Testing for Jumps in Noisy High Frequency Data*

Yacine Aït-Sahalia
Department of Economics
Bendheim Center for Finance
Princeton University
and NBER[†]

Jean Jacod
Institut de Mathématiques
de Jussieu
CNRS – UMR 7586
Université Pierre et Marie Curie
Paris 6[‡]

Jia Li
Department of Economics
Bendheim Center for Finance
Princeton University[§]

This Version: February 2, 2010

Abstract

This paper proposes a robustification of the test statistic of Aït-Sahalia and Jacod (2009) for the presence of market microstructure noise in high frequency data, based on the pre-averaging method of Jacod et al. (2009).

KEYWORDS: Semimartingale; testing for jumps; high frequency data; market microstructure noise; pre-averaging.
JEL CODES: C22.

*Financial support from the NSF under grant SES-0850533 is gratefully acknowledged.
[†]Princeton, NJ 08544-1021. E-mail: yacine@princeton.edu.
[‡]75 013 Paris. E-mail: jean.jacod@upmc.fr.
[§]Princeton, NJ 08544-1021. E-mail: jiali@princeton.edu.

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, these models often consist of semimartingales.

We focus here on testing for the presence of jumps, 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) (based on bipower variations), Jiang and Oomen (2008) (based on a swap variance) Lee and Mykland (2008) (based on detecting large increments) and Aït-Sahalia and Jacod (2009) (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 as the sampling frequency increases, which distinguishes this problem from the classical measurement error 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. There are currently three main approaches to quadratic variation estimation, using nonparametric methods that are robust to market microstructure noise: linear combination of realized volatilities obtained by subsampling (Zhang et al. (2005) and Zhang (2006)), linear combination of autocovariances (Barndorff-Nielsen et al. (2008)) and pre-averaging (Jacod et al. (2009) and Jacod et al. (2010)).

In this paper, we examine the possibility of robustifying one of these tests for jumps, that of Aït-Sahalia and Jacod (2009), using the pre-averaging method. The test, whose asymptotic properties were 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 $k^{p/2-1}$.

In this paper, we consider what happens to the test statistic when market microstructure noise is taken into account. First, we study the impact of the noise on the statistic as defined. In the presence of noise, the 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 a rounding error. Then, using the pre-averaging approach, we show how to robustify the test statistic to restore its discriminating power between jumps and no jumps even in the presence of market microstructure noise.

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. Proofs are in Section 6.

2. The setting

2.1. The underlying process.

We have 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 eventually goes to 0. 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

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

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

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

b) the processes (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

$$X_t \;=\; X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s. \tag{2}$$

In this case, we sometimes need a stronger assumption on the coefficients, namely:

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

$$\sigma_t = \sigma_0 + \int_0^t \tilde{b}_s ds + \int_0^t \tilde{\sigma}_s dW_s + M_t + \sum_{s\le t} \Delta\sigma_s\, 1_{\{|\Delta\sigma_s|>v\}}, \tag{3}$$

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 $\sum_{s\le t} 1_{\{|\Delta\sigma_s|>v\}}$ is $\int_0^t a_s' ds$, and where $\tilde{b}_t$, a_t, and a_t' are optional locally bounded processes, and $\tilde{\sigma}_t$ is optional and càdlàg, as well as b_t.

Furthermore, we suppose that the processes $\tilde{b}_t$, a_t, a_t' are locally bounded, whereas the processes b_t and $\tilde{\sigma}_t$ are left-continuous with right limits. □

2.2. The noise.

Now, the process X is observed with an error: instead of X_t we observe

$$Z_t \;=\; X_t + \chi_t. \tag{4}$$

(Of course, the error χ_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.)

Mathematically speaking, this can be formalized as follows: for each $t \ge 0$, 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" $(\chi_t : t \ge 0)$ 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:

$$\left.\begin{array}{l} \Omega \;=\; \Omega^{(0)}\times\Omega^{(1)}, \quad \mathcal{F} \;=\; \mathcal{F}^{(0)}\otimes\mathcal{F}^{(1)}, \\ \mathcal{F}_t \;=\; \mathcal{F}_t^{(0)}\otimes\sigma(\chi_s : s\in[0,t)), \quad \mathcal{G}_t \;=\; \mathcal{F}^{(0)}\otimes\sigma(\chi_s : s\in[0,t)), \\ \mathbb{P}(d\omega^{(0)},d\omega^{(1)}) = \mathbb{P}^{(0)}(d\omega^{(0)})\;\mathbb{Q}(\omega^{(0)},d\omega^{(1)}). \end{array}\right\} \tag{5}$$

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 (1), on $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\geq 0},\mathbb{P})$, and W and $\underline{\mu}$ are a Wiener process and a Poisson random measure on this extended space as well.

We make the following assumption on the noise:

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*

$$\beta(q)_t(\omega^{(0)}) \;=\; \int Q_t(\omega^{(0)},dz)\,z^q, \qquad \alpha_t \;=\; \sqrt{\beta(2)_t}, \tag{6}$$

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

$$\beta(1) \;\equiv\; 0. \tag{7}$$

The reader can look at Jacod et al. (2009) or Jacod et al. (2010) for various comments on this assumption. We assume moments of all order 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 very 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 (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* uncondationally independent.

2.3. The hypotheses to be tested.

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

$$\left.\begin{array}{rcl} \Omega_T^c & = & \{\omega^{(0)}:\ t \mapsto X_t\left(\omega^{(0)}\right) \ \text{ is continuous on } \ [0,T]\} \\ \Omega_T^j & = & \{\omega^{(0)}:\ t \mapsto X_t\left(\omega^{(0)}\right) \ \text{ is discontinuous on } \ [0,T]\} \end{array}\right\} \tag{8}$$

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

$$\left.\begin{array}{rcl} \alpha & = & \sup\left(\limsup_n \ \mathbb{P}(C_n \mid A):\ A \in \mathcal{F},\ A \subset \Omega_0,\ \mathbb{P}(A) > 0\right) \\ \beta & = & \inf\left(\liminf_n \ \mathbb{P}(C_n \mid A):\ A \in \mathcal{F},\ A \subset (\Omega_0)^c,\ \mathbb{P}(A) > 0\right). \end{array}\right\} \tag{9}$$

3. The test statistics

3.1. The case with no noise.

We briefly recall the results of Aït-Sahalia and Jacod (2009). For any process Y and any integer $i \geq 1$ and real $p > 0$ we write

$$\Delta_i^n Y \ = \ Y_{i\Delta_n} - Y_{(i-1)\Delta_n}, \quad B(Y,p,\Delta_n)_t \ = \ \sum_{i=1}^{[t/\Delta_n]} |\Delta_i^n Y|^p. \tag{10}$$

We take an integer $k \geq 2$ and consider the test statistic

$$S(p,k,\Delta_n)_n \ = \ \frac{B(X,p,k\Delta_n)_T}{B(X,p,\Delta_n)_T}. \tag{11}$$

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 behavior, when $p > 2$:

$$S(p,k,\Delta_n)_n \ \xrightarrow{\mathbb{P}} \ \begin{cases} 1 & \text{on the set } \ \Omega_T^j \\ k^{p/2-1} & \text{on the set } \ \Omega_T^c. \end{cases} \tag{12}$$

Moreover a Central Limit Theorem allows 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

$C_n = \{S(p,k,\Delta_n)_n > 1 + v_n\}$ 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 do the same thing for the rejection regions $C_n = \{S(p,k,\Delta_n)_n < k^{p/2-1} - v_n\}$.

3.2. *The behavior of the AJ test when there is noise: simulation results*

When there is noise, X is unobservable and the AJ test statistic $S(p,k,\Delta_n)_n$ is infeasible. The seemingly natural alternative would be $S'(p,k,\Delta_n)_n$, computed by replacing the unobservable underlying process X with the noisy observable process Z. Unfortunately, this naive substitute is ill-behaved. We illustrate this point by simulation, with $p = 4$ and $k = 2$ fixed. We use an observation length of 5 days, with each day consisting of 6.5 trading hours, and sample the continuous-time process at every 5 seconds. To focus on the effect of noise, we simulate the underlying process X with constant volatility. We draw i.i.d. noise from a heavy-tailed distribution in order to capture infrequent but large bounce-backs in the real data used in the empirical study of Aït-Sahalia and Jacod (2009). We use three levels of noises to cover the spectrum of cases which are relevant in practice.

Figure 1 plots the histogram of the non-standardized AJ test statistic. In the benchmark case when there is no noise (top row), the non-standardized statistic is centered at 2 when there is no jump (shaded area) and at 1 when there is on average 1 large jump per day (solid curve), which confirms the theoretical prediction of Aït-Sahalia and Jacod (2009). When there are frequent but small jumps (dashed curve), the center of the distribution drifts slightly from 1 to 2. Sampling infrequently magnified this drift (right column).

When there is noise (middle row), the distribution of the statistic on continuous path strongly deviates from the prediction of the theory in Aït-Sahalia and Jacod (2009). Although sampling infrequently helps reducing the effect of noise, it is clearly unsatisfactory (right column). We remark the resemblance between the pattern in this simulation and the empirical finding in Figure 5 of Aït-Sahalia and Jacod (2009).

As the size of noise increases (bottom row), the deviation from asymptotic theory becomes more significant. Even with infrequent sampling, the statistic seems to be unable to discriminate continuous paths from discontinuous ones.

3.3. *Pre-averaging.*

Before exhibiting the robustified test statistics against noise, we need a rather large number of notation and conventions.

First, we choose a sequence of integers k_n satisfying for some $\theta > 0$:

$$k_n \sqrt{\Delta_n} \; = \; \theta + \mathrm{o}(\Delta_n^{1/4}). \tag{13}$$

Next, we consider weight functions g on $\mathbb{R}$, satisfying

$$\left.\begin{array}{l} g \text{ is continuous, piecewise } C^1 \text{ with a piecewise Lipschitz derivative } g', \\ s \notin (0,1) \quad \Rightarrow \quad g(s) = 0, \qquad \int g(s)^2 ds \; > \; 0, \end{array}\right\} \tag{14}$$

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

$$\left.\begin{array}{llllll} g_i^n & = & g(i/k_n), & g_i'^n & = & g_i^n - g_{i-1}^n, \\ \bar{g}(p) & = & \int |g(s)|^p \, ds, & \bar{g}'(p) & = & \int |g'(s)|^p \, ds. \end{array}\right\} \tag{15}$$

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

$$\overline{Y}(g)_i^n \; = \; \sum_{j=1}^{k_n - 1} g_j^n \Delta_{i+j}^n Y, \qquad \widehat{Y}(g;k_n)_i^n \; = \; \sum_{j=1}^{k_n} (g_j'^n \Delta_{i+j}^n Y)^2 \tag{16}$$

and processes

$$V(Y,g,q,r)_t^n = \sum_{i=0}^{[t/\Delta_n]-k_n} |\overline{Y}(g;k_n)_i^n|^q \; |\widehat{Y}(g;k_n)_i^n|^r \tag{17}$$

(they – implicitly – depend on the two sequences Δ_n and k_n)).

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

$$\left.\begin{array}{l} \rho(p)_0 \; = \; 1, \\ \sum_{l=0}^{j} 2^l \; m_{2j-2l} \; C_{p-2l}^{p-2j} \; \rho(p)_l \; = \; 0, \qquad j = 1,2,\cdots,p/2, \end{array}\right\} \tag{18}$$

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),

$$\rho(4)_0 = 1, \quad \rho(4)_1 = -3, \quad \rho(4)_2 = 0.75. \tag{19}$$

Then for any process Y we set

$$\overline{V}(Y,g,p)_t^n \;=\; \sum_{l=0}^{p/2} \rho(p)_l \, V(Y,g,p-2l,l)_t^n. \tag{20}$$

3.4. *First order asymptotic properties.*

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

$$\gamma \;=\; \frac{\bar{g}(2)}{\bar{h}(2)}, \qquad \gamma' \;=\; \frac{\bar{g}(p)}{\bar{h}(p)}, \qquad \gamma'' \;=\; \frac{\gamma^{p/2}}{\gamma'}, \tag{21}$$

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. Finally, we also choose a sequence k_n satisfying (13), and our test statistics is given by

$$S(g,h,p)_n \;=\; \frac{\overline{V}(Z,g,p)_T^n}{\gamma' \, \overline{V}(Z,h,p)_T^n}. \tag{22}$$

Here we describe the limiting behavior of the test statistics $S(g,h;p)_n$ given above:

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

$$S(g,h;p)_n \;\xrightarrow{\mathbb{P}}\; \begin{cases} 1 & \text{on the set } \Omega_T^j \\ \gamma'' & \text{on the set } \Omega_T^c. \end{cases} \tag{23}$$

So we always get an asymptotic behavior which is similar to the behavior of the statistics $S(p,k,\Delta_n)_n$ when there is no noise. If we take $h(s) = g(sk)$ for some $k \neq 1$ then $\bar{g}(q) = k\bar{h}(q)$ for any $q > 0$, so $\gamma'' = k^{p/2-1}$ and we even exactly retrieve (12), except that k does not need to be an integer here.

3.5. *Second order properties.*

In order to put the above statistics in use for practical test, we need a central limit theorem associated with the convergence in (23), and there are of course two very distinct behaviors on the two sets Ω_T^c and Ω_T^d.

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

$$u_n \;=\; \alpha\,\Delta_n^{\varpi}, \qquad \text{where} \;\; \alpha>0, \;\; \varpi\in\Big(\frac{1}{12},\frac14\Big). \tag{24}$$

These will serve as truncation levels. Next we introduce a number of constants, depending on the weight functions g and h. These are quite complicated to write, although simple to numerically compute, and they will be motivated in Section 6. First we write, for any two (possibly identical) functions ϕ and ψ and any integers $w\geq 1$ and $w'\in\{0,\cdots,2w\}$:

$$\left.\begin{array}{l} a(\phi,\psi)_t=\int_{1\vee t}^{1+1\wedge t}\phi(u-1)\,\psi(u-t)\,du \\ a'(\phi,\psi;w,w')_t=\sum_{r=0}^{[w'/2]}C_{w'}^{2r}\,m_{2r}m_{2w-2r}\,a(\phi,\phi)_1^{w-w'}\,a(\phi,\psi)_t^{w'-2r}\big(a(\phi,\phi)_1\,a(\psi,\psi)_1-a(\phi,\psi)_t^2\big)^r \end{array}\right\} \tag{25}$$

(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 $w\in\mathbb{N}$:

$$\left.\begin{array}{rcl} A(g,h;w)_t &=& \sum_{l,l'\in\{0,\cdots,p/2\},l+l'\leq p-w}\sum_{w'=(2w-p+2l)^+}^{(2w)\wedge(p-2l')}\rho(p)_l\,\rho(p)_{l'}\,C_{p-2l}^{2w-w'}\,C_{p-2l'}^{w'} \\ && (2\bar g'(2))^l(2\bar h'(2))^{l'}\,a'(g,h;w,w')_t\,a'(g',h';p-l-l'-w,p-2l'-w')_t \\ A'(g,h;w) &=& \int_0^2 A(g,h;w)_t\,dt-2(m_p)^2\,\bar g(2)^{p/2}\,\bar h(2)^{p/2}\,1_{\{w=p\}}. \end{array}\right\} \tag{26}$$

We also complete the notation (17) by a truncated version:

$$V^*(Y,g,q,r)_t^n=\sum_{i=0}^{[t/\Delta_n]-k_n}|\overline{Y}(g)_i^n|^q\,1_{\{|\overline{Y}(g)_i^n|\leq u_n\}}\,|\widehat{Y}(g)_i^n|^r. \tag{27}$$

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

$$M^*(g,h,\phi)_t^n=\Delta_n^{1-p/2}\sum_{w=0}^p\frac{\theta\,A'(g,h;w)}{m_{2w}\,2^{p-w}\,\bar\phi(2)^w\,\bar\phi'(2)^{p-w}}\sum_{l=0}^w\rho(2w)_l\,V^*(Z,\phi,2w-2l,p+l-w)_t^n. \tag{28}$$

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

a) The variables

$$\frac{1}{\Delta_n^{1/4}}\,\big(S(g,h,p)_n-\gamma''\big)$$

converge stably in law, in restriction to the set Ω_T^c, towards 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 $\Sigma(g,h,p,\theta)$ (an $\mathcal{F}$-measurable positive variable).

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

$$\Sigma_n \;=\; \frac{M^*(g,g,\phi;p)_T^n-2\gamma^{p/2}M^*(g,h,\phi,p)_T^n+\gamma^pM^*(h,h,\phi;p)_T^n}{\big(\gamma'\,\Delta_n^{1-p/4}\,\overline{V}(Z,h,p)_T^n\big)^2} \tag{29}$$

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

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

$$k'_n/k_n \;\to\; \infty, \qquad k'_n\Delta_n \;\to\; 0. \tag{30}$$

Recall also the truncation level u_n of (24). We choose an arbitrary weight function ϕ (it may be g or h, or some other), and we consider the variables

$$\left.\begin{array}{lcl} \eta(\phi,0)_i^n & = & \frac{1}{k_n\,k'_n\,\Delta_n}\,\sum_{j=1}^{k'_n}\big((\overline{Z}(\phi)_{i+j}^n)^2-\frac12\,\widehat{Z}(\phi)_{i+j}^n\big)1_{\{|\overline{Z}(\phi)_{i+j}^n)|\leq u_n\}} \\ \eta(\phi,1)_i^n & = & \frac{1}{k_n\,k'_n\,\Delta_n}\sum_{j=1}^{k'_n}\widehat{Z}(\phi)_{i+j}^n\,1_{\{|\overline{Z}(\phi)_{i+j}^n)|\leq u_n\}}. \end{array}\right\} \tag{31}$$

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

$$\left.\begin{array}{lcl} N(\phi,m,-)_t^n & = & \frac{1}{k_n}\,\sum_{i=k_n+k'_n}^{[t/\Delta_n]-k_n}\big(\overline{Z}(\phi)_i^n\big)^{2p-2}\,\eta(\phi,m)_{i-k_n-k'_n}^n \\ N(\phi,m,+)_t^n & = & \frac{1}{k_n}\,\sum_{i=0}^{[t/\Delta_n]-2k_n-k'_n+1}\big(\overline{Z}(\phi)_i^n\big)^{2p-2}\,\eta(\phi,m)_{i+k_n-1}^n \end{array}\right\} \tag{32}$$

On the other hand, we introduce the numbers (recall that θ and p are fixed):

$$\left.\begin{array}{lclllcl} \Gamma(-,g)_t & = & \int_t^1 g(s)^{p-1}g(s-t)ds, & & \Gamma'(-,g)_t & = & \int_t^1 g(s)^{p-1}g'(s-t)ds \\ \Gamma(+,g)_t & = & \int_0^{1-t} g(s)^{p-1}g(s+t)ds, & & \Gamma'(+,g)_t & = & \int_0^{1-t} g(s)^{p-1}g'(s+t)ds. \\ \Psi_\pm(g,h) & = & \int_0^1\Gamma(\pm,g)_t\,\Gamma(\pm,h)_t\,dt, & & \Psi'_\pm(g,h) & = & \int_0^1\Gamma'(\pm,g)_t\,\Gamma'(\pm,h)_t\,dt \\ \Psi_\pm & = & \Psi_\pm(g,g)+\Psi_\pm(h,h)-2\Psi_\pm(g,h), & & \Psi'_\pm & = & \Psi'_\pm(g,g)+\Psi'_\pm(h,h)-2\Psi'_\pm(g,h). \end{array}\right\} \tag{33}$$

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

a) The variables

$$\frac{1}{\Delta_n^{1/4}}\,\big(S(g,h,p)_n-1\big)$$

converge stably in law, in restriction to the set Ω_T^j, towards a random variable defined on en extension of the probability space $(\Omega,\mathcal{F},\mathbb{P})$ and which, conditionally on $\mathcal{F}$, is a centered variable (nor necessarily Gaussian, unless both processes σ and α have no jumps occurring at the jump times of X), and the variance of which we denote by $\Sigma'(g,h,p,\theta)$.

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

$$\Sigma'_n=\frac{\theta\,p^2\,k_n^2\,\bar\phi(p)^2}{\big(\overline{V}(Z,\phi,p)_T^n\big)^2}\Big(\frac{\Psi_-N(\phi,1,-)_T^n+\Psi_+N(\phi,1,+)_T^n}{\bar\phi(2)\,\bar\phi(2p-2)}+\frac{\Psi'_-N(\phi,0,-)_T^n+\Psi'_+N(\phi,0,+)_T^n}{2\,\bar\phi'(2)\,\bar\phi(2p-2)}\Big) \tag{34}$$

converge in probability to the $\mathcal{F}$-conditional variance $\Sigma'(g,h,p,\theta)$, 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 (13), for defining $S(g,h,p)_n$ and for defining $V^*(Z,\phi,r,s)^n_t$ in (28) or for the processes in (32): we thus have a lot of flexibility, hence also a lot of parameters to tune.

In practice, one takes $h(s) = g(sk)$ for some $k > 1$, and also $\phi = g$ 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 Σ'_n than what is mentioned in the previous remark. For example in (31) one could truncate also $\widehat{Z}(\phi)^n_{i+j}$ at the level u_n, or leave out any truncation. In (32) we could truncate $\overline{Z}(g)^n_i$ from below, that is replace $|\overline{Z}(g)^n_i|^{2p-2}$ by $|\overline{Z}(g)^n_i|^{2p-2}\, 1_{\{|\overline{Z}(g)^n_i|>u_n\}}$. The proofs are exactly the same. □

3.6. The two tests

We start with the case where the null hypothesis is "no jump", that is Ω^c_T. As before, the two weight functions g and h are given, as well as the even integer $p \geq 4$ (typically, $p = 4$ and $h(t) = g(kt)$ for some $k > 1$; recall that in any case $\gamma'' > 1$). We use the statistics $S(g,h,p)_n$ given by (22).

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

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

$$C_n = \Big\{ S(g,h,p)_n < \gamma'' - z_\alpha\, \Delta_n^{1/4} \sqrt{\Sigma_n} \Big\}, \tag{35}$$

where Σ_n is given by (29). Then the asymptotic level of the critical region (35) for testing the null hypothesis "no jump" equals α, whereas the asymptotic power is 1.

In a second case, we set the null hypothesis to be that "there are jumps", that is Ω^j_T.

Theorem 5. *We assume Assumptions 1 and 3, and we let Σ'_n be defined by (34).*

a) The asymptotic level of the critical region defined by

$$C'_n \;=\; \Big\{ S(g,h,p)_n > 1 + \Delta_n^{1/4} \sqrt{\Sigma'_n/\alpha}\, \Big\} \tag{36}$$

for testing the null hypothesis "there are jumps" is smaller than α, and the asymptotic power is 1.

b) Suppose that we restrict our attention to models in which the processes X never jumps at the same times as either σ or α. Then the same statements are true for the (bigger) critical region

$$C_n'' \;=\; \Big\{ S(g,h,p)_n > 1 + z_\alpha \Delta_n^{1/4} \sqrt{\Sigma_n'} \Big\}. \tag{37}$$

4. Simulation results

Throughout the simulations and the empirical study, we fix $p = 4$ and weigh functions $g(x) = \max\{0.5 - |x - 0.5|, 0\}$ and $h(x) = g(2x)$. We do not provide any theoretical guideline for choosing the averaging window k_n in this paper. However, intuitively, undersmoothing (using a small k_n) induces bias from (a) insufficient noise shrinking and (b) approximation error between Riemann sum and the correponding limiting integral. On the other hand, oversmoothing (using a large k_n) induces larger dependence in the moving average sequence and reduces the quality of our Central Limit Theorem. Moreover, oversmoothing tends to kill the small jumps, which reduces the power of our test. In the simulation, we use different values of k_n for sensitivity analysis.

We use the same averaging window for the variance estimator as the window for $S(g,h,p)_n$, i.e., $k_n^* = k_n$, although this is not required by the asymptotic theory. We also link the truncation level u_n with the window k_n^* by $u_n = C(\bar{g}(2)\overline{\sigma^2})^{1/2}\Delta_n^{0.47}(k_n^*)^{1/2}$, where $\overline{\sigma^2}$ is the mean squared volatility and $C = 7$. These two restrictions make k_n the only tuning parameter in our simulation. The results are robust to deviations from this benchmark[1].

We start with a simulation using the same data generating process as in Figure 1 in order to compare our test with the AJ test. Since both the volatility $(\sigma_t)_{t\geq 0}$ and the standard deviation of noise $(\alpha_t)_{t\geq 0}$ are constant, this simple data generating process also helps us identifying potential problems of our test. Table 1 shows the results under the null hypothesis of no jumps with different levels of noises. The test statistic $S(g,h,4)_n$ is properly centered at 2 (column 3), which confirms the first-order property predicted by (23). The rejection rate in the simulation is between 4.2% and 6.5% for a 5%-level test (column 4) and is between 8.7% and 12.2% for a 10%-level test (column 5).

[1] To check robustness, we use $k_n^* = 1.5k_n$ and $k_n^* = 2k_n$. We also use $C = 6$ and $C = 8$ to determine the truncation level. The results are similar to those in the text and thus omitted to save space.

In most cases, the test slightly over-rejects. An interesting feature is that the rejection rate increases with the averaging window k_n in all cases. Figure 2 compares the Monte Carlo distribution of the standardized statistic with the asymptotic distribution.

Figure 3 plots the distribution of the non-standardized test statistic $S(g, h, 4)_n$ to illustrate how the test statistic separates continuous paths from discontinuous ones. When there is no noise (top row), the distribution of $S(g, h, 4)_n$ exhibits a similar pattern as in Figure 1, but with much larger dispersion. However, the distribution of $S(g, h, 4)_n$ is insensitive to noise (middle and bottom rows), which is a clear advantage over the AJ test statistic.

Figure 4 plots the power of the test under the null hypothesis of no jumps when the underlying process X contains compound Poisson jumps with different intensities but constant expected quadratic variation. When the jumps are infrequent with large size ($\lambda = 1$), the null hypothesis is almost always rejected. As the jumps become frequent and small, the rejection rate drops. Large averaging window reduces the power of the test, since the small jumps tend to be smoothed away.

We then investigate our test in a more realistic setting, using the same observation length and sampling frequency as in the previous simulations. We generate the underlying process X using a stochastic volatility model which is calibrated to a liquid stock trading on the NYSE. We also allow noise level α_t to be time-varying by marking it to σ_t so that the signal-to-noise ratio is kept constant. While the noise satisfies Assumption 3, they are unconditional dependent and dependent on X. The results are quite similar to what we find in the previous simulations with constant volatility and noise level. Table 2 reports the mean of the non-standardized test statistic and the rejection rate under the null hypothesis of no jumps. Figure 5 compares the Monte Carlo distribution of the standardized test statistic with the theoretical asymptotic distribution. Figure 6 shows how the non-standardized test statistic separates continuous paths from discontinuous paths. Figure 7 plots the power of the test.

It may appear counterintuitive that as k_n decreases, the type 1 error (Tables 1 and 2) and the type 2 error (Figures 4 and 7) drop simultaneously. This phenomenon may arise from our reduction of three tuning parameters – the averaging window for the test statistic, the averaging window for the variance estimator and the truncation level u_n – into a single parameter k_n as described at the beginning of this section. Consequently, the higher order effect of k_n on size and power can be very complicated. One caveat is that in practice, a researcher should not simply take a small window k_n with the attempt to minimize both type 1 and type 2 errors based on this phenomenon. When the window is too small relative to the magnitude of noise, the bias correction may be insufficient. Indeed, the upward bias of the standardized statistic is visible in Figure 5 when the noise is large (bottom row) and the window is small (left column) [2]. Although in our simulation such bias seems to help reducing the type 1 error of our one-sided test, this effect is beyond the scope of our asymptotic theory. The choice of k_n involves higher-order asymptotic theory, which is left for future research.

5. Empirical results

We now conduct the test for each of the current (October 29th, 2009) 30 Dow Jones Industrial Average (DJIA) stocks and each trading day in 2008; the data source is the TAQ database. Each day, we collect all transactions from 9:30am until 4:00pm, 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 seconds. 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 8 distributions of the non-standardized and the standard statistic under the null hypothesis of no jumps for different averaging windows, with the tuning parameters k_n, k_n^* and u_n chosen in the same way as in our simulations. The empirical distributions of these statistics are quite robust to the choice of k_n and provide evidence for the presence of jumps. Indeed, the non-standardized statistic is centered away from 2 (left column) and the distribution of the standardized statistic deviates significantly from the $\mathcal{N}(0,1)$ distribution towards the left side (or, the "jump side"). When $k_n = 100$ (middle row), the null hypothesis of no jumps is rejected 30% of the time for a 5%-level test and 41% of the time for a 10%-level test. Changing k_n to 75 or 125 affects these rejection rates by less than 1 percentage point.

[2] In the simulations of either Figure 2 or 5, if we take the window $k_n = 50$, the standardized statistic is significantly upward biased when there is noise. The bias increases with the size of noise. We omit this case in the text for the clarity of presentation.

6. The proofs.

6.1. *Some known results and their consequences.*

In this subsection we recall some properties 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 (13), are also fixed.

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

$$\frac{1}{k_n\,\bar{g}(p)}\,\overline{V}(Z,g,p)^n_T \ \stackrel{\mathbb{P}}{\longrightarrow}\ U(p) \ :=\ \sum_{s\leq T}|\Delta X_s|^p \tag{38}$$

and also, when further X is continuous:

$$\Delta_n^{1-p/4}\,\overline{V}(Z,g,p)^n_T \ \stackrel{\mathbb{P}}{\longrightarrow}\ V(g,p) := m_p\,\theta^{p/2}\,\bar{g}(2)^{p/2}\int_0^T |\sigma_s|^p\,ds. \tag{39}$$

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 (23) readily follows from (38). For the second convergence in (23) we cannot apply (39) right away, because X is not necessarily continuous, even though it is so on Ω^c_T. We set $S=\inf(t:\Delta X_t\neq 0)$ 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 (39) applied to X' that

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

Then in view of the definition of $S(g,h,p)_n$, the second convergence in (23) 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 g, we define the following Wiener integral processes

$$L(g)_t = \int g(s-t)\,dW^1_s,\quad L'(g)_t = \int g'(s-t)\,dW^2_s, \tag{40}$$

and $L(h)$ and $L'(h)$ are defined likewise with 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:

$$\left.\begin{array}{lcl} m_q(g;\eta,\zeta) &=& \mathbb{E}'\big((\eta L(g)_1 + \zeta L'(g)_1)^q\big) \\ m_{q,q'}(g,h;\eta,\zeta)_t &=& \mathbb{E}'\big((\eta L(g)_1 + \zeta L'(g)_1)^q\,(\eta L(h)_t + \zeta L'(h)_t)^{q'}\big) \\ \mu(g,h;\eta,\zeta) &=& \sum_{r,r'=0}^{p/2} \rho(p)_r \rho(p)_{r'} \left(2\zeta^2 \overline{g}'(2)\right)^r \left(2\zeta^2 \overline{h}'(2)\right)^{r'} \\ && \int_0^2 \Big(m_{p-2r,p-2r'}(g,h;\eta,\zeta)_t - m_{p-2r}(g;\eta,\zeta)\, m_{p-2r'}(h;\eta,\zeta)\Big)\, dt \\ R(g,h) &=& \theta^{1-p} \int_0^T \mu(g,h;\theta\sigma_s,\alpha_s)\, ds. \end{array}\right\} \tag{41}$$

With all this notation, we then have

$$\left(\frac{\Delta_n^{1-p/4}\, \overline{V}(Z,g,p)_T^n - V(g,p)}{\Delta_n^{1/4}}, \frac{\Delta_n^{1-p/4}\, \overline{V}(Z,h,p)_T^n - V(h,p)}{\Delta_n^{1/4}}\right) \stackrel{\mathcal{L}-(s)}{\longrightarrow} (\overline{V}(g,p), \overline{V}(h,p)) \tag{42}$$

(stable convergence in law), 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_T^c 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*

$$\Sigma(g,h,p,\theta) \;=\; \frac{R(g,g) - 2\gamma^{p/2}\, R(g,h) + \gamma^p\, R(h,h)}{\gamma'^2\, V(h,p)^2} \tag{43}$$

Finally we consider the CLT associated with the convergence (38). Under Assumptions 1 and 3,

$$\left(\frac{\overline{V}(Z,g,p)_T^n - k_n \overline{g}(p)\, U(p)}{k_n\, \Delta_n^{1/4}}, \frac{\overline{V}(Z,h,p)_T^n - k_n \overline{h}(p)\, U(p)}{k_n\, \Delta_n^{1/4}}\right) \stackrel{\mathcal{L}-(s)}{\longrightarrow} (\overline{U}(g,p), \overline{U}(h,p)) \tag{44}$$

where $(\overline{U}(g,p), \overline{U}(h,p))$ is defined on an extension of the space and, conditionally on $\mathcal{F}$, is a (usually not Gaussian, unless the processes σ and α do not jump at the same times as X) centered random vector with covariance matrix $\begin{pmatrix} D(g,g) & D(g,h) \\ D(g,h) & D(h,h) \end{pmatrix}$, where

$$D(g,h)_t = p^2\theta \sum_{s\le t} |\Delta X_s|^{2p-2} \left(\Psi_-(g,h)\sigma_{s-}^2 + \Psi_+(g,h)\sigma_s^2 + \frac{1}{\theta^2}\, \Psi'_-(g,h)\alpha_{s-}^2 + \frac{1}{\theta^2}\, \Psi'_+(g,h)\alpha_s^2\right) \tag{45}$$

(notation (33)). Therefore, exactly as for Proposition 1, we get

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

$$\Sigma'(g,h,p,\theta) \;=\; \frac{D(g,g)-2D(g,h)+D(h,h)}{U(p)^2} \tag{46}$$

Now we define Σ_n and Σ'_n by (29) and (34). Suppose for a moment that we have the following behavior:

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

$$\Sigma_n \stackrel{\mathbb{P}}{\longrightarrow} \Sigma(g,h,p,\theta) \qquad \text{on the set } \Omega_T^c. \tag{47}$$

$$\Delta_n^{1/2}\Sigma_n \stackrel{\mathbb{P}}{\longrightarrow} 0 \qquad \text{on the set } \Omega. \tag{48}$$

b) Under the assumptions of Theorem 3 we have

$$\Sigma'_n \stackrel{\mathbb{P}}{\longrightarrow} \Sigma'(g,h,p,\theta) \qquad \text{on the set } \Omega_T^j. \tag{49}$$

$$\Delta_n^{1/2}\Sigma'_n \stackrel{\mathbb{P}}{\longrightarrow} 0 \qquad \text{on the set } \Omega. \tag{50}$$

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 $T_n = (S(g,h,p)_n - \gamma'')/\Delta_n^{1/4}\sqrt{\Sigma_n}$ converge stably in law, in restriction to the set Ω_T^c 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 (23) and from (48).

For Theorem 5 the proof is exactly the same for the power of the two tests (use the second convergence in (23) and (50)), and also for the level in claim (b) because the variables $T'_n = (S(g,h,p)_n - 1)/\Delta_n^{1/4}\sqrt{\Sigma'_n}$ converge stably in law, in restriction to the set Ω_T^j to an $\mathcal{N}(0,1)$ random variable. For (a) we use the fact that T'_n above converges stably in law, on the set Ω_T^j again, to a centered variable with variance 1, and we use the Bienaymé-Tchebycheff inequality to conclude. $\square$

6.2. Preliminaries.

We first derive some estimates for the variables $\overline{Z}(g)_i^n$ and $\widehat{Z}(g)_i^n$, 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 (24), 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*:*

$$\mathbb{E}(|\overline{Z}(g)^n_{i_n}|^q) \;\leq\; K_q\big(\Delta_n^{q/4}+L^{(q-2)^+}\Delta_n^{(q/4)\wedge(1/2)}\big), \qquad \mathbb{E}(|\widehat{Z}(g)^n_{i_n}|^q) \;\leq\; K_q\Delta_n^{q/2}. \tag{51}$$

$$\mathbb{E}\big(|\overline{X}(g)^n_{i_n+j}-\sigma_{T_n}\overline{W}(g)^n_{i_n+j}|^2\wedge u_n^2\big) \;\leq\; K\Delta_n+K\Delta_n^{1/2}\,\rho_n+K\mathbb{E}\Big(\int_{T_n+(j-1)\Delta_n}^{T_n+(k_n+j-1)\Delta_n}|\sigma_s-\sigma_{T_n}|^2\,ds\Big) \tag{52}$$

$$\mathbb{E}\big(|\widehat{Z}(g)^n_{i_n}-\widehat{\chi}(g)^n_{i_n}|^q\big) \;\leq\; K\big(\Delta_n^{q/2+q\wedge 1}+\Delta_n^{3q/4+(q/2)\wedge 1}\big). \tag{53}$$

Proof. The second part of (51) 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

$$\overline{X}(g)^n_i \;=\; \int_{i\Delta_n}^{(i+k_n)\Delta_n} g_n(s-i\Delta_n)\,dX_s, \tag{54}$$

where $g_n(s)=\sum_{j=1}^{k_n-1}g_j^n\,1_{((j-1)\Delta_n,j\Delta_n]}(s)$. 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 (52) we decompose X as $X=X'+X''$, where $X'_t=\int_0^t b_s\,ds+\int_0^t\sigma_s\,dW_s$. (54) yields

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

The expectation of the squared integral above is smaller than $K\Delta_n+\mathbb{E}\big(\int_{T_n+j\Delta_n}^{T_n+(k_n+j)\Delta_n}|\sigma_s-\sigma_{T_n}|^2ds\big)$, 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 (52).

Next, observe that $\widehat{Z}(g)^n_{i_n} - \widehat{\chi}(g)^n_{i_n} = \widehat{X}(g)^n_{i_n} + a_n + a'_n$, where

$$\begin{aligned} a_n &= \textstyle\sum_{j=1}^{[k_n/2]} (g'^n_{2j})^2 \big(\Delta^n_{i_n+2j-1} X \Delta^n_{i_n+2j-1}\chi + \Delta^n_{i_n+2j-1}\chi \Delta^n_{i_n+2j-1} X\big) \\ a'_n &= \textstyle\sum_{j=0}^{[(k_n-1)/2]} (g'^n_{2j+1})^2 \big(\Delta^n_{i_n+2j} X \Delta^n_{i_n+2j}\chi + \Delta^n_{i_n+2j}\chi \Delta^n_{i_n+2j} X\big). \end{aligned}$$

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 (51), we obtain (53). $\square$

Our second preliminary concerns the behavior of the truncated variations $V^*(Z,g,q,r)^n$ of (27):

Lemma 2. *Suppose that Assumptions 4 and 5 and (24) hold, and also that X is continuous. Then if q is an even integer and $w \in \{0, \cdots, q/2\}$, we have*

$$\Delta_n^{1-q/2} V^*(Z, g, q-2w, w)^n_t \stackrel{u.c.p.}{\longrightarrow} \theta^{-q} \int_0^t (2\alpha_s^2 \overline{g}'(2))^w m_{q-2w}(g; \theta\sigma_s, \alpha_s)\, ds. \tag{55}$$

Proof. A classical localization procedure allows to suppose the strengthened Assumptions 4 and 5. We can reproduce the proof of Theorem 3.4 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

$$\Delta_n^{1-q/2} V(Z, g, q-2w, w)^n_t \stackrel{u.c.p.}{\longrightarrow} \theta^{-q} \int_0^t (2\alpha_s^2 \overline{g}'(2))^w m_{q-2w}(g; \theta\sigma_s, \alpha_s)\, ds.$$

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 $\mathbb{P}(\Omega_n) \to 1$ as $n \to \infty$. The first part of (51) applied with $q = \frac{2}{1-4\varpi}$ 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 $i = 1, \cdots, [t/\Delta_n]$ satisfies all our requirements. $\square$

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 (30). 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:

$$\begin{array}{ll} \text{case (1):} \ T_n \to T \quad \text{and} \quad T'_n \leq T \ \text{ for all } n \\ \text{case (2):} \ T'_n \to T \quad \text{and} \quad T_n \geq T \ \text{ for all } n. \end{array} \tag{56}$$

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

$$\left.\begin{array}{lll} G(g,0) = \bar{g}(2)\sigma_{T-}^2 + \frac{1}{\theta^2}\,\bar{g}'(2)\alpha_{T-}^2, & G(g,1) = \frac{2}{\theta^2}\,\bar{g}'(2)\alpha_{T-} & \text{in case (1)} \\ G(g,0) = \bar{g}(2)\sigma_{T}^2 + \frac{1}{\theta^2}\,\bar{g}'(2)\alpha_{T}^2, & G(g,1) = \frac{2}{\theta^2}\,\bar{g}'(2)\alpha_{T} & \text{in case (2)} \end{array}\right\} \tag{57}$$

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

$$\frac{1}{k_n\,k'_n\,\Delta_n}\big(V^*(Z,g,2-2w,w)^n_{T'_n+k_n\Delta_n} - V^*(Z,g,2-2w,w)^n_{T_n+k_n\Delta_n}\big) \stackrel{\mathbb{P}}{\longrightarrow} G(g,w). \tag{58}$$

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 (55) 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

$$\beta^n_i \;=\; \sigma_{i_n\Delta_n}\overline{W}(g)^n_{i_n+i}, \quad \beta'^n_i \;=\; \overline{\chi}(g)^n_{i_n+i}, \quad \widehat{\beta}^n_i \;=\; \widehat{\chi}(g)^n_{i_n+i}.$$

In this step we prove that

$$H(w)_n \;=\; \frac{1}{k_n\,k'_n\,\Delta_n}\,\mathbb{E}\Big(\sum_{i=1}^{k'_n}\Big|f_w(\overline{Z}(g)^n_{i_n+i},\widehat{Z}(g)^n_{i_n+i})\,1_{\{|\overline{Z}(g)^n_{i_n+i}|\le u_n\}} - f_w(\beta^n_i+\beta'^n_i,\widehat{\beta}^n_i)\Big|\Big) \;\to\; 0. \tag{59}$$

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 $x=\beta^n_i+\beta'^n_i$ and $y=\widehat{\beta}^n_i$ and $x'=\overline{X}(g)^n_{i_n+i}-\beta^n_i$ and $y'=\widehat{Z}(g)^n_{i_n+i}-y$ and apply (53) for $q=1$ and (51) and (52), plus the Cauchy-Schwarz inequality, to get

$$H(w)_n \;\le\; \frac{K\Delta_n^{1/2}}{k_n\Delta_n}\,\Big(\Delta_n^{3/4-\varpi}+\varepsilon+\frac{\rho_n}{\varepsilon}+\Delta_n^{1/2-2\varpi}+\rho'_n\Big),$$

where

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

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 $\limsup_n H()_n\le K\varepsilon$, and (59) 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 (59), it remains to prove that

$$\frac{1}{k_n\, k_n'\, \Delta_n} \sum_{i=1}^{k_n'} \zeta_i^n \ \stackrel{\mathbb{P}}{\longrightarrow}\ G(g,w). \tag{60}$$

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 $i \geq 1$ and ζ_i^n is $\mathcal{F}_{(i_n+i+k_n-1)\Delta_n}$-measurable, we have

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

(51) yields $\mathbb{E}(|\zeta_i^n|^2) \leq K\Delta_n$, so the right side above goes to 0 because $k_n/l_n \to 0$, and instead of (60) it is then enough to prove that

$$\frac{1}{k_n\, k_n'\, \Delta_n} \sum_{i=1}^{k_n'} \zeta_i'^n \ \stackrel{\mathbb{P}}{\longrightarrow}\ G(g,w). \tag{61}$$

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 (**??**) 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 $i = 0, \cdots, l_n$. Using (13) once more, (61) follows. □

6.3. The behavior of Σ_n.

Here we prove the first claim (a) of Proposition 3. For this, we begin by showing that $M^*(g,h,\phi)_T^n$ 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 $M^*(g,h,p)_t^n$ behaves better, at least when $p = 4$. We start with some calculations:

Lemma 4. *With the notation (26), we have*

$$\mu(g,h;\eta,\zeta) \ =\ \sum_{w=0}^{p} \eta^{2w}\, \zeta^{2p-2w}\, A'(g,h;w). \tag{62}$$

Proof. Consider the processes $L(g)$ and $L'(g)$ defined by (40). We use the notation (25). First, we have $\mathbb{E}'(L(g)_1\, L(h)_t) = a(g,h)_t$, and the process $(L(g), L(h))$ is stationary Gaussian centered. Then a well known fact about 2-dimensional centered Gaussian vectors yields that for w a nonnegative integer and $w' \in \{0, \cdots, w\}$:

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

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

$$m_q(g;\eta,\zeta) \;=\; \sum_{l=0}^{q/2} C_q^{2l}\, \eta^{2l} \zeta^{q-2l}\, m_{2l}\, m_{q-2l}\, (\bar{g}(2))^l\, (\bar{g}'(2))^{q/2-l} \tag{63}$$

$$\begin{aligned} m_{q,q'}(g,h;\eta,\zeta)_t \;&=\; \textstyle\sum_{l=0}^{q} \sum_{l'=0}^{q'} C_q^{2l}\, C_{q'}^{q'-2l'}\, \eta^{2l+2l'}\, \zeta^{q+q'-2l-2l'} \\ &\qquad a'(g,h;l/2+l'/2,l')_t\, a'(g',h';q/2+q'/2-l/2-l'/2,q'-l')_t \\ &=\; \textstyle\sum_{w=0}^{q/2+q'/2} \eta^{2w} \zeta^{q+q'-2w} \sum_{w'=(2w-q)^+}^{q'\wedge(2w)} C_q^{2w-w'}\, C_{q'}^{w'}\, \eta^{2l+2l'}\, \zeta^{q+q'-2l-2l'} \\ &\qquad a'(g,h;w,w')_t\, a'(g',h';q/2+q'/2-w,q'-w')_t. \end{aligned}$$

Using (18), we first deduce that

$$\sum_{r=0}^{p/2} \rho(p)_r (2\zeta^2 \bar{g}'(2))^r\, m_{p-2r}(g;\eta,\zeta) \;=\; m_p\, \eta^p\, (\bar{g}(2))^{p/2}. \tag{64}$$

Then in view of (26), we end up with (62). □

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

$$M^*(g,h,\phi)_T^n \;\stackrel{\mathbb{P}}{\longrightarrow}\; R(g,h) \qquad \text{on the set } \Omega_T^c. \tag{65}$$

$$\Delta_n^{p/2-1/2}\, M^*(g,h,\phi)_T^n \;\stackrel{\mathbb{P}}{\longrightarrow}\; 0. \tag{66}$$

Proof. For (65) it is enough, by the same argument as in Theorem 1, to consider the case when X is continuous. Then we can apply (55) with $q = p$ and w substituted with $p + l - w$ and sum over l between 0 and w: taking advantage of (64) with $2w$ instead of p, we readily deduce

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

At this stage we readily deduce the result from (41), (62) and the definition (28).

Now we turn to (66). Clearly, it is enough to show that $\sqrt{\Delta_n} V^*(Z, \phi, 2p-2l, l)_t^n \stackrel{\mathbb{P}}{\longrightarrow} 0$ for each $l \in \{0, \dots, p\}$. When $l = p$, by the second part of (51) with $q = p$,

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

which implies the result. When $l = 0$, by the first part of (51) with $q = 2$,

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

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

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

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

Proof of Proposition 3-(a). (47) is a straightforward consequence of (39) and (65), whereas (48) readily follows from (38) and (66). □

6.4. *The behavior of Σ_n'.*

Now we turn to the behavior of Σ_n', 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 processes which enter the definition (45) of $D(g, h)$, which are

$$\left.\begin{array}{lcl} N(\phi, 0, -)_t &=& \bar{\phi}(2)\, \bar{\phi}(2p-2) \sum_{s \leq t} \sigma_{s-}^2\, |\Delta X_s|^{2p-2} \\ N(\phi, 0, +)_t &=& \bar{\phi}(2)\, \bar{\phi}(2p-2) \sum_{s \leq t} \sigma_s^2\, |\Delta X_s|^{2p-2} \\ N(\phi, 1, -)_t &=& \frac{2}{\theta^2}\, \bar{\phi}'(2)\, \bar{\phi}(2p-2) \sum_{s \leq t} \alpha_{s-}^2\, |\Delta X_s|^{2p-2} \\ N(\phi, 1, +)_t &=& \frac{2}{\theta^2}\, \bar{\phi}'(2)\, \bar{\phi}(2p-2) \sum_{s \leq t} \alpha_s^2\, |\Delta X_s|^{2p-2} \end{array}\right\} \tag{67}$$

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

$$N(\phi, m, \pm)_T^n \;\stackrel{\mathbb{P}}{\longrightarrow}\; N(\phi, m, \pm)_T. \tag{68}$$

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 $(\underline{\mu}([0,t] \times \{z : \gamma(z) > \varepsilon\}) : t \geq 0)$, with the convention $S_0 = 0$. Let $i(n,q)$ be the random integer such that $(i(n,q)-1)\Delta_n < S_q \leq i(n,q)\Delta_n$. Our aim is to prove that for $m = 0,1$ and all $q \geq 1$ we have

$$\eta(\phi,m)^n_{i(n,q)-k_n-k'_n} \stackrel{\mathbb{P}}{\longrightarrow} \overline{\eta}_q(m) \;=\; \begin{cases} \bar{\phi}(2)\,\sigma^2_{S_q-} & \text{if } m=0 \\ \frac{2}{\theta^2}\,\bar{\phi}'(2)\,\alpha^2_{S_q-} & \text{if } m=1 \end{cases} \tag{69}$$

For this, consider the processes

$$X(\varepsilon)_t \;=\; X_t - \sum_{q\geq 1} \Delta X_{S_q}\, 1_{\{S_q \leq t\}}, \qquad Z(\varepsilon)_t \;=\; X(\varepsilon)_t + \chi_t, \tag{70}$$

and denote by $\eta'(\phi,m)^n_i$ the variables defined by (31), 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 (59) 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 (58), written for $Z' = X\varepsilon) + \chi$ instead of $Z = X + \chi$, 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 (69) follows from the convergence (58), 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 (69) is proved.

Step 2) This step is devoted to some estimates, in the same setting as in the previous step. We denote by Ω_n the set on which all $S_q \leq T$ satisfy $S_{q-1} + (2k_n + k'_n)\Delta_n < S_q < T - k_n\Delta_n$, so $\mathbb{P}(\Omega_n) \to 1$. We also set

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

and use the notation $\rho'(m)^n_i$ if we substitute Z with $Z(\varepsilon)$. Using the initial filtration, and also the big filtration $(\mathcal{G}_t)$ and the fact that if $1 \leq j \leq k_n$ we have $\eta'(\phi,m)^n_{i_n+j} = \eta(\phi,m)_{i_n+j}$ when we are on the set Ω_n and $S_q \leq T$ and $j = 1, \cdots, k_n$, we deduce from (51) that for any $i \geq 1$ and $j = 1, \cdots, k_n + k'_n$,

$$\mathbb{E}(|\rho(m)_i^n|) \ \leq \ K\Delta_n^{1/2}, \qquad \mathbb{E}\big(|\rho(m)_{i_n+j}^n| \, 1_{\{S_q \leq T\} \cap \Omega_n}\big) \ \leq \ K\Delta_n^{1/2}. \tag{71}$$

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

$$\mathbb{E}(|\eta(\phi, m)_i^n|) \ \leq \ K, \qquad \mathbb{E}\big(|\eta(\phi, m)_{i_n+j}^n| \, 1_{\{S_q \leq T\} \cap \Omega_n}\big) \ \leq \ K. \tag{72}$$

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

$$\eta(\phi, m)_{i_n-j}^n - \eta(\phi, m)_{i_n}^n \ = \ \frac{1}{k_n \, k'_n \, \Delta_n}\Big(\sum_{l=1-j}^{0} \rho(m)_{i_n+l}^n - \sum_{l=k'_n-j+1}^{k'_n} \rho(m)_{i_n+l}^n\Big),$$

and each sum above has at most k_n summands. Therefore (71) yields

$$\mathbb{E}\big(|\eta(\phi, m)_{i_n-j}^n - \eta(\phi, m)_{i_n}^n| \, 1_{\{S_q \leq T\} \cap \Omega_n}\big) \ \leq \ K k_n / k'_n.$$

Since $k_n / k'_n \to 0$ and $\mathbb{P}(\Omega_n) \to 1$, we then deduce from (69) that

$$\frac{1}{k_n} \sum_{j=1}^{k_n} \eta(\phi, m)_{i-k_n-k'_n-j}^n \, |\phi_j^n|^{2p-2} \ \xrightarrow{\mathbb{P}} \ \bar{\phi}(2p-1) \, \overline{\eta}_q(m). \tag{73}$$

Finally, since $p \geq 4$ and since (51) is also true if we take the conditional expectation with respect to $\mathcal{G}_{i_n \Delta_n}$ when we consider $Z(\varepsilon)$, that is

$$\mathbb{E}\big(|\overline{Z(\varepsilon)}(\phi)_i^n|^{2p-2} \mid \mathcal{G}_{i\Delta_n}\big) \ \leq \ K\big(\Delta_n^{p/2-1/2} + \varepsilon^{2p-4} \Delta_n^{1/2}\big) \ \leq \ K\Delta_n^{1/2}. \tag{74}$$

Step 3) Recall that the variables in (69) implicitly depend on ε, and set

$$\begin{aligned} N(m, \varepsilon)_t \ &= \ \bar{\phi}(2p-2) \textstyle\sum_{q: S_q \leq t} \overline{\eta}_q(m) \, |\Delta X_{S_p}|^{2p-2} \\ B(m, \varepsilon)_t^n \ &= \ \frac{1}{k_n} \textstyle\sum_{i=k_n+k'_n}^{[t/\Delta_n]-k_n} \eta(\phi, m)_{i-k_n-k'_n}^n \, |\overline{Z(\varepsilon)}(\phi)_i^n|^{2p-2} \\ N(m, \varepsilon)_t^n \ &= \ N(\phi, m, -)_t^n - B(m, \varepsilon)_t^n. \end{aligned}$$

By virtue of (67) and (69) and the dominated convergence theorem, we have

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

Observing that $\eta(\phi, m)_{i-k_n-k'_n}^n$ is $\mathcal{F}_{i\Delta_n}$-measurable, by successive conditioning we deduce from (72) and (74) that

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

which implies

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

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

$$N(m,\varepsilon)^n_T\ \stackrel{\mathbb{P}}{\longrightarrow}\ N(m,\varepsilon)_T. \tag{75}$$

Step 4) Now we proceed to proving (75). 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

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

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 (75) will follow if we prove $\xi(m)^n_q\ \stackrel{\mathbb{P}}{\longrightarrow}\ \bar{\phi}(2p-2)\overline{\eta}_q(m)\,|\Delta X_{S_q}|^{2p-2}$. In view of (73), it thus remains to prove that

$$\xi'(m)^n_q\ =\frac{1}{k_n}\sum_{j=1}^{k_n}\eta(\phi,m)^n_{i(n,q)-k_n-k'_n-j}\,\big(|\overline{Z}(\phi)^n_{i(n,q)-j}|^{2p-2}-|\overline{Z(\varepsilon)}(\phi)^n_{i(n,q)-j}|^{2p-2}-|\phi^n_j\,\Delta X_{S_q}|^{2p-2}\big)\ \stackrel{\mathbb{P}}{\longrightarrow}\ 0$$

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

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

and since $|w^n_j|\leq K$, and upon using (74), it follows that for $1\leq j\leq k_n$ we have

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

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 (72) that

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

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

Proof of Proposition 3-(b). (49) is a straightforward consequence of (38) and (68), because $D(g,g)-2D(g,h)+D(h,h)$ is equal to

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

As for (50), it needs to be proved on the set Ω_T^c only (on Ω_T^j it follows from (49)). So by our usual argument we can assume that X is continuous. Then (39) shows that it is enough to prove that

$$\Delta_n^{3/2-p/2}\, N(\phi,m,\pm)_T \ \stackrel{\mathbb{P}}{\longrightarrow}\ 0 \tag{76}$$

for $m=0,1$. We can again suppose Assumptions 4 and 5, by localization. Then if we combine (51) (with $L=0$ because X is continuous) and (72), we readily obtain that $\mathbb{E}(|N(\phi,m,\pm)_T|) \leq KT\Delta_n^{p/2-1}$, and thus (76) holds. □

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., Jacod, J., 2009. 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.

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. Journal of Finance 58, 2581–2610.

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, forthcoming .

Jacod, J., Shiryaev, A. N., 2003. Limit Theorems for Stochastic Processes, 2nd Edition. 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.

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.

Std. Dev. of Noise	$\mathbf{k_n}$	Mean Value of $\mathbf{S(g,h,4)_n}$	Rejection Rate 5%	10%
$\alpha = 0$	75	2.011	0.050	0.101
	100	2.003	0.058	0.115
	125	2.001	0.059	0.116
$\alpha = \sigma\sqrt{5\text{ sec}}$	75	2.013	0.054	0.105
	100	2.004	0.061	0.115
	125	2.003	0.065	0.114
$\alpha = 2\sigma\sqrt{5\text{ sec}}$	75	2.018	0.042	0.087
	100	2.001	0.055	0.107
	125	1.995	0.061	0.122

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 continuous underlying process X is generated using the same model as in Figure 1. The truncation level is $u_n = 7(\bar{g}(2))^{1/2}\sigma\Delta_n^{0.47}k_n^{1/2}$ in the estimation of variance.

Std. Dev. of Noise	$\mathbf{k_n}$	Mean Value of $\mathbf{S(g,h,4)_n}$	Rejection Rate 5%	10%
$\alpha_t = 0$	75	2.009	0.054	0.100
	100	1.999	0.064	0.117
	125	1.999	0.069	0.125
$\alpha_t = \sigma_t\sqrt{5 \text{ sec}}$	75	2.010	0.052	0.098
	100	2.000	0.064	0.118
	125	1.999	0.068	0.127
$\alpha_t = 2\sigma_t\sqrt{5 \text{ sec}}$	75	2.028	0.043	0.085
	100	2.004	0.059	0.115
	125	2.000	0.065	0.123

Table 2. **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, when there is no noise (top row), the time-varying standard deviation of noise α_t is equal to (middle row) or twice as large as (bottom row) the standard deviation of the increment of the continuous martingale part over 5 seconds. The data generating process is the stochastic volatility model $d\log(X_t) = \sigma_t\, dW_t$, with $\sigma_t = v_t^{1/2}$, $dv_t = \kappa(\beta - v_t)\, dt + \gamma v_t^{1/2}\, dB_t$, $E[dW_t\, dB_t] = \rho\, dt$, $\beta^{1/2} = 0.4$, $\gamma = 0.5$, $\kappa = 5$, $\rho = -0.5$. The parameter values are realistic for a stock based on the evidence reported in Aït-Sahalia and Kimmel (2007). The observed price Z_t is generated by $\log(Z_t) = \log(X_t) + \chi_t$, where the noise χ_t is drawn independently from a $t(2.5)$ distribution, properly scaled so that it has standard deviation α_t, and then trimmed with threshold being $100\sigma_t\sqrt{5 \text{ second}}$. We compute the test statistic using $\log(Z_t)$. The truncation level is $u_n = 7(\bar{g}(2)\beta)^{1/2}\Delta_n^{0.47}k_n^{1/2}$ in the estimation of variance.

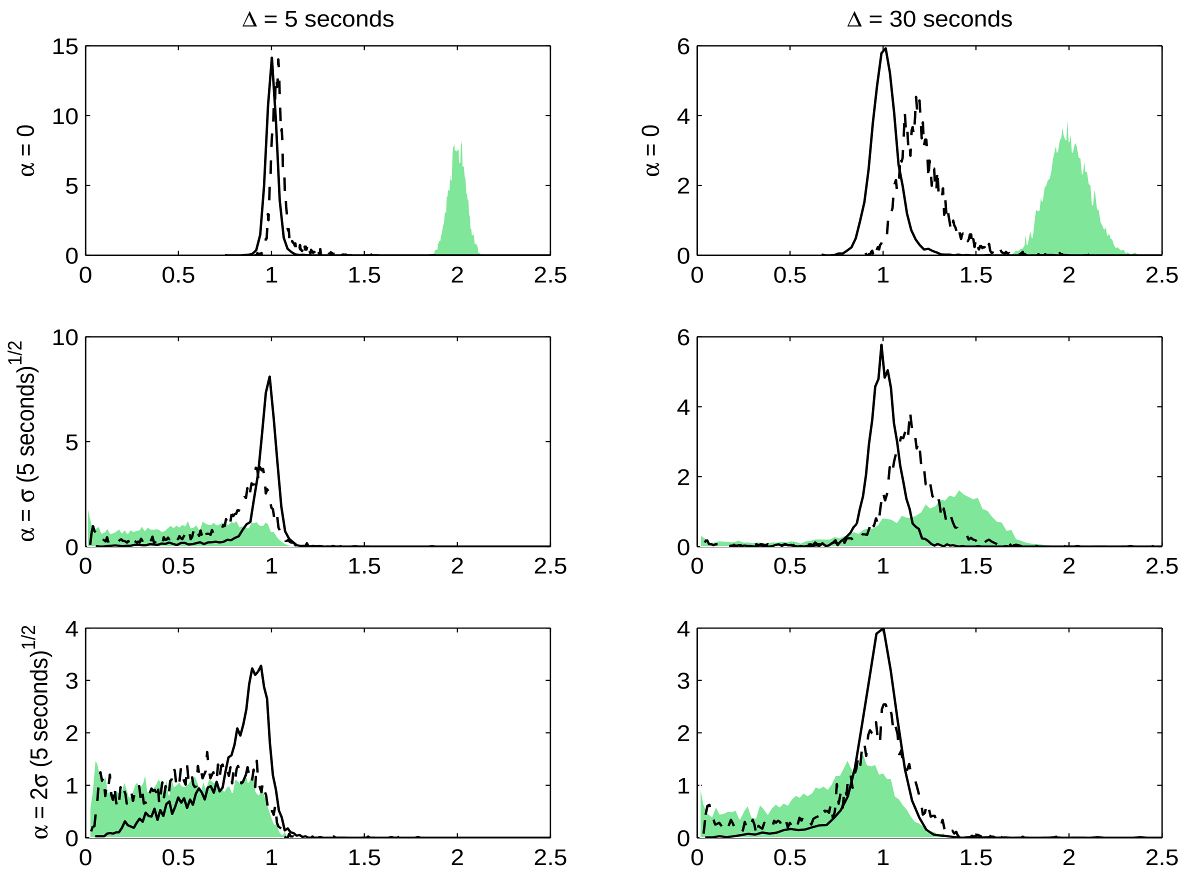

Fig. 1. Simulations of the uncorrected AJ test statistic. The noise level increases from the top to the bottom row; the sampling frequency decreases from the left to the right column.

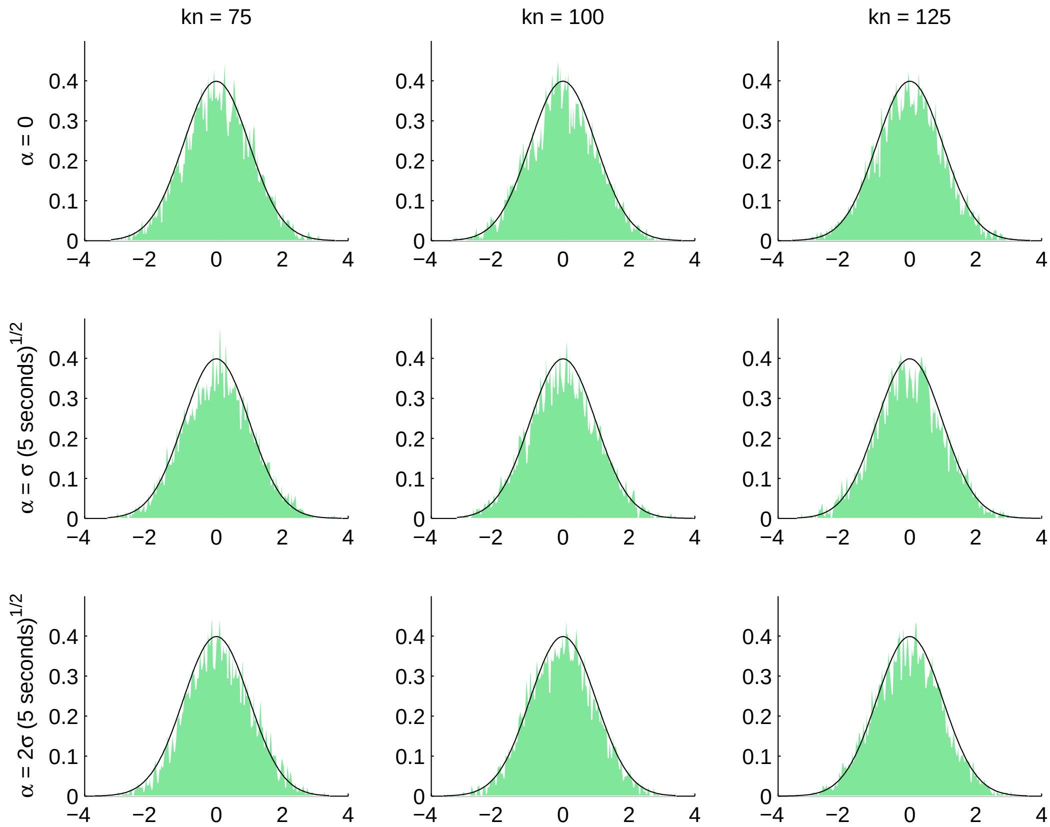

Fig. 2. Monte Carlo (5,000 simulations) and theoretical asymptotic distributions of the standardized test statistic under the null hypothesis of no jumps. The continuous process X is generated using the same model as in Figure 1. We gradually increase the noise level α (from top to bottom) and the averaging window k_n (from left to right). The solid curve is the $\mathcal{N}(0,1)$ density.

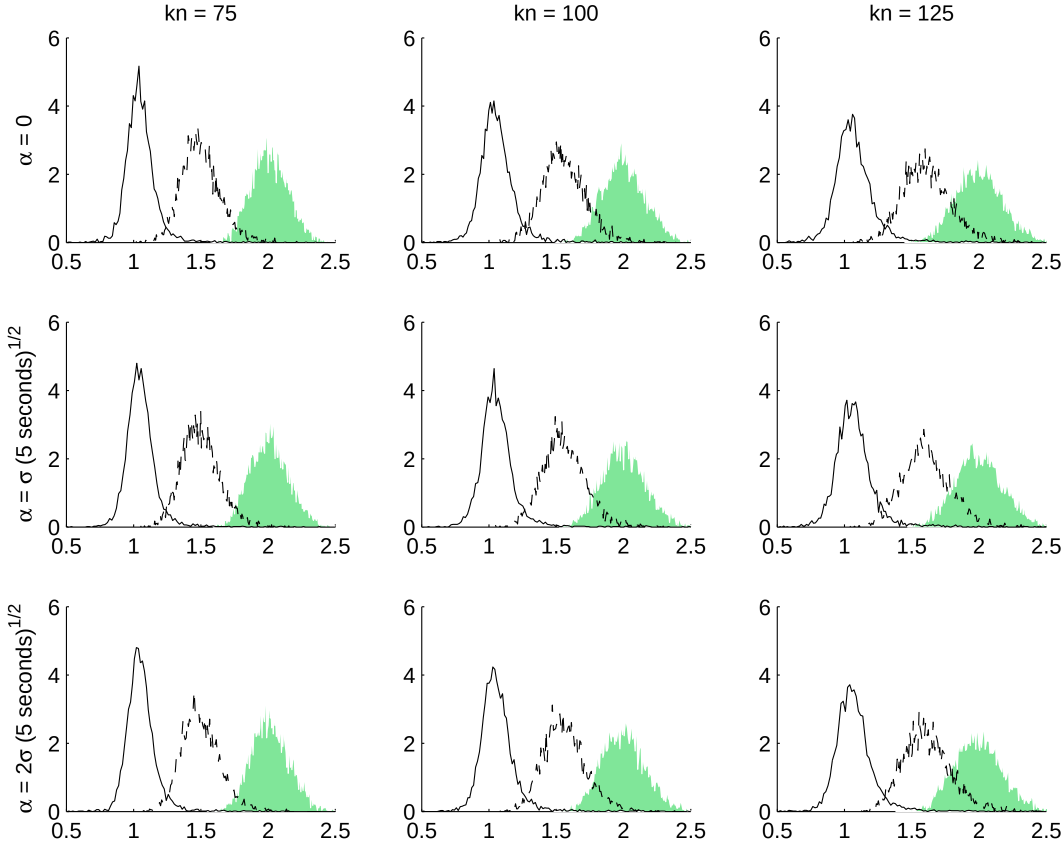

Fig. 3. Monte Carlo (5,000 simulations) distribution of the non-standardized test statistic $S(g, h, 4)_n$, using the same data generating process as in Figure 1. In particular, the jump process is a compound Poisson process with intensity $\lambda = 0$ (shaded area), 1 (solid curve) or 20 (dashed curve). As λ increases, we shrink the jump size to keep the expected quadratic variation of jumps constant. We gradually increase the noise level α (from top to bottom) and the averaging window k_n (from left to right).

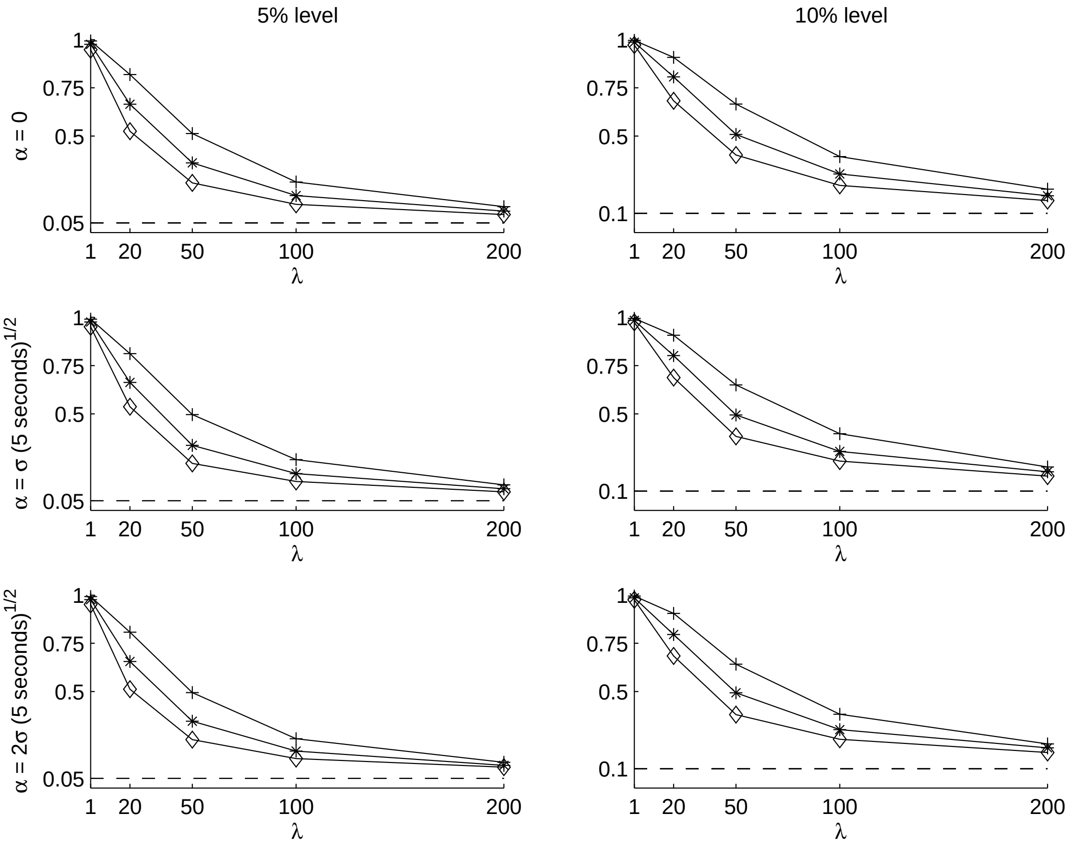

Fig. 4. Monte Carlo (5,000 simulations) rejection rates of 5%-level (left column) and 10%-level (right column) tests under the null hypothesis of no jumps when the underlying process X contains jumps, using the same data generating process as in Figure 1. In particular, the jump process is a compound Poisson process with intensity λ (horizontal axis). As λ increases, we shrink the size of jumps to keep the expected quadratic variation of jumps constant. We gradually increase the noise level α (from top to bottom) and the averaging window k_n: $k_n = 75$ $(+)$, $k_n = 100$ $(*)$ and $k_n = 125$ $(\diamond)$. The dashed line indicates the asymptotic level of the test.

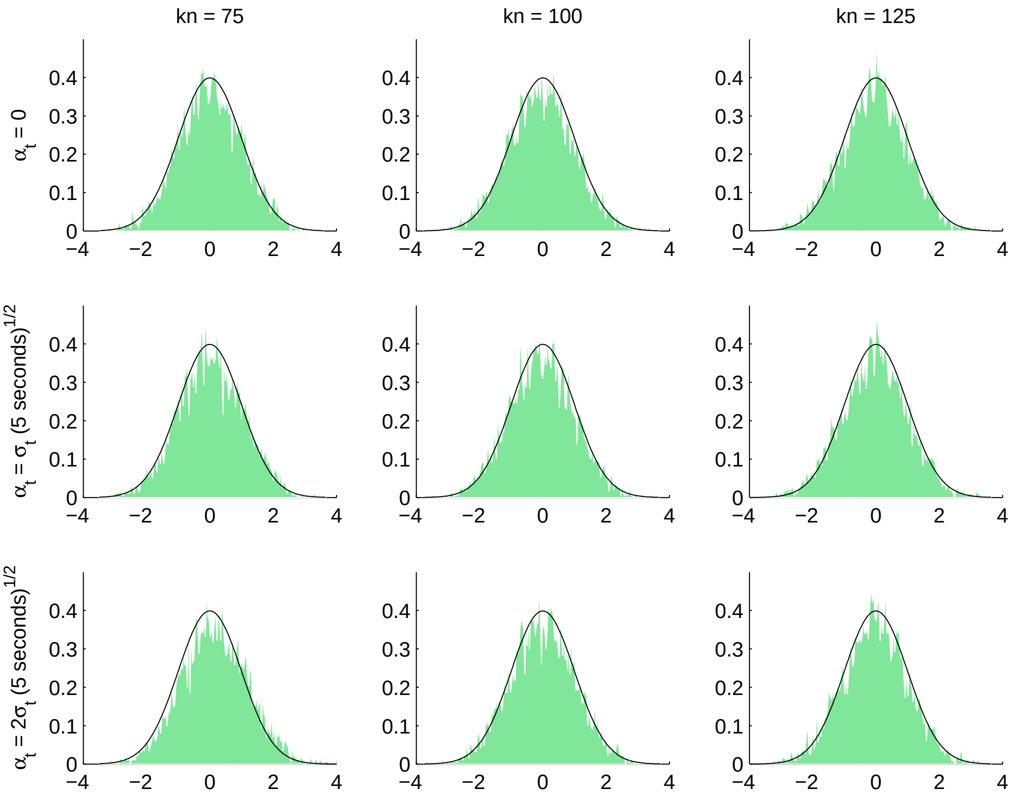

Fig. 5. Monte Carlo (5,000 simulations) and theoretical asymptotic distributions of the standardized test statistic under the null hypothesis of no jumps. The continuous process X is generated using the same model as in Table 2. We gradually increase the noise level α (from top to bottom) and the averaging window k_n (from left to right). The solid curve is the $\mathcal{N}(0,1)$ density.

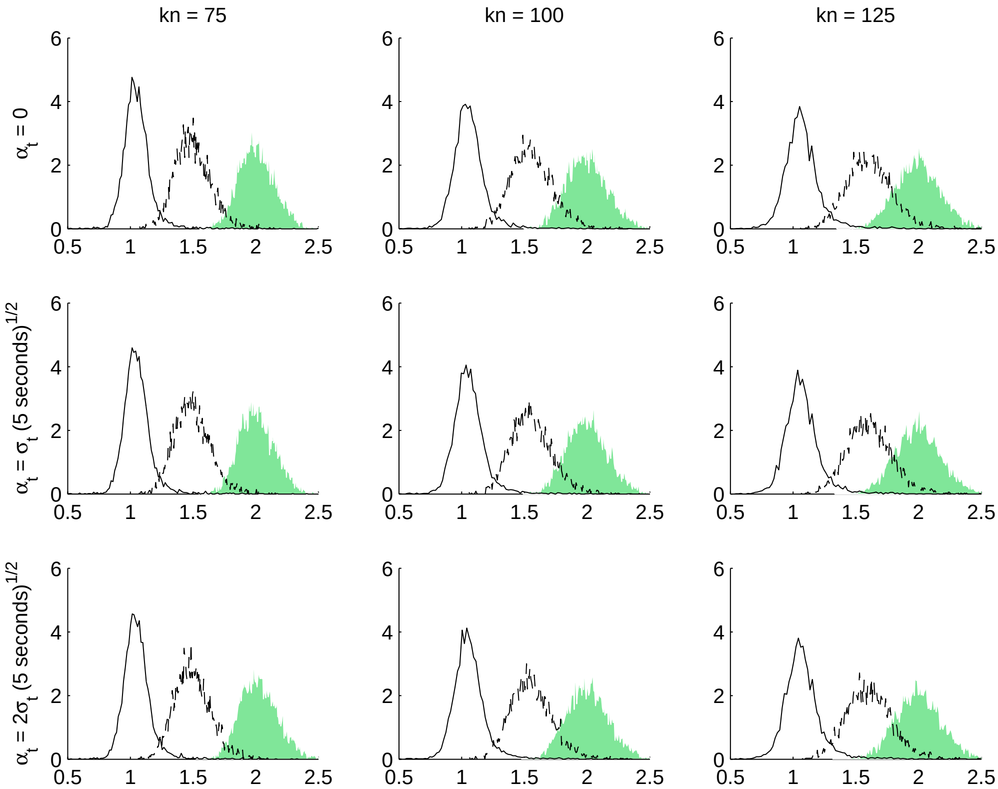

Fig. 6. Monte Carlo (5,000 simulations) distribution of the non-standardized test statistic $S(g,h,4)_n$. The underlying price process X is generated by $dX_t/X_t = \sigma_t dW_t + J_t dN_t$, where σ_t is the same as in Table 2, J_t is the product of a uniformly distributed variable on $[-2,-1] \cup [1,2]$ times a constant J_S and N is a Poisson process with intensity $\lambda = 0$ (shaded area), 1 (solid curve) or 20 (dashed curve) jumps per day. When $\lambda > 0$, J_S is determined by $\beta^2 = (7/3)J_S^2\lambda$ so that the continuous martingale part contributes 50% of total expected quadratic variation. 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 the sample, the number of jumps is slightly higher than specified. We compute the test statistic using $\log(Z_t)$, which is generated in the same way as in Table 2. We gradually increase the noise level α (from top to bottom) and the averaging window k_n (from left to right).

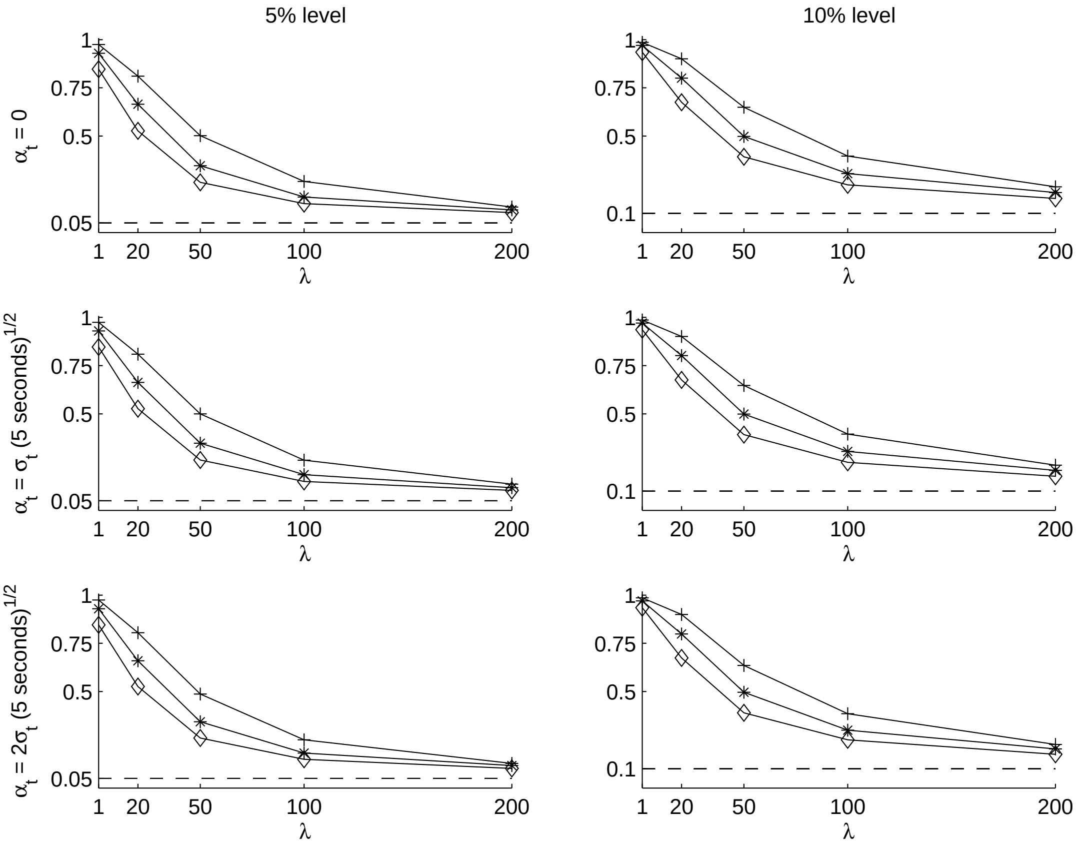

Fig. 7. Monte Carlo (5,000 simulations) rejection rates of 5%-level (left column) and 10%-level (right column) tests under the null hypothesis of no jumps when the underlying process X contains jumps, using the same data generating process as in Figure 6. In particular, the jump process is a compound Poisson process with intensity λ (horizontal axis). As λ increases, we shrink the size of jumps to keep the expected quadratic variation of jumps constant. We gradually increase the noise level α (from top to bottom) and the averaging window k_n: $k_n = 75$ (+), $k_n = 100$ ($*$) and $k_n = 125$ ($\diamond$). The dashed line indicates the asymptotic level of the test.

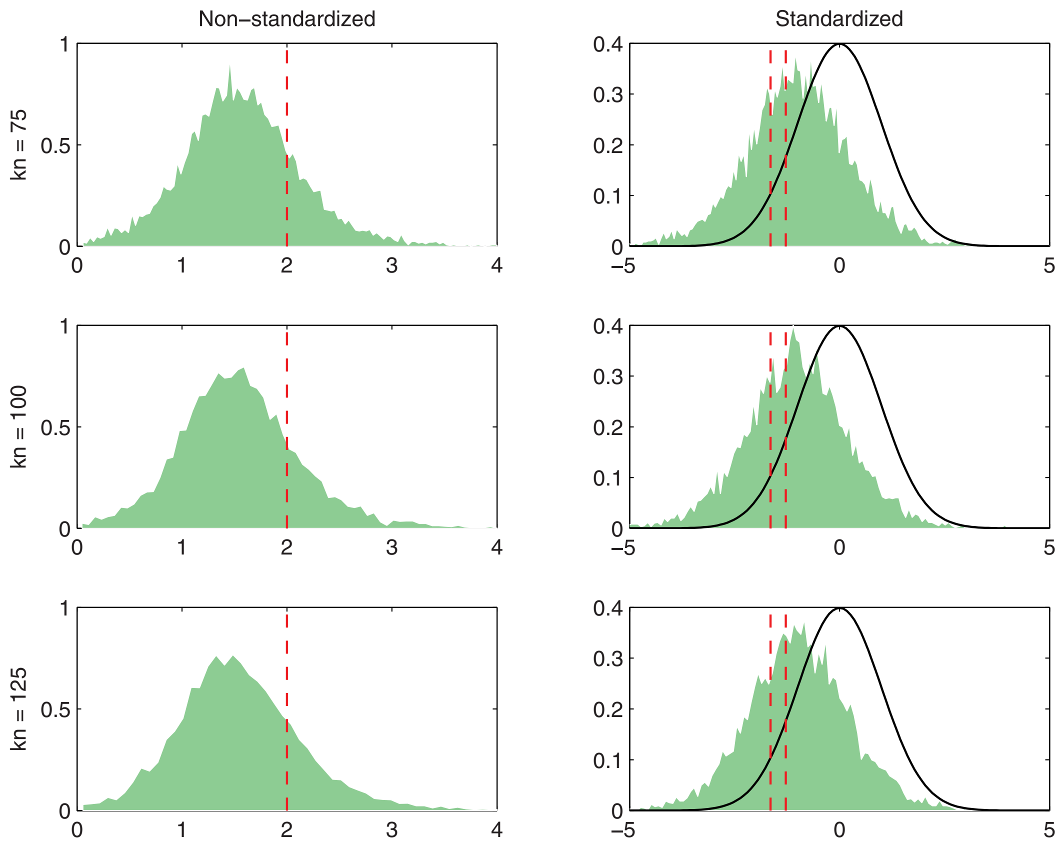

Fig. 8. Empirical distributions of the non-standardized statistic $S(g, h, 4)_n$ (left column) and the standardized statistic under the null hypothesis of no jumps (right column) for different averaging windows: $k_n = 75$ (top row), $k_n = 100$ (middle row) and $k_n = 125$ (bottom row). In the left column, the dashed line indicates the limit of the non-standardized statistic under the null hypothesis of no jumps. In the right column, the solid curve is the $\mathcal{N}(0, 1)$ density and the dashed lines indicate the 5% and the 10% quantiles of a standard normal variable. Each sample point is computed using all the transactions for one of the current (October 29th, 2009) 30 DJIA stocks observed over one trading day in 2008. This produces 7590 realizations of the statistics. We use the same window for computing the test statistic $S(g, h, 4)$ and the variance Σ_n. For each day, the truncation level u_n is determined by $u_n = C(\bar{g}(2)\overline{\sigma^2})^{1/2}\Delta_n^{0.47}k_n^{1/2}$, where $C = 7$ and the mean squared volatility $\overline{\sigma^2}$ is approximated by $k_n^{-1}\bar{g}(2)^{-1}\overline{V}(Z, g, 2; k_n)_T^n$.