phi,1)_i^n&=\frac{1}{k_n\,k_n'\,\Delta_n}\sum_{j=1}^{k_n'}\widehat{Z}(\phi)_{i+j}^n\,1_{\{|\overline{Z}(\phi)_{i+j}^n|\le u_n\}}.\end{aligned}\right\} \tag{3.22}$$

This allows to define four processes (below, m is either 0 or 1) as follows:

$$\left.\begin{aligned}N(\phi,m,-)_t^n&=\frac{1}{k_n}\sum_{i=k_n+k_n'}^{[t/\Delta_n]-k_n}\left(\overline{Z}(\phi)_i^n\right)^{2p-2}\\&\quad\eta(\phi,m)_{i-k_n-k_n'}^n\\ N(\phi,m,+)_t^n&=\frac{1}{k_n}\sum_{i=0}^{[t/\Delta_n]-2k_n-k_n'+1}\left(\overline{Z}(\phi)_i^n\right)^{2p-2}\\&\quad\times\eta(\phi,m)_{i+k_n-1}^n.\end{aligned}\right\} \tag{3.23}$$

On the other hand, we introduce the numbers (recall that p is fixed):

$$\left.\begin{aligned}\Gamma(-,g)_t&=\int_t^1 g(s)^{p-1}g(s-t)ds,\\ \Gamma'(-,g)_t&=\int_t^1 g(s)^{p-1}g'(s-t)ds\\ \Gamma(+,g)_t&=\int_0^{1-t} g(s)^{p-1}g(s+t)ds,\\ \Gamma'(+,g)_t&=\int_0^{1-t} g(s)^{p-1}g'(s+t)ds.\\ \Psi_\pm(g,h)&=\int_0^1\Gamma(\pm,g)_t\,\Gamma(\pm,h)_t\,dt,\\ \Psi'_\pm(g,h)&=\int_0^1\Gamma'(\pm,g)_t\,\Gamma'(\pm,h)_t dt\\ \Psi_\pm&=\Psi_\pm(g,g)+\gamma'^2\Psi_\pm(h,h)-2\gamma'\Psi_\pm(g,h),\\ \Psi'_\pm&=\Psi'_\pm(g,g)+\gamma'^2\Psi'_\pm(h,h)-2\gamma'\Psi'_\pm(g,h).\end{aligned}\right\} \tag{3.24}$$

Theorem 3. *Suppose Assumptions 1 and 3 hold.*

(a) The variables

$$\frac{1}{\Delta_n^{1/4}}\left(S_{RJ}(g,h,p)_n-1\right)$$

converge stably in law, in restriction to the set Ω_T^j, toward a random variable defined on an extension of the probability space $(\Omega,\mathcal{F},\mathbb{P})$ and which, conditionally on $\mathcal{F}$, is a centered Gaussian variable, the variance of which we denote by $\Sigma_{RJ}^j(g,h,p,\theta)$.

(b) For any choice of the auxiliary weight function ϕ, the variables

$$\begin{aligned}\Sigma_{RJ,n}^j&=\frac{\theta p^2k_n^2}{\left(\overline{V}(Z,g,p)_T^n\right)^2}\\&\quad\times\left(\frac{\Psi_-N(\phi,0,-)_T^n+\Psi_+N(\phi,0,+)_T^n}{\bar{\phi}(2)\bar{\phi}(2p-2)}\right.\\&\quad+\left.\frac{\Psi'_-N(\phi,1,-)_T^n+\Psi'_+N(\phi,1,+)_T^n}{2\bar{\phi}'(2)\bar{\phi}(2p-2)}\right)\end{aligned} \tag{3.25}$$

converge in probability to the $\mathcal{F}$-conditional variance $\Sigma_{RJ}^j(g,h,p,\theta)$, in restriction to the set Ω_T^j.

Remark 1. As we can see from the previous formulas, we have the two weight functions g and h used for our basic test statistics, and another one ϕ used to compute the estimators for the variance. We could also choose different sequences k_n, with different values θ in (3.4), for defining $S_{RJ}(g,h,p)_n$ and for defining $V^*(Z,\phi,r,s)_t^n$ in (3.19) or for the processes in (3.23): we thus have a lot of flexibility, hence also a lot of parameters to tune. In practice one takes $\phi=g$ or $\phi=h$, with the same sequence k_n all the time.

Remark 2. There is even more flexibility for the estimators of the conditional variance $\Sigma_{RJ,n}^j$ than what is mentioned in the previous remark. For example in (3.22) one could truncate also $\widehat{Z}(\phi)_{i+j}^n$ at the level u_n, or leave out any truncation. In (3.23) we could truncate $\overline{Z}(g)_i^n$ from below, that is replacing $|\overline{Z}(g)_i^n|^{2p-2}$ with $|\overline{Z}(g)_i^n|^{2p-2}\,1_{\{|\overline{Z}(g)_i^n|>u_n\}}$. The proofs are exactly the same.

3.5. The two tests

We start with the case where the null hypothesis is "no jump", that is Ω_T^c. As before, the two weight functions g and h are given, as well as the even integer $p\ge 4$ (typically, $p=4$ and $h(t)=g(kt)$ for some $k>1$; recall that in any case $\gamma''>1$). We use the statistics $S_{RJ}(g,h,p)_n$ given by (3.13).

With the aim of constructing a test with a given asymptotic level $\alpha\in(0,1)$, we denote by z_α the corresponding quantile of $\mathcal{N}(0,1)$, that is the positive number such that $\mathbb{P}(U\ge z_\alpha)=\alpha$, where U is $\mathcal{N}(0,1)$.

Theorem 4. *We assume Assumptions 2 and 3, and we set*

$$C_n^c=\left\{S_{RJ}(g,h,p)_n<\gamma''-z_\alpha\,\Delta_n^{1/4}\sqrt{\Sigma_{RJ,n}^c}\right\}, \tag{3.26}$$

where $\Sigma_{RJ,n}^c$ is given by (3.20). Then the asymptotic level of the critical regions (3.26) for testing the null hypothesis "no jump" equals α, and those tests are consistent for the alternative.

In the second case, we set the null hypothesis to be that "there are jumps", that is Ω_T^j.

Theorem 5. *We assume Assumptions 1 and 3, and we let $\Sigma^j_{RJ,n}$ be defined by (3.25). Then the asymptotic level of the critical regions defined by*

$$C^j_n = \left\{ S_{RJ}(g,h,p)_n > 1 + z_\alpha \Delta_n^{1/4} \sqrt{\Sigma^j_{RJ,n}} \right\} \tag{3.27}$$

for testing the null hypothesis "there are jumps" equals α, and those tests are consistent for the alternative.

4. Simulation results

We now examine the validity of the asymptotic theory above in a simulation setting designed to approximate the constraints faced in a typical real life application of jump tests to financial data. The log price Z_t is generated according to the following model:

$$Z_t = X^c_t + J_t + \epsilon_t,$$
$$X^c_t = X_0 + \int_0^t \sigma_s dW_s, \qquad \sigma_t = v_t^{1/2},$$
$$v_t = \kappa(\beta - v_t)\,dt + \gamma v_t^{1/2} dB_t, \qquad E[dW_t dB_t] = \rho dt,$$
$$\text{where } \beta^{1/2} = 0.4, \qquad \gamma = 0.5, \qquad \kappa = 5, \qquad \rho = -0.5,$$
$$X_0 = \log(25).$$

Here, X^c_t is the continuous part with instantaneous volatility σ_t, J_t is a pure jump process and ϵ_t is the additive noise. The drift in X^c_t is excluded because it plays little role in the high frequency setting. Parameters governing the stochastic volatility process are calibrated according to the estimates in Aït-Sahalia and Kimmel (2007). We use an observation length of $T = 5$ days, with each day consisting of 6.5 h, and sample the continuous-time process at every 5 s. There are 5000 simulations in each experiment.

We consider four settings for the additive noise:

$$\epsilon_t = \begin{cases} 0 \quad \text{(No noise)} \\ 2\sigma_t \Delta_n^{1/2} \epsilon^A_t \quad \text{(Gaussian noise)} \\ 2\sigma_t \Delta_n^{1/2} \epsilon^B_t \Big/ \sqrt{\dfrac{df}{df-2}} \quad (T\text{-distributed noise}) \\ 2\sigma_t \Delta_n^{1/2} \left(\epsilon^A_t + \epsilon^B_t \Big/ \sqrt{\dfrac{df}{df-2}} \right) \quad (\text{Gaussian-}T \text{ mixture noise}), \end{cases} \tag{4.1}$$

where ϵ^A_t and ϵ^B_t are mutually independent i.i.d. draws from an $\mathcal{N}(0,1)$ distribution and a t-distribution with degree of freedom $df = 2.5$, respectively. The instantaneous standard deviations of the Gaussian noise and the t-distributed noise are 2 times that of the diffusive increment, i.e., $\sigma_t \Delta_n^{1/2}$.[1] This experimental design allows temporal heteroskedasticity and dependence in ϵ_t. The t-distributed noise is introduced to capture the large bouncebacks commonly observed in transaction data. Fig. 1 plots one realization of the returns of the process $X^c_t + \epsilon_t$. Comparing the log returns of the clean price (top panel) with those of the noisy price, we note that in our design, the microstructure noise clearly dominates the diffusive increment. Moreover, with the t-distributed noise present, one could observe many large returns even in the absence of jumps. Naively equating large increments with jumps could lead to very misleading results.

We simulate the jump process J_t from a centered symmetric α-stable process with activity index 0.5, 1, 1.5 or 1.75. To compare results across various activity levels, we scale J_t so that in each realization, the realized quadratic variation of J_t is fixed at $\beta T/3$, that is one third of the mean quadratic variation of the continuous part. We plot one realization of the returns of J_t in Fig. 2. We note that our design allows for a wide spectrum of jump behavior. When the jump activity level is 0.5 (top panel), the jump process is dominated by a few big jumps, featuring the situation with "infrequent big jumps"; when the activity is high (bottom panel), jumps have relatively similar sizes, featuring the situation with "many small jumps". As the activity index approaches 2, that is the index of the Brownian motion, jumps become more difficult to detect.

Besides the additive noise, we also consider rounding on the price level. That is the situation in which we only observe the price level rounded to the nearest multiple of the tick size κ, where we set $\kappa = 0.01, 0.02$ or 0.03. In practice, the rounding threshold is typically 0.01 dollar for stock prices and $1/32 \approx 0.03$ dollar for bonds. We note that when the additive noise $\epsilon_t \equiv 0$, we are in the situation with pure rounding. Although our asymptotic theory does not cover the pure rounding case, it is included here as a robustness check.[2]

Throughout the simulations and the empirical study, we fix $p = 4$ and weight functions $g(x) = (0.5 - |x - 0.5|)^+$, $h(x) = g(2x)$ and $\phi = g$. In particular, when there is no jump, the unstandardized statistic $S_{RJ}(g,h,p)_n$ converges to $\gamma'' = 2$. We set the averaging windows $k_n = 80$, 100, or 120, and $k'_n = 3k_n$. The truncation level u_n in (3.15) is fixed at $C\left(\bar{V}(Z,g,2)^n_T/T\right)^{1/2} \Delta_n^{0.49}$, with $C = 4$, 5 or 6. In the discussion below, we often focus on the case with $k_n = 100$ and $C = 5$ to save space.

Fig. 3 plots the finite-sample distribution of the unstandardized test statistic $S_{RJ}(g,h,4)_n$. When there is no jump (solid line), the distribution of $S_{RJ}(g,h,4)_n$ is correctly centered at 2 (sample mean = 2.00), as predicted by the asymptotic theory. However, when there are jumps, the distribution of $S_{RJ}(g,h,4)_n$ tends to be biased toward 2 instead of centered at the theoretical limit 1. The bias increases as the jump activity approaches 2: When the jump activity is 0.5 (dash-circled line), the sample mean of $S_{RJ}(g,h,4)_n$ in the simulation is 1.07; when the jump activity is 1.75 (dashed line), the sample mean is 1.53.

Table 1 reports the finite-sample size of the test $H_0: \omega \in \Omega^c_T$ vs. $H_1: \omega \in \Omega^j_T$ at the 5% level. First consider the case without rounding (upper-left panel). We find that the rejection rate is close to the nominal level in all settings. Importantly, the result is quite robust across various types of additive noise as desired. The result is also robust to moderate perturbations on the averaging window k_n and the truncation level C. Now consider the case with rounding. The rejection rate changes surprisingly little when the price is rounded at 0.01 or 0.02 (upper-right and lower-left panels), even with pure rounding. In the case with coarse pure rounding, i.e. no additive noise and rounding at 0.03, we find that the test slightly over-rejects. However, the over-rejection disappears as soon as the additive noise is included. Overall, the simulation evidence suggests that the test has good size control in a variety of practically relevant settings.

Table 2 reports the finite-sample power of this test at the same nominal level. To save space, we only report results with $k_n = 100$, noting that results with $k_n = 80$ or 120 show a similar pattern. Again, we find that the rejection rate is fairly robust across various types of additive noise. The rejection rate is almost 100% when the jump activity is 0.5, and decreases as the activity level approaches 2. The rejection rate also decreases in the truncation level C. However, this finding does not imply that one should mechanically use a small C to increase the power of the test. When C is too small, the truncation not only eliminates jumps,

[1] In the simulations, we truncate the t-distributed noise ϵ^B_t at $50\sqrt{df/(df-2)}$ in order to verify Assumption 3. Hence, the instantaneous standard deviation of the noise is slightly lower than $2\sigma_t \Delta_n^{1/2}$.

[2] We thank an anonymous referee for the suggestion that rounding should be considered in our simulations.

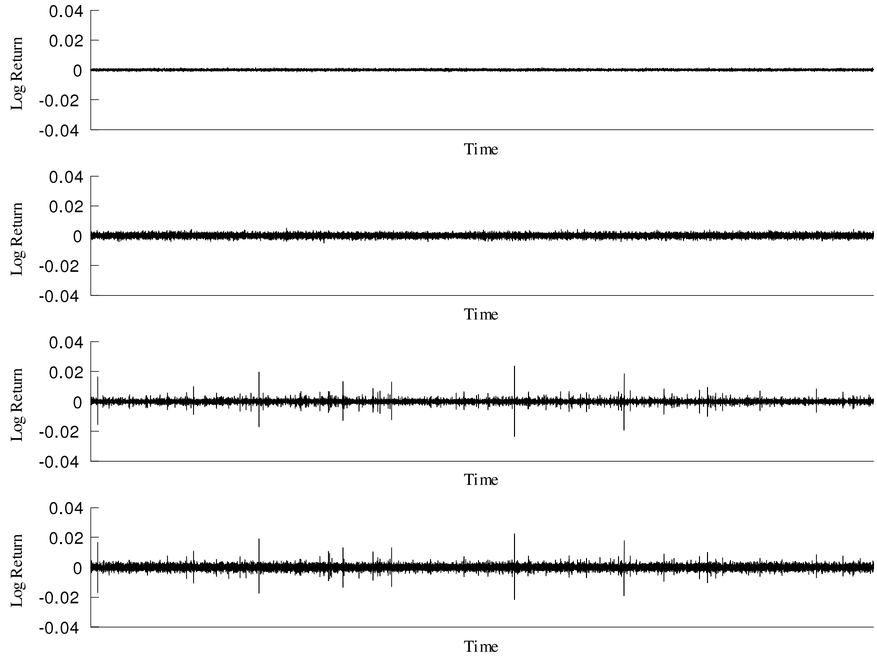

Fig. 1. A realization of returns of the noisy continuous part of the log price ($X_t^c + \epsilon_t$). From top to bottom, the continuous part is contaminated with no noise, Gaussian noise, T-distributed noise, and Gaussian-T mixture noise.

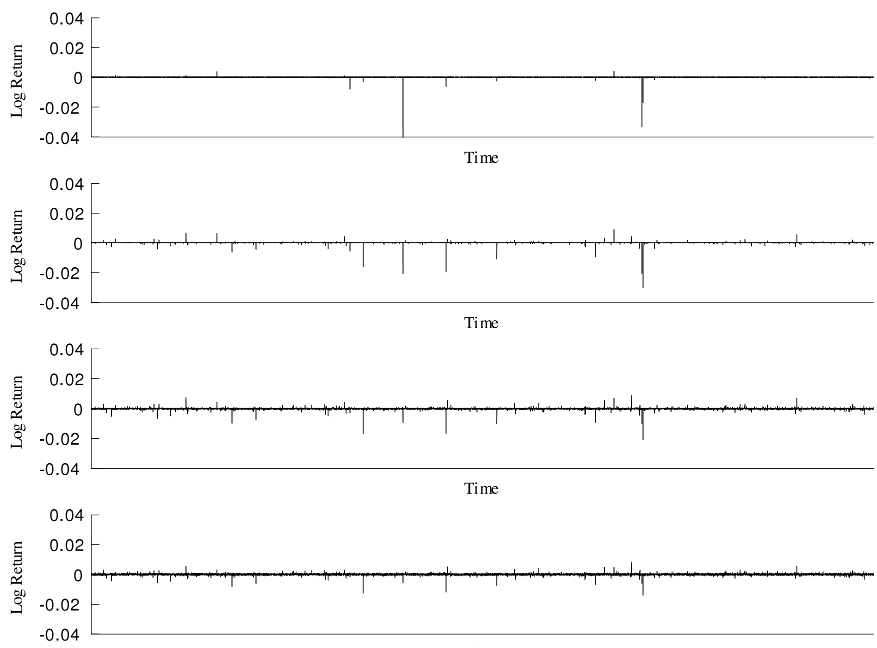

Fig. 2. A realization of returns of the jump process J_t. From top to bottom, the activity index of the jump process is 0.5, 1, 1.5, 1.75. The jump processes are scaled so that its realized quadratic variation is fixed at $\beta T/3$.

but also the diffusive increments, and thus leads to size distortion.[3] Finally, we note that the result is not sensitive to rounding.

Fig. 4 compares finite-sample distributions of the standardized test statistic with the limiting $\mathcal{N}(0, 1)$ distribution. In the absence of jumps (solid line), the finite-sample distribution is fairly close to the $\mathcal{N}(0, 1)$ distribution (shaded area) as predicted by Theorem 2.

[3] In the same setting as in Table 1, we also consider the case with $C = 2$ and $C = 3$ (not reported here). We find mild over-rejection when $C = 3$. However, when $C = 2$, the rejection rate is about 23% at the 5% nominal level.

We now turn to the testing problem for $H_0: \omega \in \Omega_T^j$ vs. $H_1: \omega \in \Omega_T^c$. Since the simulation result is somewhat pessimistic, we only briefly discuss the result to document the problems. Table 3 shows the finite-sample size of the test. The setting in Panel A is the same as that of Table 2. We find large size distortions especially when the jump activity is high. We further increase the strength of the jump signal in the simulation by rescaling the jump process so that the realized quadratic variation in each simulation is fixed at $3\beta T$. As shown in Panel B of Table 3, in this case, the size distortion become much smaller. Fig. 5 plots finite-sample distributions of the standardized test

Table 1
Finite-sample size (%) of the 5%-level test for the null hypothesis of no jump.

Additive noise	k_n	No rounding			Rounding at 0.01		
		$C = 4$	$C = 5$	$C = 6$	$C = 4$	$C = 5$	$C = 6$
No noise	80	4.8	4.8	4.8	4.9	4.9	4.9
	100	5.7	5.6	5.6	5.6	5.5	5.5
	120	6.3	6.3	6.3	6.2	6.2	6.2
Gaussian	80	5.0	5.0	5.0	5.0	5.0	5.0
	100	5.7	5.7	5.7	5.8	5.7	5.7
	120	6.1	6.1	6.1	6.1	6.1	6.1
T-distributed	80	4.8	4.8	4.8	4.8	4.8	4.8
	100	5.2	5.2	5.2	5.2	5.2	5.2
	120	6.0	6.0	6.0	6.0	6.0	6.0
Gaussian-T mixture	80	4.8	4.8	4.8	4.8	4.8	4.8
	100	5.4	5.4	5.4	5.4	5.4	5.4
	120	6.0	6.0	6.0	6.0	6.0	6.0
		Rounding at 0.02			Rounding at 0.03		
No noise	80	4.8	4.8	4.8	8.3	8.3	8.3
	100	5.5	5.5	5.5	7.6	7.5	7.5
	120	6.1	6.1	6.1	7.4	7.4	7.4
Gaussian	80	4.8	4.8	4.8	5.0	5.0	5.0
	100	5.7	5.6	5.6	5.8	5.8	5.8
	120	6.1	6.1	6.1	6.0	6.0	6.0
T-distributed	80	4.8	4.8	4.8	5.4	5.4	5.4
	100	5.4	5.3	5.3	5.9	5.9	5.9
	120	6.1	6.1	6.1	6.3	6.3	6.3
Gaussian-T mixture	80	4.8	4.8	4.8	4.8	4.8	4.8
	100	5.4	5.4	5.4	5.4	5.4	5.4
	120	6.0	6.0	6.0	5.8	5.8	5.8

Table 2
Finite-sample power (%) of the 5%-level test for the null hypothesis of no jump. In each realization, the quadratic variation of the jump process is fixed to be $\beta T/3$, i.e., one third of the mean quadratic variation of the continuous part. In all cases, we set the averaging window $k_n = 100$.

Additive noise	Jump activity	No rounding			Rounding at 0.03		
		$C = 4$	$C = 5$	$C = 6$	$C = 4$	$C = 5$	$C = 6$
No noise	0.50	99.9	99.9	99.2	100.0	99.9	99.1
	1.00	99.7	98.1	92.3	99.7	98.1	92.4
	1.50	91.3	82.7	68.6	91.4	82.9	69.0
	1.75	67.6	55.9	44.9	68.5	57.1	45.9
Gaussian	0.50	99.9	99.9	99.1	99.9	99.8	99.2
	1.00	99.6	98.0	92.1	99.6	98.0	92.1
	1.50	91.0	82.4	68.3	91.1	82.3	68.4
	1.75	67.1	55.5	44.5	66.9	55.4	44.4
T-distributed	0.50	99.9	99.9	99.1	99.9	99.9	99.1
	1.00	99.6	98.1	92.2	99.6	98.0	92.1
	1.50	91.4	82.4	68.3	91.3	82.5	68.4
	1.75	67.4	55.8	44.9	67.6	55.9	44.8
Gaussian-T mixture	0.50	99.9	99.8	99.1	99.9	99.8	99.0
	1.00	99.5	98.0	92.0	99.5	98.0	92.1
	1.50	91.2	82.0	67.8	91.0	82.0	67.8
	1.75	66.9	55.2	44.3	66.8	55.0	44.2

statistic $(S_{RJ}(g, h, 4)_n - 1)/\Delta_n^{1/4}\sqrt{\Sigma^j_{RJ,n}}$. When the jump signal is weak (Panel A), the finite-sample distributions deviate from the limiting $\mathcal{N}(0, 1)$ distribution to an undesirable extent; when the jump signal gets stronger (Panel B), the quality of the asymptotic approximation becomes better. The size distortion observed here might not be surprising. Testing under the null with jumps is a problem with composite null $\Delta X \neq 0$ against simple alternative $\Delta X = 0$,[4] while testing under the null with no jumps is exactly the opposite. Intuitively, the large size in the former testing problem and the low power in the latter reflect the difficulty in finding small jumps with noisy data. Nonetheless, we consider Theorem 3 as a relevant benchmark for future research on testing and estimation of the fine structure of jumps. Because of the priority given to size control in the frequentist testing paradigm, we suggest testing under the null hypothesis of no jumps in practical applications.

5. Empirical results

We now conduct the test for jumps for each of the 30 Dow Jones Industrial Average (DJIA) stocks and each trading day in 2008; the data source is the TAQ database. Because the composition of the DJIA changes over time, we use the 30 stocks that are the components of the index as of October 29th, 2009. For each trading day in 2008, we collect all transactions from 9:30 am until 4:00 pm, and compute the volume-weighted average of transaction prices at each time stamp for each one of these stocks. We sample in calendar time every 5 s. Each day and stock is treated on its own, so there are 7590 stock-day pairs in total. We use filters to eliminate clear data errors (price set to zero, etc.) as is standard in

[4] Of course, strictly speaking, both the null hypothesis and the alternative are composite, because we leave the instantaneous drift and volatility processes unspecified.

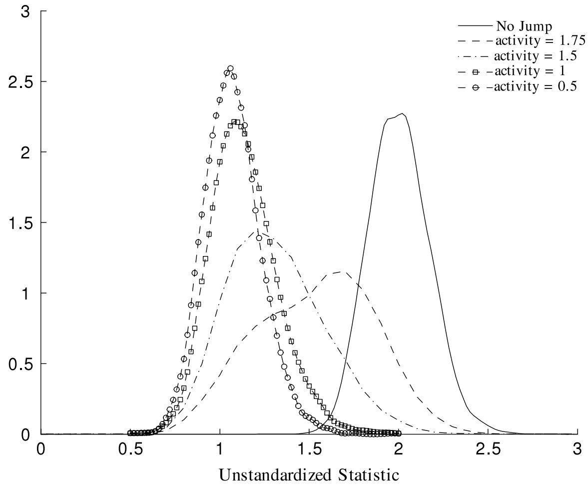

Fig. 3. Finite-sample distributions of the unstandardized test statistic $S_{RJ}(g, h, 4)_n$. We plot the distribution in the continuous case (solid line). In the discontinuous case, the jump process is α-stable with activity 0.5 (dash-circle), 1 (dash-square), 1.5 (dash-dot) and 1.75 (dash). In each experiment, the jump process is scaled so that its realized quadratic variation is $\beta T/3$, i.e., one third of the mean quadratic variation of the continuous part. In all cases, the averaging window is $k_n = 100$ and the efficient price is contaminated with the Gaussian-T mixture noise without rounding.

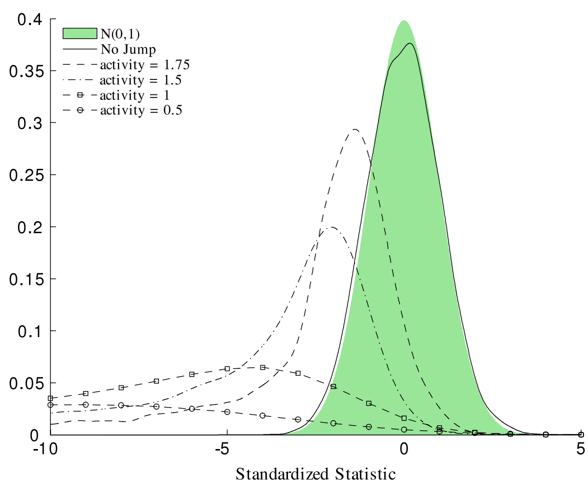

Fig. 4. Finite-sample distributions of the standardized test statistic for testing the null hypothesis with no jump. The standardized statistic is $\left(S_{RJ}(g, h, 4)_n - 2\right) / \Delta_n^{1/4}\sqrt{\Sigma^c_{RJ,n}}$. We plot the distribution in the continuous case (solid line). In the discontinuous case, the jump process is α-stable with activity 0.5 (dash-circle), 1 (dash-square), 1.5 (dash-dot) and 1.75 (dash). In each experiment, the jump process is scaled so that its realized quadratic variation is $\beta T/3$, i.e., one third of the mean quadratic variation of the continuous part. In all cases, the averaging window is $k_n = 100$, the truncation level is $C = 5$, and the efficient price is contaminated with the Gaussian-T mixture noise without rounding.

the empirical market microstructure literature. We use the same tuning parameters as in the simulation.

Table 4 summarizes the results. We find statistical evidence for the presence of jumps. In particular, when the averaging window $k_n = 100$ and the truncation level $C = 5$, we reject the null hypothesis of no jump 32.6% of the time at the 5% nominal level. Fig. 6 shows the empirical distributions of the non-standardized (top) and the standardized (bottom) test statistics. Relative to the limiting distribution under the null hypothesis of no jump (shaded area), the empirical distribution of the standardized test statistic deviates toward the left side, that is, "the jump side". We also find that the distribution of the unstandardized statistic has a unimodal distribution centered around 1.5, instead of having two modes at 1 and 2. In view of the simulation pattern in Fig. 3, our conjecture is that small jumps are often present in the data but we only reject about one third of the time because the test has low power when the jump signal is weak. This conjecture is consistent with the estimates of jump activity in Aït-Sahalia and Jacod (2009a). Noise-robust inference about the fine structure of jumps is a natural step for future research.

Table 3
Finite-sample size (%) of the 5%-level test for the null hypothesis of jumps. In each realization, the realized quadratic variation of the jump process is fixed to be $\beta T/3$ (Panel *A*) or $3\beta T$ (Panel *B*). The additive noise is generated from the Gaussian-T mixture distribution. In the case with rounding, the rounding level is 0.03. In all cases, we set the averaging window $k_n = 100$ and the truncation level $C = 5$.

Additive noise?	Rounding?	Jump activity			
		0.50	1.00	1.50	1.75
Panel A. Realized jump QV =$\beta T/3$.					
No	No	17.2	28.6	57.4	79.3
No	Yes	17.6	28.4	57.2	79.2
Yes	No	17.0	28.5	56.7	78.8
Yes	Yes	17.1	28.2	57.1	79.0
Panel B. Realized jump QV =$3\beta T$.					
No	No	8.7	9.7	15.5	35.5
No	Yes	9.0	9.7	15.4	35.8
Yes	No	8.7	9.6	15.4	36.1
Yes	Yes	8.8	9.7	15.3	35.8

Table 4
Summary statistics of the empirical results for 30 stocks of DJIA in 2008. The summary statistics are computed for all 7590 stock-day pairs. We set the averaging window $k_n = 80$, 100, 120 and the truncation level $C = 4, 5, 6$.

Panel A. Empirical mean of the unstandardized statistic.			
Mean	$k_n = 80$	$k_n = 100$	$k_n = 120$
	1.55	1.53	1.53
Panel B. Empirical mean of the standardized statistic.			
$C = 4$	−1.36	−1.25	−1.16
$C = 5$	−1.16	−1.12	−1.07
$C = 6$	−1.09	−1.08	−1.04
Panel C. Empirical rejection rate (%) for $H_0: \omega \in \Omega^c_T$.			
$C = 4$	39.7	37.2	34.5
$C = 5$	33.5	32.6	30.7
$C = 6$	31.6	31.4	29.7

6. Conclusions

We provide a robustification of the test statistic for jumps of Aït-Sahalia and Jacod (2009b) for the presence of market microstructure noise in high frequency data, based on the pre-averaging method of Jacod et al. (2010). When noise dominates, the base test statistic is no longer able to disentangle between the two situations where the sample path is either continuous or discontinuous. After robustification, the test statistic is once again able to separate these two hypotheses.

Simulation evidence reveals that the robustified test statistic performs well under the null hypothesis of continuous paths. When the null hypothesis includes jumps, its performance is more mixed, since the pre-averaging step tends to affect the fine structure of the sample path. Empirically, we find that the conclusion reached using the non-robust version of the test statistic remains, namely that jumps are likely present in the dataset considered.

Appendix. Proofs

In Appendix A, we prove the theorems in the text, recall some known results in the literature, and state some preliminary results.

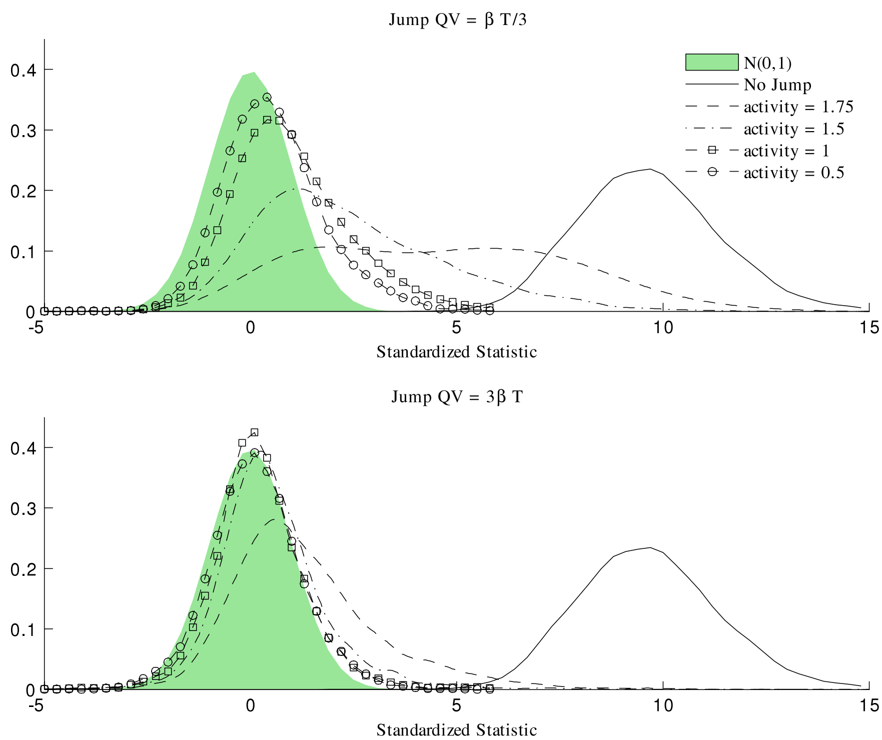

Fig. 5. Finite-sample distributions of the standardized test statistic for testing the null hypothesis with jumps. The standardized statistic is $\left(S_{RJ}\left(g,h,4\right)_n - 1\right) / \Delta_n^{1/4}\sqrt{\Sigma^j_{RJ,n}}$. We plot the distribution in the continuous case (solid line). In the discontinuous case, the jump process is α-stable with activity 0.5 (dash-circle), 1 (dash-square), 1.5 (dash-dot) and 1.75 (dash). In each experiment, the jump process is scaled so that its realized quadratic variation is $\beta T/3$ (top) or $3\beta T$ (bottom), i.e., 1/3 or 3 times of the mean quadratic variation of the continuous part, respectively. In all cases, the averaging window is $k_n = 100$, the truncation level is $C = 5$, and the efficient price is contaminated with the Gaussian-T mixture noise without rounding.

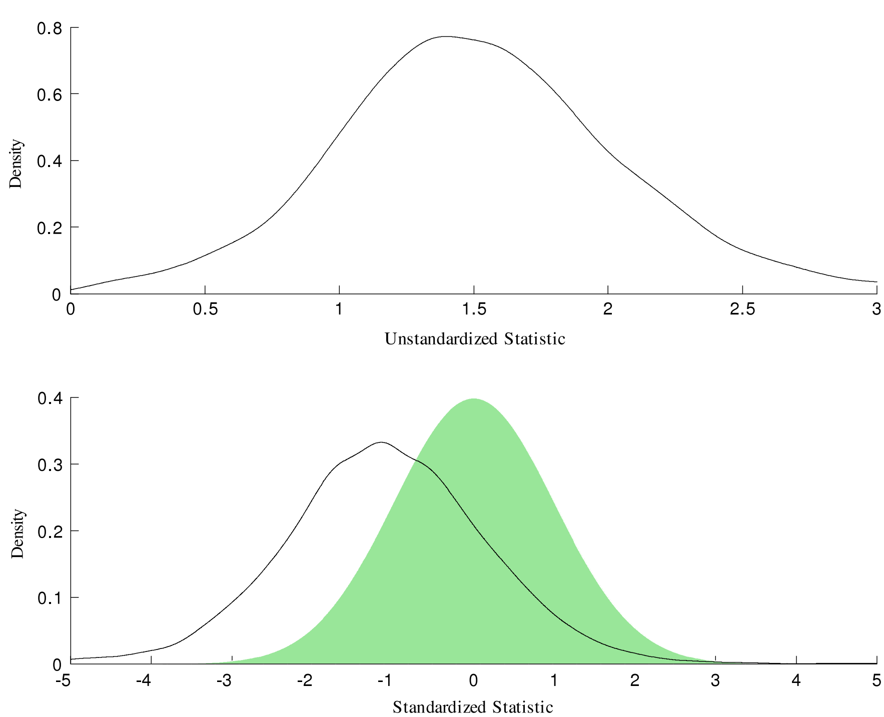

Fig. 6. Testing results for the 30 stocks of DJIA in 2008. In the top panel, we plot the empirical distribution of the unstandardized test statistic $S_{RJ}\left(g,h,4\right)_n$ for all 7590 stock-day pairs. In the bottom panel, we plot the empirical distribution of the standardized test statistic $\left(S_{RJ}\left(g,h,4\right)_n - 2\right) / \Delta_n^{1/4}\sqrt{\Sigma^c_{RJ,n}}$ (solid line) and, for comparison, the $\mathcal{N}\left(0,1\right)$ distribution (shaded area).

The rest of the Appendix is devoted to the proof of Proposition 3. After introducing some technical lemmas in Appendix B, we prove Proposition 3(a) and (b) in Appendices C and D, respectively.

Appendix A. Some known results and their consequences

We start by recalling some properties of the pre-averaging construction from Jacod et al. (2010). Below, g and h are two given weight functions, and $p \geq 4$ is an even integer. The sequence k_n, hence the number θ coming in (3.4), are also fixed.

First, we have some laws of large numbers. Namely, under Assumptions 1 and 3 we have

$$\frac{1}{k_n\,\bar{g}(p)}\,\overline{V}(Z,g,p)^n_T \xrightarrow{\mathbb{P}} U(p) := \sum_{s\leq T} |\Delta X_s|^p \tag{A.1}$$

and also, when further X is continuous:

$$\Delta_n^{1-p/4}\,\overline{V}(Z,g,p)^n_T \xrightarrow{\mathbb{P}} V(g,p) := m_p\,\theta^{p/2}\,\bar{g}(2)^{p/2}\int_0^T |\sigma_s|^p\,ds. \tag{A.2}$$

These two facts allow for a simple proof of Theorem 1:

Proof of Theorem 1. Since $U(p) > 0$ on the set Ω^j_T, the first convergence in (3.14) readily follows from (A.1). For the second convergence in (3.14) we cannot apply (A.2) right away, because X is not necessarily continuous, even though it is so on Ω^c_T. We set $S = \inf(t\colon \Delta X_t \neq 0)$ and $X'_t = X_{t\wedge S} - \Delta X_S\,1_{\{S\leq t\}}$. The process X' is a continuous semimartingale satisfying Assumption 1 by construction, with the volatility process $\sigma_t\,1_{\{t\leq S\}}$. Furthermore on the set Ω^c_T we have $S \geq T$ and $X'_t = X_t$ for all $t \leq T$: so the variables $\overline{V}(Z,g,p)^n_T$ associated with X and with X' coincide on the set Ω^c_T, as well as $\int_0^T |\sigma_s|^p\,ds$. Hence we deduce from (A.2) applied to X' that

$$\Delta_n^{1-p/4}\,\overline{V}(Z,g,p)^n_T \xrightarrow{\mathbb{P}} V(g,p) \quad \text{in restriction to } \Omega^c_T.$$

Then in view of the definition of $S_{RJ}(g,h,p)_n$, the second convergence in (3.14) is obvious. □

Second, we have some central limit theorems. First, assume that X is continuous and satisfies Assumptions 2 and 3. To describe the limit we need some notation. Consider two independent Brownian motions W^1 and W^2, given on another auxiliary filtered probability space $(\Omega', \mathcal{F}', (\mathcal{F}'_t)_{t\geq 0}, \mathbb{P}')$. Associated with the weight function g, we define the following Wiener integral processes

$$L(g)_t = \int g(s-t)\,dW^1_s, \qquad L'(g)_t = \int g'(s-t)\,dW^2_s, \tag{A.3}$$

and $L(h)$ and $L'(h)$ are defined likewise with the weight function h instead of g, with the same W^1 and W^2. The four dimensional process $(L(g), L'(g), L(h), L'(h))$ is continuous stationary centered Gaussian. We then set for $\eta,\zeta \in \mathbb{R}$ and q, q' even integers:

$$\left.\begin{aligned}
&m_q(g;\eta,\zeta) = \mathbb{E}'\big((\eta L(g)_1 + \zeta L'(g)_1)^q\big)\\
&m_{q,q'}(g,h;\eta,\zeta)_t\\
&\quad = \mathbb{E}'\Big((\eta L(g)_1 + \zeta L'(g)_1)^q\,(\eta L(h)_t + \zeta L'(h)_t)^{q'}\Big)\\
&\mu(g,h;\eta,\zeta) = \sum_{r,r'=0}^{p/2} \rho(p)_r\,\rho(p)_{r'}\\
&\quad \big(2\zeta^2\bar{g}'(2)\big)^r\,\big(2\zeta^2\bar{h}'(2)\big)^{r'}\\
&\int_0^2 \Big(m_{p-2r,p-2r'}(g,h;\eta,\zeta)_t\\
&\quad - m_{p-2r}(g;\eta,\zeta)\,m_{p-2r'}(h;\eta,\zeta)\Big)\,dt\\
&R(g,h) = \theta^{1-p}\int_0^T \mu(g,h;\theta\sigma_s,\alpha_s)\,ds.
\end{aligned}\right\} \tag{A.4}$$

With all this notation, we then have

$$\left(\frac{\Delta_n^{1-p/4}\,\overline{V}(Z,g,p)^n_T - V(g,p)}{\Delta_n^{1/4}},\ \frac{\Delta_n^{1-p/4}\,\overline{V}(Z,h,p)^n_T - V(h,p)}{\Delta_n^{1/4}}\right) \xrightarrow{\mathcal{L}-(s)} (\overline{V}(g,p), \overline{V}(h,p)) \tag{A.5}$$

where $\mathcal{L}-(s)$ denotes stable convergence in law, and where $(\overline{V}(g,p), \overline{V}(h,p))$ is defined on an extension of the space and, conditionally on $\mathcal{F}$, is a Gaussian centered vector with the covariance matrix

$$\begin{pmatrix} R(g,g) & R(g,h)\\ R(g,h) & R(h,h)\end{pmatrix}.$$

The following proposition is then a simple consequence of this result plus the delta method, together with the fact that in restriction to Ω^c_T we can argue as if the process X were everywhere continuous, exactly as in the proof of Theorem 1:

Proposition 1. *We have (a) of Theorem 2, with*

$$\Sigma(g,h,p,\theta) = \frac{R(g,g) - 2\gamma^{p/2}R(g,h) + \gamma^p R(h,h)}{\gamma'^2\,V(h,p)^2} \tag{A.6}$$

Finally we consider the CLT associated with the convergence (A.1). Under Assumptions 1 and 3,

$$\left(\frac{\overline{V}(Z,g,p)^n_T - k_n\bar{g}(p)\,U(p)}{k_n\,\Delta_n^{1/4}},\ \frac{\overline{V}(Z,h,p)^n_T - k_n\bar{h}(p)\,U(p)}{k_n\,\Delta_n^{1/4}}\right) \xrightarrow{\mathcal{L}-(s)} (\overline{U}(g,p), \overline{U}(h,p)) \tag{A.7}$$

where $(\overline{U}(g,p), \overline{U}(h,p))$ is defined on an extension of the space and, conditionally on $\mathcal{F}$, is a centered random vector with covariance matrix

$$\begin{pmatrix} D(g,g) & D(g,h)\\ D(g,h) & D(h,h)\end{pmatrix},$$

where

$$D(g,h)_t = p^2\theta\sum_{s\leq t}|\Delta X_s|^{2p-2}\Big(\Psi_-(g,h)\sigma^2_{s-} + \Psi_+(g,h)\sigma^2_s + \frac{1}{\theta^2}\Psi'_-(g,h)\alpha^2_{s-} + \frac{1}{\theta^2}\Psi'_+(g,h)\alpha^2_s\Big) \tag{A.8}$$

(notation (3.24)). Therefore, exactly as for Proposition 1, we get

Proposition 2. *We have (a) of Theorem 3, with*

$$\Sigma^j_{RJ}(g,h,p,\theta) = \frac{D(g,g) - 2\gamma' D(g,h) + (\gamma')^2 D(h,h)}{(\bar{g}\,(p)\,U(p))^2}. \tag{A.9}$$

Now we define $\Sigma^c_{RJ,n}$ and $\Sigma^j_{RJ,n}$ by (3.20) and (3.25). Suppose for a moment (we will prove this result below) that we have the following behavior:

Proposition 3. (a)*Under the assumptions of Theorem 2 we have*

$$\Sigma^c_{RJ,n} \xrightarrow{\mathbb{P}} \Sigma^c_{RJ}(g,h,p,\theta) \qquad \text{on the set } \Omega^c_T. \tag{A.10}$$

$$\Delta_n^{1/2}\,\Sigma^c_{RJ,n} \xrightarrow{\mathbb{P}} 0 \qquad \text{on the set } \Omega. \tag{A.11}$$

(b)*Under the assumptions of Theorem 3 we have*

$$\Sigma^j_{RJ,n} \xrightarrow{\mathbb{P}} \Sigma^j_{RJ}(g,h,p,\theta) \qquad \text{on the set } \Omega^j_T. \tag{A.12}$$

$$\Delta_n^{1/2}\Sigma^j_{RJ,n} \xrightarrow{\mathbb{P}} 0 \qquad \text{on the set } \Omega. \tag{A.13}$$

Then Theorems 2 and 3 immediately follow from the previous three propositions, whereas we have:

Proof of Theorems 4 and 5. By Theorem 2, the standardized variables

$$T^c_n = (S_{RJ}(g,h,p)_n - \gamma'')/\Delta_n^{1/4}\sqrt{\Sigma^c_{RJ,n}}$$

converge stably in law, in restriction to the set Ω^c_T to an $\mathcal{N}(0,1)$ random variable. The claim about the asymptotic level in Theorem 4 is then obvious. As for the claim about the asymptotic power in the same theorem, it follows from the first convergence in (3.14) and from (A.11). For Theorem 5 the proof is exactly the same for the power of the two tests (use the second convergence in (3.14) and (A.13)), and also for the level because the variables

$$T^j_n = (S_{RJ}(g,h,p)_n - 1)/\Delta_n^{1/4}\sqrt{\Sigma^j_{RJ,n}}$$

converge stably in law, in restriction to the set Ω^j_T to an $\mathcal{N}(0,1)$ random variable. □

Appendix B. Preliminary results

We first derive some estimates for the variables $\overline{Z}(g)^n_i$ and $\widehat{Z}(g)^n_i$, where g is a generic weight function. Those estimates are valid under strengthened versions of our assumptions, namely

Assumption 4. We have Assumption 1 and the processes (b_t), σ_t and $\sup_{z\in E}|\delta(\omega^{(0)},t,z)|/\gamma(z)$ are bounded.

Assumption 5. We have Assumption 3 and the processes $\beta(q)_t$ are all bounded (by a constant depending on q).

Below the constants are denoted by K and vary from line to line, and may depend on the characteristics of the process to which they apply, and on the weight function which is used. They are written K_q if they depend on some extra parameter q.

Lemma 1. *Suppose that Assumptions 4 and 5 hold. Let $i_n \geq 1$ be (possibly random) indices, such that $T_n = i_n\Delta_n$ are stopping times. Then, recalling u_n in (3.15), we have for all $q > 0$ and $j \geq 1$ and for some sequence $\rho_n \to 0$ and with L denoting a bound for the jump sizes of X:*

$$\begin{aligned}
&\mathbb{E}(|\overline{Z}(g)^n_{i_n}|^q) \leq K_q\big(\Delta_n^{q/4} + L^{(q-2)^+}\Delta_n^{(q/4)\wedge(1/2)}\big),\\
&\mathbb{E}(|\widehat{Z}(g)^n_{i_n}|^q) \leq K_q\Delta_n^{q/2}.
\end{aligned} \tag{B.1}$$

$$\begin{aligned}
&\mathbb{E}\big(|\overline{X}(g)^n_{i_n+j} - \sigma_{T_n}\overline{W}(g)^n_{i_n+j}|^2 \wedge u_n^2\big) \leq K\Delta_n + K\Delta_n^{1/2}\rho_n\\
&\quad + K\mathbb{E}\left(\int_{T_n+(j-1)\Delta_n}^{T_n+(k_n+j-1)\Delta_n}|\sigma_s - \sigma_{T_n}|^2 ds\right)
\end{aligned} \tag{B.2}$$

$$\mathbb{E}\big(|\widehat{Z}(g)^n_{i_n} - \widehat{\epsilon}(g)^n_{i_n}|^q\big) \leq K\big(\Delta_n^{q/2+q\wedge 1} + \Delta_n^{3q/4+(q/2)\wedge 1}\big). \tag{B.3}$$

Proof. The second part of (B.1) is (5.43) of Jacod et al. (2010), and for the first part we use (5.3) and (5.4) of that paper, the latter being

$$\overline{X}(g)^n_i = \int_{i\Delta_n}^{(i+k_n)\Delta_n} g_n(s - i\Delta_n)\,dX_s, \tag{B.4}$$

where $g_n(s) = \sum_{j=1}^{k_n-1} g^n_j\,1_{((j-1)\Delta_n,j\Delta_n]}(s)$. Then the result follows from the Burkholder–Davis–Gundy inequality and $|g_n| \leq K$ (the fact that i_n is random changes nothing, since $i_n\Delta_n$ is a stopping time). For (B.2) we decompose X as $X = X' + X''$, where $X'_t = \int_0^t b_s\,ds + \int_0^t \sigma_s\,dW_s$. (B.4) yields

$$\begin{aligned}
\overline{X}(g)^n_{i_n+j} - \sigma_{T_n}\overline{W}(g)^n_{i_n+j} &= \int_{T_n+(j-1)\Delta_n}^{T_n+(k_n+j-1)\Delta_n} g_n(s-T_n)\big(b_s\,ds\\
&\quad + (\sigma_s - \sigma_{T_n})\,dW_s\big) + \overline{X''}(g)^n_{i_n+j}.
\end{aligned}$$

The expectation of the squared integral above is smaller than $K\Delta_n + \mathbb{E}\big(\int_{T_n+j\Delta_n}^{T_n+(k_n+j)\Delta_n}|\sigma_s - \sigma_{T_n}|^2 ds\big)$, because of Assumption 4. On the other hand, an easy adaptation of (6.25) of Jacod (2007) shows that $\mathbb{E}(|\overline{X''}(g)^n_{i_n+j}|^2 \wedge u_n^2) \leq \Delta_n^{1/2}\rho_n$ for some sequence $\rho_n \to 0$. These two properties yield (B.2). Next, observe that $\widehat{Z}(g)^n_{i_n} - \widehat{\epsilon}(g)^n_{i_n} = \widehat{X}(g)^n_{i_n} + a_n + a'_n$, where

$$a_n = 2\sum_{j=1}^{[k_n/2]}(g'^n_{2j})^2\big(\Delta^n_{i_n+2j}X\,\Delta^n_{i_n+2j}\epsilon\big),$$

$$a'_n = 2\sum_{j=0}^{[(k_n-1)/2]}(g'^n_{2j+1})^2\big(\Delta^n_{i_n+2j+1}X\,\Delta^n_{i_n+2j+1}\epsilon\big).$$

The summands of a^n_i are martingale increments. Then by the Burkholder–Davis–Gundy and Hölder inequalities, plus Assumption 5 and the well known property $\mathbb{E}(|\Delta^n_{i_n+j}X|^q) \leq K_q\Delta_n^{(q/2)\wedge 1}$ and also $|g'^n_j| \leq K/k_n$, we get $\mathbb{E}(|a_n|^q) \leq K_q\Delta_n^{3q/4+(q/2)\wedge 1}$ for $q \geq 1$, hence also for $q \in (0,1)$ by Hölder's inequality again. The same holds for a'_n, and another application of Hölder's inequality yields $\mathbb{E}(|\widehat{X}(g)^n_{i_n}|^q) \leq K_q\Delta_n^{q/2+q\wedge 1}$. Then, upon using the last part of (B.1), we obtain (B.3). □

Our second preliminary concerns the behavior of the truncated variations $V^*(Z,g,q,r)^n$ of (3.18):

Lemma 2. *Suppose that Assumptions 4 and 5 and (3.15) hold, and also that X is continuous. Then if q is an even integer and $w \in \{0,\ldots,q/2\}$, we have*

$$\begin{aligned}
&\Delta_n^{1-q/4}V^*(Z,g,q-2w,w)^n_t\\
&\quad\xrightarrow{u.c.p.}\theta^{-q/2}\int_0^t(2\alpha_s^2\bar{g}'(2))^w m_{q-2w}(g;\theta\sigma_s,\alpha_s)\,ds.
\end{aligned} \tag{B.5}$$

Proof. A classical localization procedure allows to suppose the strengthened Assumptions 4 and 5. We can reproduce the proof of Theorem 3.3 in Jacod et al. (2010) with the functions $f_n(x,y,z) = f(x,y,z) = |x(0)+y(0)|^{q-2w}|z(0)|^w$ to get

$$\begin{aligned}
&\Delta_n^{1-q/4}V(Z,g,q-2w,w)^n_t\\
&\quad\xrightarrow{u.c.p.}\theta^{-q/2}\int_0^t(2\alpha_s^2\bar{g}'(2))^w m_{q-2w}(g;\theta\sigma_s,\alpha_s)\,ds.
\end{aligned}$$

Therefore it remains to prove that for any $t > 0$ we have $V^*(Z,g,q-2w,w)^n_s = V(Z,g,q-2w,w)^n_s$ for all $s \leq t$, on a set Ω^n_t which satisfies $\mathbb{P}(\Omega^n_t) \to 1$ as $n \to \infty$. The first part of (B.1) applied with $q = \frac{8}{1-4\varpi}$ and $L = 0$ and Markov's inequality yield $\mathbb{P}(|\overline{Z}(\phi)^n_i| > u_n) \leq K\Delta_n^2$. Therefore the set Ω^n_t on which $|\overline{Z}(\phi)^n_i| \leq u_n$ for all $i = 1,\ldots,[t/\Delta_n]$ satisfies all our requirements. □

We need a "local" result of the same type, at least when $q = 2$, but when X has jumps. We still have our random integers i_n such that $T_n = i_n\Delta_n$ is a stopping time. We also have integers $k'_n \geq k_n$ satisfying (3.21). Then $T'_n = T_n + k'_n\Delta_n$ is also a stopping time, and we consider two cases, where T is again a stopping time:

$$\begin{aligned}
&\text{case (1)}: T_n \to T \text{ and } T'_n \leq T \text{ for all } n\\
&\text{case (2)}: T'_n \to T \text{ and } T_n \geq T \text{ for all } n.
\end{aligned} \tag{B.6}$$

Below, w takes the values 0 or 1. We consider the variables

$$\left.\begin{array}{ll} G(g,0)=\bar g(2)\sigma_{T-}^2+\frac{1}{\theta^2}\,\bar g'(2)\alpha_{T-}^2, & \\ \quad G(g,1)=\frac{2}{\theta^2}\,\bar g'(2)\alpha_{T-} & \text{in case (1)} \\ G(g,0)=\bar g(2)\sigma_{T}^2+\frac{1}{\theta^2}\,\bar g'(2)\alpha_{T}^2, & \\ \quad G(g,1)=\frac{2}{\theta^2}\,\bar g'(2)\alpha_{T} & \text{in case (2).} \end{array}\right\} \tag{B.7}$$

Lemma 3. *Suppose that Assumptions 4 and 5 hold and $k'_n/k_n \to \infty$. Let w be either 0 or 1. Then*

$$\frac{1}{k_n\,k'_n\,\Delta_n}\big(V^*(Z,g,2-2w,w)^n_{T'_n+k_n\Delta_n} - V^*(Z,g,2-2w,w)^n_{T_n+k_n\Delta_n}\big) \xrightarrow{\mathbb{P}} G(g,w). \tag{B.8}$$

This is, for $q = 2$, the local version of the previous lemma, since it can be easily checked (see later an explicit expression for $m_q(g;\eta,\zeta)$) that $G(g,w)$ in case (2) for example is the value of the integrand in the right side of (B.5) evaluated at time T.

Proof. (1) We can again assume Assumptions 4 and 5. Set $f_0(x,y)=x^2$ and $f_1(x,y)=y$ and

$$\beta_i^n=\sigma_{i_n\Delta_n}\overline{W}(g)^n_{i_n+i},\qquad \beta_i'^n=\overline{\epsilon}(g)^n_{i_n+i},\qquad \widehat{\beta}_i^n=\widehat{\epsilon}(g)^n_{i_n+i}.$$

In this step we prove that

$$H(w)_n=\frac{1}{k_n\,k'_n\,\Delta_n}\,\mathbb{E}\left(\sum_{i=1}^{k'_n}\Big|f_w(\overline{Z}(g)^n_{i_n+i},\widehat{Z}(g)^n_{i_n+i})\,1_{\{|\overline{Z}(g)^n_{i_n+i}|\le u_n\}} - f_w(\beta_i^n+\beta_i'^n,\widehat{\beta}_i^n)\Big|\right)\to 0. \tag{B.9}$$

By virtue of the definition of f_w we have for $w=0,1$:

$$|f_w(x+x',y+y')1_{\{|x+x'|\le u_n\}}-f_w(x,y)| \le |y'|+\frac{|y|\,|x+x'|}{u_n}+\varepsilon x^2+\frac{2(x'^2\wedge u_n^2)}{\varepsilon}+\frac{2x^4}{u_n^2}+\frac{x^2|x+x'|}{u_n}$$

for all $\varepsilon\in(0,1)$. Then we take $x=\beta_i^n+\beta_i'^n$ and $y=\widehat{\beta}_i^n$ and $x'=\overline{X}(g)^n_{i_n+i}-\beta_i^n$ and $y'=\widehat{Z}(g)^n_{i_n+i}-y$ and apply (B.3) for $q=1$ and (B.1) and (B.2), plus the Cauchy–Schwarz inequality, to get

$$H(w)_n\le\frac{K\Delta_n^{1/2}}{k_n\Delta_n}\left(\Delta_n^{1/4-\varpi}+\varepsilon+\frac{\rho_n+\rho'_n}{\varepsilon}+\Delta_n^{1/2-2\varpi}\right),$$

where

$$\rho'_n=\frac{1}{k'_n\Delta_n^{1/2}}\sum_{i=1}^{k'_n}\mathbb{E}\left(\int_{T_n+(i-1)\Delta_n}^{T_n+(k_n+i-1)\Delta_n}|\sigma_s-\sigma_{T_n}|^2\,ds\right) \le K\,\mathbb{E}\big(\sup_{T_n\le s\le T'_n}|\sigma_s-\sigma_{T_n}|^2\big).$$

Since σ_t is càdlàg and bounded, we see that $\rho'_n\to 0$ in both cases (1) and (2). Then since $\varpi<1/4$ we get $\limsup_n H(w)_n\le K\varepsilon$, and (B.9) follows because ε is arbitrarily small. (2) If $\zeta_i^n=f_w(\beta_i^n+\beta_i'^n,\widehat{\beta}_i^n)$, and by (B.9), it remains to prove that

$$\frac{1}{k_n\,k'_n\,\Delta_n}\sum_{i=1}^{k'_n}\zeta_i^n\xrightarrow{\mathbb{P}}G(g,w). \tag{B.10}$$

Set $\zeta_i'^n=\mathbb{E}(\zeta_i^n\mid\mathcal{F}_{(i_n+i-1)\Delta_n})$ and $\zeta_i''^n=\zeta_i^n-\zeta_i'^n$. Since $(i_n+i-1)\Delta_n$ is a stopping time for each $i\ge 1$ and ζ_i^n is $\mathcal{F}_{(i_n+i+k_n-1)\Delta_n}$-measurable, we have

$$\mathbb{E}\left(\left(\frac{1}{k_n\,k'_n\,\Delta_n}\sum_{i=1}^{k'_n}\zeta_i''^n\right)^2\right) \le\frac{2}{k_n^2\,k_n'^2\,\Delta_n^2}\sum_{i=1}^{k'_n}\sum_{j=0}^{(k_n-1)\wedge(k'_n-i)}\big|\mathbb{E}(\zeta_i^n\,\zeta_{i+j}^n)\big| \le\frac{2k_n}{k_n^2\,k_n'^2\,\Delta_n^2}\sum_{i=1}^{k'_n}\mathbb{E}(|\zeta_i^n|^2).$$

(B.1) yields $\mathbb{E}(|\zeta_i^n|^2)\le K\Delta_n$, so the right side above goes to 0 because $k_n/k'_n\to 0$, and instead of (B.10) it is then enough to prove that

$$\frac{1}{k_n\,k'_n\,\Delta_n}\sum_{i=1}^{k'_n}\zeta_i'^n\xrightarrow{\mathbb{P}}G(g,w). \tag{B.11}$$

(3) Due to the special form of f_w, we can calculate $\zeta_i'^n$ explicitly:

$$\zeta_i'^n=\begin{cases}\sum_{j=1}^{k_n}(g_j^n)^2\sigma_{T_n}^2\Delta_n+\sum_{j=1}^{k_n}(g_j'^n)^2\alpha^2_{(i_n+i+j-1)\Delta_n} & \text{if } w=0\\ \sum_{j=1}^{k_n}(g_j'^n)^2\big(\alpha^2_{(i_n+i+j-1)\Delta_n}+\alpha^2_{(i_n+i+j)\Delta_n}\big) & \text{if } w=1\end{cases}$$

(we heavily use the independence and centering properties of the noise, see Assumption 3). We also observe that, due to the properties of the weight function,

$$k_n\sum_{j=1}^{k_n}(g_j'^n)^2\to\bar g'(2)$$

and

$$\frac{1}{k_n}\sum_{j=1}^{k_n}(g_j^n)^2\to\bar g(2).$$

Since σ_t and α_t are càdlàg and (3.4) holds, we readily deduce that in case (1), $|k_n\zeta_i'^n-\theta^2 G(g,w)|$ goes to 0 (pathwise) and stays bounded, uniformly in $i=0,\ldots,k'_n$. Using (3.4) once more, (B.11) follows. □

Appendix C. The behavior of $\Sigma^{c}_{RJ,n}$

Here we prove the first claim (a) of Proposition 3. For this, we begin by showing that $M^*(g,h,\phi)^n_T$ is an estimator for $R(g,h)$: in fact another estimator for $R(g,h)$ is already provided in Jacod et al. (2010); however simulation studies suggest that $M^*(g,h,p)^n_t$ behaves better, at least when $p = 4$. We start with some calculations:

Lemma 4. *With the notation (3.17), we have*

$$\mu(g,h;\eta,\zeta)=\sum_{w=0}^{p}\eta^{2w}\,\zeta^{2p-2w}\,A'(g,h;w). \tag{C.1}$$

Proof. Consider the processes $L(g)$ and $L'(g)$ defined by (Eq. (A.3)). We use the (3.16). First, we have $\mathbb{E}'(L(g)_1\,L(h)_t)=a(g,h)_t$, and the process $(L(g),L(h))$ is stationary centered Gaussian. Then a well known fact about 2-dimensional centered Gaussian vectors yields that for w a nonnegative integer and $w' \in \{0, \dots, w\}$:

$$w \text{ even} \Rightarrow \mathbb{E}'(L(g)_t^w) = m_w\, \bar{g}(2)$$
$$\mathbb{E}'(L(g)_1^{w-w'}\, L(h)_t^{w'}) = a'(g, h; w/2, w')_t$$
$$w \text{ odd} \Rightarrow \mathbb{E}'(L(g)_t^w) = 0$$
$$\mathbb{E}'(L(g)_1^{w-w'}\, L(h)_t^{w'}) = 0.$$

We have the same for $L'(g)$ and $L'(h)$, provided we substitute (g, h) with (g', h') in the right hand sides above. Since further $(L(g), L(h))$ is independent from $(L'(g), L'(h))$, and upon using the binomial formula, we deduce that for q, q' even integers

$$m_q(g; \eta, \zeta) = \sum_{l=0}^{q/2} C_q^{2l}\, \eta^{2l} \zeta^{q-2l}\, m_{2l}\, m_{q-2l}\, (\bar{g}(2))^l\, (\bar{g}'(2))^{q/2-l} \tag{C.2}$$

$$\begin{aligned} m_{q,q'}(g, h; \eta, \zeta)_t &= \sum_{l=0}^{q} \sum_{l'=0}^{q'} 1_{\{l+l' \text{ is even}\}} C_q^l\, C_{q'}^{l'}\, \eta^{l+l'}\, \zeta^{q+q'-l-l'} \\ &\quad \times a'(g, h; l/2 + l'/2, l')_t\, a'(g', h'; q/2 \\ &\quad + q'/2 - l/2 - l'/2, q' - l')_t \\ &= \sum_{w=0}^{q/2+q'/2} \eta^{2w} \zeta^{q+q'-2w} \sum_{w'=(2w-q)^+}^{q' \wedge (2w)} C_q^{2w-w'}\, C_{q'}^{w'} \\ &\quad \times a'(g, h; w, w')_t\, a'(g', h'; q/2 \\ &\quad + q'/2 - w, q' - w')_t. \end{aligned}$$

Using (3.9), we first deduce that

$$\sum_{r=0}^{p/2} \rho(p)_r (2\zeta^2 \bar{g}'(2))^r\, m_{p-2r}(g; \eta, \zeta) = m_p\, \eta^p\, (\bar{g}(2))^{p/2}. \tag{C.3}$$

Then in view of (3.17), we end up with (C.1). $\square$

Lemma 5. *Under Assumptions 1 and 3, we have*

$$M^*(g, h, \phi)_T^n \xrightarrow{\mathbb{P}} R(g, h) \quad \text{on the set } \Omega_T^c, \tag{C.4}$$

$$\Delta_n^{p/2-1/2}\, M^*(g, h, \phi)_T^n \xrightarrow{\mathbb{P}} 0. \tag{C.5}$$

Proof. For (C.4) it is enough, by the same argument as in Theorem 1, to consider the case when X is continuous. Then we can apply (B.5) with $q = p$ and w substituted with $p + l - w$ and sum over l between 0 and w: taking advantage of (C.3) with $2w$ instead of p, we readily deduce

$$\begin{aligned} &\Delta_n^{1-p/2} \sum_{l=0}^{w} \rho(2w)_l\, V^*(Z, \phi, 2w - 2l, p + l - w)_t^n \\ &\xrightarrow{u.c.p.} \frac{2^{p-w}}{\theta^p}\, (\bar{\phi}(2))^w\, (\bar{\phi}'(2))^{p-w}\, m_{2w} \int_0^t \sigma_s^{2w}\, \alpha_s^{2p-2w}\, ds. \end{aligned}$$

At this stage we readily deduce the result from (A.4) to (C.1) and the Definition (3.19). Now we turn to (C.5). Clearly, it is enough to show that $\sqrt{\Delta_n} V^*(Z, \phi, 2p - 2l, l)_t^n \xrightarrow{\mathbb{P}} 0$ for each $l \in \{0, \dots, p\}$. When $l = p$, by the second part of (B.1) with $q = p$,

$$\mathbb{E}\big(V^*(Z, \phi, 0, p)_t^n\big) \le Kt\Delta_n^{p/2-1},$$

which implies the result. When $l = 0$, by the first part of (B.1) with $q = 2$,

$$\mathbb{E}\big(V^*(Z, \phi, 2p, 0)_t^n\big) \le u_n^{2p-2} \sum_{i=0}^{[t/\Delta_n]-k_n} \mathbb{E}\big(|\overline{Z}(\phi)_i^n|^2\big) \le Kt\Delta_n^{(2p-2)\varpi-1/2}$$

which again implies the result because $\varpi > 1/12$ and $p \ge 4$. If $1 \le l \le p - 1$, by Hölder's inequality,

$$\sqrt{\Delta_n} V^*(Z, \phi, 2p - 2l, l)_t^n \le \big(\sqrt{\Delta_n} V^*(Z, \phi, 2p, 0)_t^n\big)^{\frac{p-l}{p}} \big(\sqrt{\Delta_n} V^*(Z, \phi, 0, p)_t^n\big)^{\frac{l}{p}}.$$

Therefore the result for these values of l follows from the result for $l = 0$ and $l = p$. $\square$

Proof of Proposition 3-(a). (A.10) is a straightforward consequence of (A.2) and (C.4), whereas (A.11) readily follows from (A.1) to (C.5). $\square$

Appendix D. The behavior of $\Sigma_{RJ,n}^j$

Now we turn to the behavior of $\Sigma_{RJ,n}^j$, that is we prove (b) of Proposition 3. As in the previous subsection, this essentially amounts to finding the behavior of the processes $N(\phi, \pm)_T^n$ and $N'(\phi, \pm)_T^n$, in connection with the four variables which enter the Definition (A.8) of $D(g, h)$, which are

$$\left.\begin{aligned} N(\phi, 0, -)_t &= \bar{\phi}(2)\, \bar{\phi}(2p-2) \sum_{s \le t} \sigma_{s-}^2\, |\Delta X_s|^{2p-2} \\ N(\phi, 0, +)_t &= \bar{\phi}(2)\, \bar{\phi}(2p-2) \sum_{s \le t} \sigma_s^2\, |\Delta X_s|^{2p-2} \\ N(\phi, 1, -)_t &= \frac{2}{\theta^2}\, \bar{\phi}'(2)\, \bar{\phi}(2p-2) \sum_{s \le t} \alpha_{s-}^2\, |\Delta X_s|^{2p-2} \\ N(\phi, 1, +)_t &= \frac{2}{\theta^2}\, \bar{\phi}'(2)\, \bar{\phi}(2p-2) \sum_{s \le t} \alpha_s^2\, |\Delta X_s|^{2p-2} \end{aligned}\right\} \tag{D.1}$$

Lemma 6. *Under Assumptions 1 and 3 we have for $m = 0, 1$:*

$$N(\phi, m, \pm)_T^n \xrightarrow{\mathbb{P}} N(\phi, m, \pm)_T. \tag{D.2}$$

Proof. We prove only the statements about $N(\phi, m, -)_T^n$, the others being similar (and in fact slightly simpler: we would not need to introduce below the "bigger" filtration $(\mathcal{G}_t)$). By localization again, we may suppose Assumptions 4 and 5. *Step* (1) We fix $\varepsilon \in (0, 1)$, and we denote by S_q the successive jump times of the Poisson process $(\underline{\mu}([0, t] \times \{z: \gamma(z) > \varepsilon\}): t \ge 0)$, with the convention $S_0 = 0$. Let $i(n, q)$ be the random integer such that $(i(n, q) - 1)\Delta_n < S_q \le i(n, q)\Delta_n$. Our aim is to prove that for $m = 0, 1$ and all $q \ge 1$ we have

$$\eta(\phi, m)_{i(n,q)-k_n-k_n'}^n \xrightarrow{\mathbb{P}} \overline{\eta}_q(m) = \begin{cases} \bar{\phi}(2)\, \sigma_{S_q-}^2 & \text{if } m = 0 \\ \dfrac{2}{\theta^2}\, \bar{\phi}'(2)\, \alpha_{S_q-}^2 & \text{if } m = 1. \end{cases} \tag{D.3}$$

For this, consider the processes

$$X(\varepsilon)_t = X_t - \sum_{q \ge 1} \Delta X_{S_q}\, 1_{\{S_q \le t\}}, \qquad Z(\varepsilon)_t = X(\varepsilon)_t + \epsilon_t, \tag{D.4}$$

and denote by $\eta'(\phi, m)_i^n$ the variables defined by (3.22), with Z substituted with $Z(\varepsilon)$. On the set $\{S_{q-1} < S_q - (2k_n + k_n')\Delta_n\}$, whose probability goes to 1 as $n \to \infty$, we have $\eta(\phi,m)^n_{i(n,q)-k_n-k'_n} = \eta'(\phi,m)^n_{i(n,q)-k_n-k'_n}$. Therefore it is enough to prove (B.9) with $\eta(\phi,m)^n_{i(n,q)-k_n-k'_n}$ substituted with $\eta'(\phi,m)^n_{i(n,q)-k_n-k'_n}$. Set $i_n = i(n,q) - k_n - k'_n$ and $T_n = i_n\Delta_n$. We observe that the left side of (B.8), written for $Z' = X(\varepsilon) + \epsilon$ instead of $Z = X + \epsilon$, is equal to $\eta'(\phi,0)^n_{i_n} - \frac{1}{2}\eta'(\phi,1)^n_{i_n}$ when $w = 0$ and to $\eta'(\phi,1)^n_{i_n}$ when $w = 1$. Then (D.3) follows from the convergence (B.8), provided we can apply Lemma 3 with $X(\varepsilon)$ and the above random i_n. To check this point, we call $(\mathcal{G}_t)$ the smallest filtration containing $(\mathcal{F}_t)$ and such that all stopping times S_q are $\mathcal{G}_0$-measurable. On the one hand, T_n is obviously a stopping time with respect to this bigger filtration. On the other hand, it is well known that $X(\varepsilon)$ is a $(\mathcal{G}_t)$-semimartingale satisfying Assumption 4 with respect to this bigger filtration, whereas Assumption 5 is obviously satisfied with respect to $(\mathcal{G}_t)$ as well. Therefore we are in a position to apply Lemma 3 and (D.3) is proved. *Step* (2) This step is devoted to some estimates, in the setting of the previous step. We denote by Ω_n the set on which, for all $q \geq 1$, either $S_q > T$ or $S_{q-1} + (2k_n + k'_n)\Delta_n < S_q < T - k_n\Delta_n$, so $\mathbb{P}(\Omega_n) \to 1$. We also set

$$\rho(0)^n_i = \left((\overline{Z}(\phi)^n_i)^2 - \frac{1}{2}\widehat{Z}(\phi)^n_i\right)1_{\{|\overline{Z}(\phi)^n_i| \leq u_n\}},$$
$$\rho(1)^n_i = \widehat{Z}(\phi)^n_i\, 1_{\{|\overline{Z}(\phi)^n_{i+j}| \leq u_n\}},$$

and use the notation $\rho'(m)^n_i$ if we substitute Z with $Z(\varepsilon)$. Using the initial filtration, and also the bigger filtration $(\mathcal{G}_t)$ and the fact that $\rho(m)^n_{i(n,q)-j} = \rho(m)'_{i(n,q)-j}$ when we are on the set Ω_n and $S_q \leq T$ and $k_n \leq j \leq 2k_n + k'_n$, we deduce from (B.1) that for any $i \geq 1$ and $k_n \leq j \leq 2k_n + k'_n$,

$$\mathbb{E}(|\rho(m)^n_i|) \leq K\Delta_n^{1/2},$$
$$\mathbb{E}\left(|\rho(m)^n_{i(n,q)-j}|\, 1_{\{S_q \leq T\}\cap\Omega_n}\right) \leq K\Delta_n^{1/2}. \tag{D.5}$$

This readily gives the following, for any $i \geq 1$ and $j = 1, \ldots, k_n$ and $m = 0, 1$:

$$\mathbb{E}(|\eta(\phi,m)^n_i|) \leq K,$$
$$\mathbb{E}\left(|\eta(\phi,m)^n_{i(n,q)-k_n-k'_n-j}|\, 1_{\{S_q \leq T\}\cap\Omega_n}\right) \leq K. \tag{D.6}$$

On the set $\{S_q \leq T\} \cap \Omega_n$ we have for $j = 1, \ldots, k_n$:

$$\eta(\phi,m)^n_{i(n,q)-k_n-k'_n-j} - \eta(\phi,m)^n_{i(n,q)-k_n-k'_n}$$
$$= \frac{1}{k_n\, k'_n\, \Delta_n}\sum_{l=0}^{j-1}\left(\rho(m)^n_{i(n,q)-k_n-k'_n-l} - \rho(m)^n_{i(n,q)-k'_n-l}\right)$$

and each sum above has at most k_n summands. Therefore (D.5) yields

$$\mathbb{E}\left(|\eta(\phi,m)^n_{i(n,q)-k_n-k'_n-j} - \eta(\phi,m)^n_{i(n,q)-k_n-k'_n}|\, 1_{\{S_q \leq T\}\cap\Omega_n}\right)$$
$$\leq Kk_n/k'_n.$$

Since $k_n/k'_n \to 0$ and $\mathbb{P}(\Omega_n) \to 1$, we then deduce from (D.3) that

$$\frac{1}{k_n}\sum_{j=1}^{k_n}\eta(\phi,m)^n_{i(n,q)-k_n-k'_n-j}\, |\phi^n_j|^{2p-2} \xrightarrow{\mathbb{P}} \bar{\phi}(2p-2)\,\overline{\eta}_q(m). \tag{D.7}$$

Finally, since $p \geq 4$, (B.1) applied to $Z(\varepsilon)$ with the filtration $(\mathcal{G}_t)$ gives

$$\mathbb{E}\left(|\overline{Z(\varepsilon)}(\phi)^n_i|^{2p-2} \mid \mathcal{G}_{i\Delta_n}\right) \leq K\left(\Delta_n^{p/2-1/2} + \varepsilon^{2p-4}\Delta_n^{1/2}\right)$$
$$\leq K\Delta_n^{1/2}. \tag{D.8}$$

Step (3) Recall that the variables in (D.3) implicitly depend on ε, and set

$$N(m,\varepsilon)_t = \bar{\phi}(2p-2)\sum_{q:S_q \leq t}\overline{\eta}_q(m)\,|\Delta X_{S_p}|^{2p-2}$$

$$B(m,\varepsilon)^n_t = \frac{1}{k_n}\sum_{i=k_n+k'_n}^{[t/\Delta_n]-k_n}\eta(\phi,m)^n_{i-k_n-k'_n}\,|\overline{Z(\varepsilon)}(\phi)^n_i|^{2p-2}$$

$$N(m,\varepsilon)^n_t = N(\phi,m,-)^n_t - B(m,\varepsilon)^n_t.$$

By virtue of (D.1) and (D.3) and the dominated convergence theorem, we have

$$N(m,\varepsilon)_T \to N(\phi,m,-)_T \quad \text{pointwise, as } \varepsilon \to 0.$$

Observing that $\eta(\phi,m)^n_{i-k_n-k'_n}$ is $\mathcal{F}_{i\Delta_n}$-measurable, by successive conditioning we deduce from (D.6) to (D.8) that

$$\mathbb{E}\left(|B(m,\varepsilon)^n_T|\right) \leq K\,T\,\left(\Delta_n^{p/2-3/2} + \varepsilon^{2p-2}\right),$$

which implies

$$\lim_{\varepsilon\to 0}\limsup_{n\to\infty}\mathbb{E}\left(|B(m,\varepsilon)^n_T|\right) = 0.$$

Therefore it remains to prove that, for $m = 0, 1$ and ε fixed and as $n \to \infty$, we have

$$N(m,\varepsilon)^n_T \xrightarrow{\mathbb{P}} N(m,\varepsilon)_T. \tag{D.9}$$

Step(4) Now we proceed to proving (D.9). Again ε is fixed, and we use the notation S_q and $i(n,q)$ of Step 1. In restriction to the set Ω_n, we have $N(m,\varepsilon)^n_T = \sum_{q\geq 1}\xi(m)^n_q\, 1_{\{S_q \leq T\}}$ where

$$\xi(m)^n_q = \frac{1}{k_n}\sum_{j=1}^{k_n}\eta(\phi,m)^n_{i(n,q)-k_n-k'_n-j}$$
$$\times\left(|\overline{Z}(\phi)^n_{i(n,q)-j}|^{2p-2} - |\overline{Z(\varepsilon)}(\phi)^n_{i(n,q)-j}|^{2p-2}\right)$$

because $\overline{Z}(\phi)^n_i = \overline{Z(\varepsilon)}(\phi)^n_i$ when i is not between $i(n,q) - k_n + 1$ and $i(n,q)$ for some q. Thus (D.9) will follow if we prove $\xi(m)^n_q \xrightarrow{\mathbb{P}} \bar{\phi}(2p-2)\overline{\eta}_q(m)\,|\Delta X_{S_q}|^{2p-2}$. In view of (D.7), it thus remains to prove that

$$\xi'(m)^n_q = \frac{1}{k_n}\sum_{j=1}^{k_n}\eta(\phi,m)^n_{i(n,q)-k_n-k'_n-j}\left(|\overline{Z}(\phi)^n_{i(n,q)-j}|^{2p-2}\right.$$
$$\left. - |\overline{Z(\varepsilon)}(\phi)^n_{i(n,q)-j}|^{2p-2} - |\phi^n_j\,\Delta X_{S_q}|^{2p-2}\right)$$

goes to 0 in probability. When $1 \leq j \leq k_n$, and on the set $\{S_q \leq T\} \cap \Omega_n$, we have $\overline{Z}(\phi)^n_{i(n,q)-j} = \overline{Z(\varepsilon)}(\phi)^n_{i(n,q)-j} + w^n_j$, where $w^n_j = \phi^n_j\Delta X_{S_q}$. Since for any reals x, y we have the estimate

$$\left||x+y|^{2p-2} - |x|^{2p-2} - |y|^{2p-2}\right|$$
$$\leq K\left(|x|^{2p-2} + |y|^{2p-3}\,|x|\right),$$

and since $|w^n_j| \leq K$, and upon using (B.1) and (D.8), it follows that for $1 \leq j \leq k_n$ we have

$$\mathbb{E}\left(\left||\overline{Z}(\phi)^n_{i(n,q)-j}|^{2p-2} - |\overline{Z(\varepsilon)}(\phi)^n_{i(n,q)-j}|^{2p-2}\right.\right.$$
$$\left.\left. - |\phi^n_j\,\Delta X_{S_q}|^{2p-2}\right|\mathcal{G}_{(i(n,q)-j)\Delta_n}\right)1_{\{S_q \leq T\}\cap\Omega_n} \leq K\Delta_n^{1/4}.$$

Since $\eta(\phi,m)^n_{i(n,q)-k_n-k'_n-j}$ is $\mathcal{G}_{(i(n,q)-j)\Delta_n}$-measurable by successive conditioning we deduce from the above and from (D.6) that

$$\mathbb{E}\left(|\xi'(m)^n_i|\, 1_{\{S_q \leq T\}\cap\Omega_n}\right) \leq K\Delta_n^{1/4},$$

and $\xi(m)^n_q \xrightarrow{\mathbb{P}} 0$ follows because $\mathbb{P}(\Omega_n) \to 1$. □

Proof of Proposition 3-(b). (A.12) is a straightforward consequence of (A.1) and (D.2), because $D(g,g) - 2\gamma' D(g,h) + (\gamma')^2 D(h,h)$ is equal to

$$\theta\, p^2 \left(\frac{\Psi_- N(\phi,0,-)_T + \Psi_+ N(\phi,0,+)_T}{\bar{\phi}(2)\,\bar{\phi}(2p-2)} + \frac{\Psi'_- N(\phi,1,-)_T + \Psi'_+ N(\phi,1,+)_T}{2\,\bar{\phi}'(2)\,\bar{\phi}(2p-2)} \right).$$

As for (A.13), it needs to be proved on the set Ω_T^c only (on Ω_T^j it follows from (A.12)). So by our usual argument we can assume that X is continuous. Then (A.2) shows that it is enough to prove that

$$\Delta_n^{3/2-p/2}\, N(\phi,m,\pm)_T \xrightarrow{\mathbb{P}} 0 \tag{D.10}$$

for $m = 0, 1$. We can again suppose Assumptions 4 and 5, by localization. Then if we combine (B.1) (with $L = 0$ because X is continuous) and (D.6), we readily obtain that $\mathbb{E}(|N(\phi,m,\pm)_T|) \le KT\,\Delta_n^{p/2-1}$, and thus (D.10) holds. □

References

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

Aït-Sahalia, Y., Jacod, J., 2009a. Estimating the degree of activity of jumps in high frequency financial data. Annals of Statistics 37, 2202–2244.

Aït-Sahalia, Y., Jacod, J., 2009b. Testing for jumps in a discretely observed process. Annals of Statistics 37, 184–222.

Aït-Sahalia, Y., Kimmel, R., 2007. Maximum likelihood estimation of stochastic volatility models. Journal of Financial Economics 83, 413–452.

Aït-Sahalia, Y., Mykland, P.A., Zhang, L., 2005. How often to sample a continuous-time process in the presence of market microstructure noise. Review of Financial Studies 18, 351–416.

Andersen, T.G., Bollerslev, T., Diebold, F.X., 2007. Roughing it up: including jump components in the measurement, modeling, and forecasting of return volatility. Review of Economics and Statistics 89, 701–720.

Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A., Shephard, N., 2008. Designing realized kernels to measure ex-post variation of equity prices in the presence of noise. Econometrica 76, 1481–1536.

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

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

Huang, X., Tauchen, G.T., 2005. The relative contribution of jumps to total price variance. Journal of Financial Econometrics 4, 456–499.

Jacod, J., 2007. Statistics and high frequency data: SEMSTAT Seminar. Tech. rep., Université de Paris-6.

Jacod, J., Li, Y., Mykland, P.A., Podolskij, M., Vetter, M., 2009. Microstructure noise in the continuous case: the pre-averaging approach. Stochastic Processes and Their Applications 119, 2249–2276.

Jacod, J., Podolskij, M., Vetter, M., 2010. Limit theorems for moving averages of discretized processes plus noise. Annals of Statistics 38, 1478–1545.

Jacod, J., Shiryaev, A.N., 2003. Limit Theorems for Stochastic Processes, second ed.. Springer-Verlag, New York.

Jiang, G.J., Oomen, R.C., 2008. Testing for jumps when asset prices are observed with noise—a swap variance approach. Journal of Econometrics 144, 352–370.

Lee, S., Mykland, P.A., 2008. Jumps in financial markets: a new nonparametric test and jump dynamics. Review of Financial Studies 21, 2535–2563.

Lee, S.S., Hannig, J., 2010. Detecting jumps from Lévy jump diffusion processes. Journal of Financial Economics 96, 271–290.

Li, Y., Mykland, P.A., 2007. Are volatility estimators robust with respect to modeling assumptions? Bernoulli 13, 601–622.

Xiu, D., 2010. Quasi-maximum likelihood estimation of volatility with high frequency data. Journal of Econometrics 159, 235–250.

Zhang, L., 2006. Efficient estimation of stochastic volatility using noisy observations: a multi-scale approach. Bernoulli 12, 1019–1043.

Zhang, L., Mykland, P.A., Aït-Sahalia, Y., 2005. A tale of two time scales: determining integrated volatility with noisy high-frequency data. Journal of the American Statistical Association 100, 1394–1411.