SEMIMARTINGALE: ITÔ OR NOT ?

By Yacine Aït-Sahalia and Jean Jacod

Princeton University and UPMC (Université Paris-6)

[1] Itô semimartingales are the semimartingales whose characteristics are absolutely continuous with respect to Lebesgue measure. We study the importance of this assumption for statistical inference on a discretely sampled semimartingale in terms of the identifiability of its characteristics, their estimation, and propose tests of the Itô property against the non-Itô alternative when the observed semimartingale is continuous, or discontinuous with finite activity jumps, and under a number of technical assumptions.

1. Introduction. Semimartingales play an important role among stochastic processes, in particular in financial applications where they are the proper class of processes able to represent arbitrage-free asset prices, and in many other applied fields as well. It turns out that, among all semimartingales, the class of *Itô semimartingales* plays a very special role. An Itô semimartingale is a semimartingale which has absolutely continuous characteristics, or equivalently which is driven by a Brownian motion and a (usually compensated) Poisson random measure. The name comes from Itô (1951), in which Markov-type solutions of stochastic differential equations for these processes were introduced.

Although there is generally no theoretical reason for using Itô semimartingales rather than general semimartingales, the Itô framework is nearly universally adopted when semimartingales are employed in practice. They are much simpler and arise when one considers stochastic differential equations driven by some white noise random input such as a Brownian motion or a Poisson random measure. They are also (so far, at least) the only class of semimartingales amenable to statistical analysis based on discrete observations, or to approximation schemes such as the Euler-Maruyama scheme for SDEs if one desires a controlled rate of convergence.

This paper examines in what circumstances the Itô special structure among semimartingales impacts the identifiability of the characteristics of the process, whether it can be tested, and when it can, proposes tests of the Itô property against the non-Itô alternative. The latter aspect is related to Duvernet et al. (2010), who devised a statistical procedure to test whether the data generating process is an Itô

---

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

Key words and phrases. Semimartingale, Itô, discrete sampling, high frequency, absolute continuity, Cantor set.

semimartingale, against the specific alternative hypothesis that it consists of a multifractal random walk.

Let us mention right away that the identifiability analysis is conducted in the simple case of a continuous process with independent (non stationary) increments. The tests are for semimartingales that are continuous, or discontinuous but with finite activity jumps only; they need rather usual hypotheses (local boundedness and path regularity of the volatility), under the Itô hypothesis, and additionally (and unfortunately) a number of technical assumptions under the non-Itô alternative.

Consider a one-dimensional semimartingale X, with characteristics (B, C, ν) where B is the drift, C is the quadratic variation of the continuous martingale part, and ν is the compensator of the jump measure μ of X (see, e.g., Aït-Sahalia and Jacod (2014) for any unexplained, but classical, notation). By definition, the semimartingale is Itô if its characteristics factorize as follows, up to a null set:

$$(1) \quad B(\omega)_t = \int_0^t b(\omega)_s \, ds, \qquad C(\omega)_t = \int_0^t c(\omega)_s \, ds, \qquad \nu(\omega, dt, dx) = dt \, F_{t,\omega}(dx).$$

The processes b_t and c_t and the random measures F_t can be chosen to be predictable, although in many instances it is easier to assume that they are optional only (both versions actually exist). If X is not Itô, at least one of the three properties in (1) fails, but in any case C_t is continuous, whereas B_t and ν are predictable. The absolute continuity with respect to Lebesgue measure of the characteristics required in (1) for a semimartingale to be Itô may at first seem like a minor technical regularity condition, but it is not. For example, it leads to a natural definition of the "spot" characteristics (the terms b_s, c_s, F_s above), such as the spot volatility c_s as opposed to just the integrated volatility C_t. And we will see that when inference is performed on the basis of discrete observations of the process, typically regularly or nearly regularly spaced, whether the Itô property is satisfied or not does have important and somewhat surprising consequences for the asymptotic properties of the estimators.

Note that we observe only a single path of X. Furthermore, this path is only (partially) observed on the time interval $[0, 1]$, so whether (1) holds or not for $t > 1$ cannot be decided and is indeed irrelevant. Therefore, the Itô property means that (1) holds on the time interval $[0, 1]$ only. In order to set up a benchmark for what can and cannot be decided upon the observation of X at discrete times, one can first describe what happens in the idealized situation where the whole path $t \mapsto X_t$ for $t \in [0, 1]$ is observed. This amounts to analyzing the extent to which the characteristics can be identified for any given outcome ω on the basis of the knowledge of the complete path $t \mapsto X(\omega)_t$ (outside a null ω-set, of course) on a finite time interval. The answer to this question is summarized as follows:

1. The second characteristic C, which is the "continuous part" of the quadratic variation $[X, X]$, is always identifiable. Therefore deciding whether C is of the form (1) is, in principle, a solvable problem.

2. The behavior near the origin of the measures $\nu([0,t],dx)$, for all $t \in [0,1]$, can be retrieved to some extent: for example, the (possibly random and time-varying) Blumenthal-Getoor (BG) index of jump activity and the corresponding intensity process A_t are identifiable (see Chapter 11.2 in Aït-Sahalia and Jacod (2014)). In the Itô case, A should be of the form $A_t = \int_0^t a_s\,ds$, and deciding whether A is of this form is again a solvable problem.

3. The other parts of the characteristics (the drift B, or the measures $\nu([0,t],dx)$ outside an arbitrary neighborhood of 0) are typically *not identifiable*. So in particular if the jumps have finite activity, there is *no way* to test whether those jumps have the Itô property or not (meaning, ν satisfies (1) or not).

The previous statement about ν is always true in the sense that, except for very specific (and rather far-fetched) models nothing can be said about ν far from 0. On the other hand, in some not so strange cases we will see that the drift B turns out to be identifiable even on a finite time interval, when the Itô property no longer holds. A related example of the impact of the Itô property or lack thereof lies in the Euler-Maruyama discretization scheme: along a regular grid, rates of convergence differ based on the singularity of the characteristics of the process in the non-Itô case.

Even in the classical Itô semimartingale case, estimating the BG index and the associated intensity when observations are discrete is not a trivial task, which in addition requires somehow restrictive assumptions on the model (see Chapter 11.3 in Aït-Sahalia and Jacod (2014)). Checking whether A_t is absolutely continuous is even more difficult and so far we have no idea as how to tackle this question. Therefore, below we restrict our attention to the question of deciding whether the second characteristic C is absolutely continuous or not, and the consequences of the Itô property or lack thereof for the estimation of C.

The paper is organized as follows. We first describe in Section 2 the sampling scheme and the discrete variations to be used. Then we study in Section 3 the situation where the process has independent increments, covering in particular Lévy processes. The main results of the paper are contained in Sections 4 and 5, which are concerned respectively with continuous semimartingales, and with discontinuous semimartingales with jumps of finite activity in Section 5. Section 6 examines the performance of the test in Monte Carlo simulations. Section 7 contains concluding remarks. Proofs are in Section 9.

2. Discrete sampling. The precise assumptions on X will be specified below. As far as sampling is concerned, the process will be observed at the regularly spaced times $i\Delta_n$ for $i = 0, 1, \ldots$, within the time interval $[0,1]$, and the ith increment at stage n, for $i \geq 1$, is $\Delta_i^n X = X_{i\Delta_n} - X_{(i-1)\Delta_n}$, and $\Delta_i^n Y$ is defined in the same way for any other process Y. The asymptotic setting is high frequency, i.e., $\Delta_n \to 0$. However, for the problem at hand it is convenient to assume that 1 is an observation time, and also that going from stage n to stage $n+1$ amounts to adding new observations. That means that one assumes $\Delta_n = 1/a^n$ for some integer $a \geq 2$, which is not a serious restriction. Note that a^n consecutive returns are observed within $[0,1]$. Since X is only observed on the interval $[0,1]$, one might assume that the process is indexed by the time interval $[0,1]$, or alternatively that it takes a dummy value (such as $X_t = X_1$ for example) for all $t > 1$.

We will make use of some variational quantities related to the discretization procedure, and state a result describing the limiting behavior of these quantities. The following conventions are used throughout: $0^q = +\infty$ if $q < 0$, while $0^0 = 1$, $0/0 = 0$ and $x/0 = +\infty$ if $x > 0$. For any two processes Y, Z and any $p > 0$ we set

$$\begin{gathered} V(Y,p)_n = \sum_{i=1}^{a^n} |\Delta_i^n Y|^p \\ \overline{V}(Y,p)_n = \Delta_n^{1-p}\, V(Y;p)_n, \qquad \overline{V}(Y,Z,p)_n = \sum_{i=1}^{a^n} |\Delta_i^n Y|^p\, |\Delta_i^n Z|^{1-p}. \end{gathered} \tag{2}$$

Note that $\overline{V}(Y,p)_n = \overline{V}(Y,Z,p)$ with $Z_t = t$ and that $V(Y,p)_n$ and $\overline{V}(Y,p)_n$ are finite-valued, whereas $\overline{V}(Y,Z,p)_n$ may take the value $+\infty$. By Hölder's inequality, we necessarily have

$$1 \leq p < q \quad \Rightarrow \quad \overline{V}(Y,p)_n^{1/p} \leq \overline{V}(Y,q)_n^{1/q}. \tag{3}$$

When Y is a càdlàg process whose total variation $V(Y)$ over $[0,1]$ is finite, the sequence $V(Y,1)_n$ is non-decreasing with limit $V(Y)$, and is even identically equal to $V(Y)$ when Y is either non-increasing or non-decreasing. More generally with Y as before and Z increasing, the absolute continuity of Y with respect to Z is connected with the behavior of the sequence $\overline{V}(Y,Z,p)_n$ as follows (the next lemma will be used extensively in the rest of the paper and, although we are unable to trace a reference for the specific form stated below, its proof in Section 9 follows very classical lines):

Lemma 1. *Let Y, Z be two càdlàg functions with $Y_0 = Z_0 = 0$ and Z non-decreasing and $Z_1 > 0$. If $p > 1$, the sequence $\overline{V}(Y,Z,p)_n$ increases to a limit $\overline{V}(Y,Z,p)_\infty$. Moreover, this limit is finite if and only if $Y_t = \int_0^t y_s\, dZ_s$ for some function y satisfying $\int_0^1 |y_s|^p\, dZ_s < \infty$, in which case $\overline{V}(Y,Z,p)_\infty = \int_0^1 |y_s|^p\, dZ_s$.*

In all what follows, all processes are defined on a standard filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$.

3. The case of independent increments. We start with a simple scenario, where X is is a process with independent increments (PII, in short) and also is continuous. The reason for this is that a PII which is also a semimartingale has non-random characteristics, hence deciding whether (1) holds or not should in principle be simpler, since at least this property does not depend on the specific observed outcome ω.

By convention, a PII starts at $X_0=0$ and, when continuous, can be written as

$$
(4)\qquad \begin{array}{l} X=B+M, \quad \text{where } B_0=M_0=0,\ B \text{ is a continuous function,} \\ M \text{ a continuous Gaussian martingale} \\ \text{and } C_t=\mathbb{E}(M_t^2) \text{ is non-decreasing continuous.} \end{array}
$$

Conversely, (4) implies that X is a continuous PII, and it is also a semimartingale if and only if B is of (locally) finite variation, which we will assume from now on.

In the Lévy case, $B_t=bt$ and $C_t=ct$ for two constants $b\in\mathbb{R}$ and $c\geq 0$, and one knows how to efficiently estimate $c=C_1$ on the basis of our discrete observation scheme, and also that there is no consistent estimator for $b=B_1$ within this observation setting. In the Itô case, $B_t=\int_0^t b_s\,ds$ and $C_t=\int_0^t c_s\,ds$, with b_s and c_s possibly time-varying but non-random, and the same properties also apply at least when c_t is bounded away from 0.

As it turns out, things can be different in the non-Itô case, although the problem is still fully tractable when X is a continuous PII. This may help to cast some light on the differences between Itô and non Itô cases; that is why we divert for a while from our main testing problem, and consider estimation of the drift and volatility parameters in a parametric setting and, next, the estimation of the integrated volatility in a semi-parametric setting.

3.1. *Drift and volatility estimation in a fully parametric setting.* The parametric problem which we consider is as follows:

$$
(5)\qquad \begin{array}{l} B=b\,G, \quad C=c\,H, \quad \text{where} \\ G \text{ is a continuous function with finite variation and } G_0=0 \\ H \text{ is a continuous increasing function and } H_0=0,\ H_1>0. \end{array}
$$

Here, G and H are known, whereas $b\in\mathbb{R}$ and $c>0$ are the two unknown parameters to be estimated. The requirements $H_1>0$ rules out the trivial situation where X is a pure drift.

The following sequences play a fundamental role:

$$
(6)\qquad v_n=\sqrt{\overline{V}(G,H,2)_n},\quad N_n=\#I_n,\quad \text{where}\quad I_n=\{i:\,1\leq i\leq a^n,\ \Delta_i^n H>0\}.
$$

Both v_n and N_n are non-decreasing (by Lemma 1 for v_n).

Theorem 1. *We have three cases:*

(i) $v_n=\infty$ for all $n\geq n_0$: then at any stage $n\geq n_0$ there is an estimator $\widehat{b}_n$ satisfying $\widehat{b}_n=b$ almost surely; There are also estimators $\widehat{c}_n$ (for example the MLE) such that $\sqrt{N_n}\,(\widehat{c}_n-c)$ converges in law to $\mathcal{N}(0,2c^2)$, the centered Gaussian distribution with variance $2c^2$.

(ii) $v_n\uparrow\infty$ and $v_n<\infty$ for all n: there are estimators $(\widehat{b}_n,\widehat{c}_n)$ (for example the MLE) such that $(v_n(\widehat{b}_n-b),\sqrt{N_n}\,(\widehat{c}_n-c))$ converges in law to a centered Gaussian vector whose two components are independent and with respective variances c and $2c^2$.

(iii) $v_n \uparrow u < \infty$, *which is equivalent to having* $G_t = \int_0^t g_s\,dH_s$ *and* $\int_0^1 g_s^2\,dH_s < \infty$*: then the value* b *cannot be estimated in a consistent way, and there are estimators* $\widehat{c}_n$ *(for example the MLE computed with* $b = 0$*) such that* $\sqrt{N_n}\,(\widehat{c}_n - c)$ *converges in law to to* $\mathcal{N}(0, 2c^2)$ *.*

In other words, the drift can become identified in some cases ((i) and (ii) above) despite sampling taking place on a time interval of finite length. This happens for instance if C is absolutely continuous, but B is not, in which case it is impossible to have $G_t = \int_0^t g_s\,dH_s$.

3.2. *Integrated volatility estimation in a non-parametric setting.* We come back to the general situation (4) of a continuous PII, and additionally assume that B and C are unknown but subject to

$$B_t = \int_0^t r_s\,dC_s, \qquad \text{with} \quad \int_0^1 r_s^2\,dC_s < \infty \tag{7}$$

where r_s is again a non-random function. This is a natural restriction in financial applications: in these, the no-arbitrage condition implies the existence of an equivalent martingale measure, and by Girsanov's theorem this property (in the continuous PII case) is in fact equivalent to having (7). If further (5) holds, then (7) means that we are in case (iii) of Theorem 1, hence B cannot be estimated consistently.

Our aim is to estimate the "integrated volatility" C_1 at time 1. This is now a semi-parametric problem, and the so-called realized volatility

$$\widehat{C}_n = \sum_{i=1}^{a^n} (\Delta_i^n X)^2 \tag{8}$$

is always consistent for this (at least in the absence of microstructure noise, which is outside the scope of the present investigation). However, the rate of convergence of $\widehat{C}_n$ strongly depends upon whether we are in the Itô case or not. More precisely, we have:

Theorem 2. *Assume (4) and (7) and* $C_1 > 0$*, and set* $v_n = 1/\sqrt{2V(C,2)_n}$*.*

a) We always have $v_n \to \infty$*, and* $v_n \asymp \Delta_n^{-1/2}$ *if and only if* $C_t = \int_0^t c_s\,ds$ *with* $\int_0^1 c_s^2\,ds < \infty$*, whereas* $v_n\sqrt{\Delta_n} \to 0$ *otherwise.*

b) The sequence of variables $v_n(\widehat{C}_n - C_1)$ *is tight, and it also converges in law to* $\mathcal{N}(0,1)$ *as soon as we have for some real* $p > 2$*:*

$$\frac{V(C,p)_n}{V(C,2)_n^{p/2}} \;\to\; 0. \tag{9}$$

So, as far as volatility estimation is concerned, the genuine boundary is not between Itô and non Itô PIIs, but rather between PIIs having an absolutely continuous integrated volatility with a square-integrable density, and all others (Itô or not).

Of course, it might happen that the rate v_n is not sharp, meaning that the convergence holds with a strictly larger rate, but it is sharp under (9). This additional condition is satisfied (for all $p > 2$ at once) if

$$\frac{1}{V(C,2)_n} \sup_{i=1,\dots,a^n} (\Delta_i^n C)^2 \to 0 \tag{10}$$

(this is because, with $w_n = \sup_{i=1,\dots,a^n} \Delta_i^n C$, we have $V(C,p)_n \leq w_n^{p-2} V(C,2)_n = (w'_n)^{p/2-1} V(C,2)_n^{p/2}$, where w'_n is the left side of (10)). Under the additional assumption that the function C is Hölder with index ρ, (10) is satisfied when $\rho > 1/2$, because $\Delta_n^{-1} V(C,2)_n = \overline{V}(C,2)_n \geq \overline{V}(C,1)_n^2 = C_1^2$ is bounded away from 0.

Remark 1. *In the parametric setting (5) with, say, $G \equiv 0$ and a function H which is strictly increasing, we have $N_n = a^n = \Delta_n^{-1}$ and the MLE $\widehat{c}_n$ is asymptotically efficient and converges with rate $\Delta_n^{-1/2}$ and asymptotic variance $2c^2$. Translated into estimators for $C_1 = cH_1$, this gives the (efficient) estimators $\widehat{C}'^n = \widehat{c}_n H_1$ which satisfy*

$$\frac{1}{\sqrt{\Delta_n}} (\widehat{C}'^n - C_1) \stackrel{\mathcal{L}}{\to} \mathcal{N}(0, 2(C_1)^2),$$

irrespective of whether H is absolutely continuous or not. In contrast, the approximate quadratic variation estimators $\widehat{C}_n$ are not rate-efficient unless H is absolutely continuous with square-integrable density h_t, in which case they achieve the efficient asymptotic variance only if h_t is almost everywhere constant. This, however, does not contradict the asymptotic efficiency of $\widehat{C}_n$ when C has a square-integrable density, because in the non-parametric case the MLE $\widehat{c}_n$ is simply not available.

3.3. *Testing the Itô property for a PII.* We next turn to testing the Itô property (1). We still restrict our attention to continuous PIIs satisfying (4) and (7). Although this can be viewed as a special case of what will follow for more general semimartingales, it also helps us understand better in a simple setting the conditions which will be needed later on.

Since the Itô property is expressed in terms of the function C (because (7) holds), and more precisely in terms of the power variations of C, through Lemma 1, we have to get a handle on these power variations through the increments (or power variations) of X itself. For this, a key property is the following lemma, where m_p denotes the pth absolute moment of $\mathcal{N}(0,1)$:

Lemma 2. *Assume (4) and (7) and $C_1 > 0$. For any $p > 2$ we have the implication*

$$\frac{V(C,p)_n}{(V(C,p/2)_n)^2} \to 0 \implies \frac{V(X,p)_n}{V(C,p/2)_n} \stackrel{\mathbb{P}}{\to} m_p. \tag{11}$$

As already emphasized, the property we really want to prove or disprove is that C is absolutely continuous with a square-integrable density. When this fails we have $\overline{V}(C,2)_n \to \infty$ by Lemma 1, but if this convergence is too slow one can easily imagine that it is difficult to decide whether the sequence actually goes to ∞ or is bounded. This is why we introduce a condition expressing the fact that $\overline{V}(C,2)_n$ goes "fast enough" to ∞. In the sequel we in fact need a similar condition for other powers as well, so for all $p>2$ we state:

Property (A-p). *We have* $\limsup_n \overline{V}(C,p/2)_{n-1}/\overline{V}(C,p/2)_n < 1$.

Then, recalling that $V(C,p)_n = a^{n(1-p)}\overline{V}(C,p)_n$, under the condition of the previous lemma, including the left side of (11), plus (A-p), it is clear that for some number $\theta=\theta(p)>0$ we have

$$\mathbb{P}\left(\frac{V(X,p)_{n-1}}{V(X,p)_n} > a^{p/2-1}-\theta\right) \;\to\; 0. \tag{12}$$

On the other hand, when C has a p-integrable density, we have $\overline{V}(C,p)_{n-1}/\overline{V}(C,p)_n \to 1$ and thus $V(X,p)_{n-1}/V(X,p)_n \stackrel{\mathbb{P}}{\to} a^{p/2-1}$. So a natural test statistic, in the spirit of Aït-Sahalia and Jacod (2009) and Duvernet et al. (2010), is

$$S(p)_n \;=\; \frac{V(X,p)_{n-1}}{V(X,p)_n}, \tag{13}$$

and should in principle allow us to separate the null hypothesis "C has a p-integrable density" from the alternative "(A-p) holds".

This is indeed the case. Let us introduce the following two constants, where Ψ and Ψ' are two independent $\mathcal{N}(0,1)$ variables,

$$\begin{aligned}\gamma_p &= a^{p-2}((a+1)(m_{2p}-(m_p)^2)) - 2a^{p/2-1}\,\widetilde{m}_{2p}\\ \widetilde{m}_{2p} &= \mathbb{E}\left(|\Psi|^p\,|\Psi+\sqrt{a-1}\Psi'|^p\right) - a^{p/2}(m_p)^2,\end{aligned} \tag{14}$$

There is a Central Limit Theorem for the statistic $S(p)_n$ under the null hypothesis, and we need to estimate its (normalized) asymptotic variance. Toward this aim we introduce the "truncated" power variations

$$V(X,p;u_n)_n = \sum_{i=1}^{a^n} |\Delta_i^n X|^p\, 1_{\{|\Delta_i^n X|\le u_n\}}, \tag{15}$$

for a suitable sequence u_n in $(0,\infty]$ (so we recover $V(X,p)_n$ when $u_n=\infty$). Then, if z_α for $\alpha\in(0,1)$ is the number such that $\mathbb{P}(U>z_\alpha)=\alpha$ for an $\mathcal{N}(0,1)$ variable U, we have the following, whose Part (i) is a special case of Theorem 10.15 in Aït-Sahalia and Jacod (2014):

Theorem 3. *Assume (4) and (7), with $C_1>0$. Let $p>2$ and $u_n\in(0,\infty]$ be such that $u_n a^{n\varpi}\to\infty$ for some $\varpi\in(0,1/2)$, and define the critical regions*

$$\mathcal{C}(p)_n \;=\; \left\{S(p)_n < a^{p/2-1} - z_\alpha\sqrt{\frac{\gamma_p\,V(X,2p;u_n)_n}{m_{2p}\,V(X,p)_n^2}}\right\}\;. \tag{16}$$

(i) These critical regions have the asymptotic size α for the null hypothesis that $C_t = \int_0^t c_s\,ds$ with a piecewise Lipschitz function c_t.

(ii) They are asymptotically consistent for the alternative hypothesis that (A-p) holds together with $\frac{V(C,p)_n}{V(C,p/2)_n^2} \to 0$.

It is possible to take $u_n = \infty$ above but, as discussed later in Remark 4 in a more general situation, the power of the test is improved by taking u_n going to 0 and as small as possible, subject to the stated conditions.

We end this Section with an example of a function C satisfying all requirements of (ii) above:

Example 1. *We take for C the distribution function of the standard Cantor measure, say F_C, supported by the triadic Cantor set on $[0,1]$. Recall that $F_C(1) = 1$, and since any $t \in [0,1)$ can be uniquely written as $t = \sum_{j\geq 1} v_j/3^j$ with each v_j in $\{0,1,2\}$ and the "stationary" sequences with $v_j = 2$ for all j large being excluded, we have $F_C(t) = \sum_{1\leq j\leq k}(v_j \wedge 1)/2^j$ where $k = \inf(j : v_j = 1)$. (see, e.g., Dovgoshey et al. (2006)). F_C is also the limit of the following sequence: start with $F_0(t) = t$, then let*

$$F_{n+1}(t) = \begin{cases} F_n(3t)/2 & \text{if } 0 \leq t \leq 1/3 \\ 1/2 & \text{if } 1/3 \leq t \leq 2/3 \\ 1/2 + F_n(3t-2)/2 & \text{if } 2/3 \leq t \leq 1 \end{cases} \tag{17}$$

Then, using $a = 3$ to simplify the analysis, we see that $\Delta_i^n C$ equals 2^{-n} for 2^n values of i ranging from 1 to 3^n, and 0 otherwise. It follows that $V(C,p)_n = 2^{n(1-p)}$, hence $\frac{V(C,p)_n}{V(C,p/2)_n^2} = 2^{-n} \to 0$ and $\frac{\overline{V}(C,p/2)_{n-1}}{\overline{V}(C,p/2)_n} = \left(\frac{2}{3}\right)^{p/2-1} < 1$.

4. The case of a continuous semimartingale.

4.1. *The model assumptions.* In this Section we suppose that X is a continuous semimartingale. Its characteristics are $(B,C,0)$, with B continuous with finite variation and C continuous increasing, and $B_0 = C_0 = 0$. For the same reasons as in the PII case, we additionally assume (7), although now r_t is a random process.

The continuous increasing process C has a Lebesgue decomposition (pathwise, for each ω) where:

$$C = C' + C'', \quad C'_t = \int_0^t 1_A(s)\,dC_s = \int_0^t c'_s\,ds, \quad C''_t = \int_0^t 1_{A^c}(s)\,dC_s, \quad dC''_t \perp dt \tag{18}$$

for some random (optional) set A. Since (7) is assumed, we have the associated decompositions for B and X (with $b'_t = r_t c'_t$ below):

$$B = B' + B'', \qquad \begin{cases} B'_t = \int_0^t 1_A(s)\,dB_s = \int_0^t b'_s\,ds = \int_0^t r_s\,dC'_s \\ B''_t = \int_0^t 1_{A^c}(s)\,dB_s = \int_0^t r_s\,dC''_s \end{cases} \tag{19}$$

$$X = X_0 + X' + X'', \qquad X'_t = \int_0^t 1_A(s)\,dX_s, \qquad X''_t = \int_0^t 1_{A^c}(s)\,dX_s. \tag{20}$$

Here X' is an Itô semimartingale and X'' is not (unless of course $X'' \equiv 0$). Moreover, the stopping time

$$\tau \;=\; 1 \bigwedge \inf(t:\, C''_t > 0 \;\text{ or }\; B''_t \neq 0) \tag{21}$$

can be interpreted as the biggest time such that X restricted to $[0,\tau]$ is Itô, in the sense that the stopped process $X_{t\wedge S}$, with S a stopping time is an Itô semimartingale if and only if $S \leq \tau$ a.s.

The necessary technical assumptions on X are most easily stated in terms of the previous decomposition:

Assumption (H). *We have $X = X_0 + X' + X''$, where*

$$X'_t = \int_0^t r'_s c'_s\, ds + \int_0^t \sigma'_s\, dW'_s, \qquad X''_t = \int_0^t r''_s c''_s\, dH_s + \int_0^t \sigma''_s\, dM_s, \tag{22}$$

and

(i) W' is a Brownian motion and r' is a locally bounded optional process and $c' = \sigma'^2$ is a semimartingale satisfying (H-2), meaning it is an Itô semimartingale with characteristics which are "locally bounded" in a suitable sense (see Aït-Sahalia and Jacod (2014), p.480).

(ii) M is a continuous local martingale of the form $M_t = W''_{H_t}$ where W'' is a Brownian motion independent of W', and H is non-decreasing, continuous, null at 0, and with purely singular paths (that is, $dH_t \perp dt$) and is independent of both W' and W''.

(iii) σ'' is adapted, càdlàg, non-vanishing, and the left limit σ''_- does not vanish either, and $c''_t = \sigma''^2_t$ and r'' is a locally bounded optional process.

In (22) we could take $r'_t = r''_t$ without loss of generality.

Remark 2. *The assumptions (i) about the Itô semimartingale X' are equivalent to Assumption (K') in Aït-Sahalia and Jacod (2014), plus of course (7), and (iii) is quite weak. In contrast, (ii) is more restrictive, from a structural viewpoint, the reason being that we want the increments $\Delta_i^n X''$ to be "almost" Gaussian: such a property is true for a continuous Itô semimartingale, but not in general. The quadratic variation of the continuous martingale part X''^c of X'' is $C''_t = \int_0^t c''_s\, dH_s$, which is singular, in accordance with (19). Now, any continuous local martingale X''^c with quadratic variation C'' has the form $X''^c_t = W''_{C''_t}$ for a Brownian motion W'', so it looks like it satisfies (22) with $H = C''$ and $\sigma'' \equiv 1$. However, in general there might be dependencies between C'' and W'', and also between (C'', W'') and the Brownian motion W' driving X'. So (ii) implies the (needed) singularity of C'', and also puts restrictions on the above dependencies. These restrictions are mitigated by the fact that we allow σ'' to be random and arbitrarily dependent on everything else.*

4.2. *Testing the Itô property for a continuous semimartingale.* With the time horizon 1, our (theoretical) aim is to decide to which one of the following two complementary sets the observed outcome ω belongs:

$$\begin{aligned} \Omega_{\text{Itô}} &= \{\tau = 1\} = \{X''_t = 0 \ \forall\, t \leq 1\}, \\ \Omega_{\text{No-Itô}} &= \{\tau < 1\} = \{\exists\, t \leq 1: \ X''_t \neq 0\}. \end{aligned} \tag{23}$$

(the above equalities are almost sure only, but we do not care about null sets here).

The present setting is a direct extension of the continuous PII setting, so, exactly as in Theorem 3, testing the hypotheses (23) is impossible as such, and we will have to restrict those two sets. For example in this theorem we needed $C_1 > 0$, hence here we need to intersect the above sets with $\{C_1 > 0\}$. This is innocuous since, on the set $\{C_1 = 0\}$, we have $B_t = 0$ and thus $X_t = X_0$ for all $t \in [0,1]$, a case where one would never do any testing! We do need stronger restrictions on the alternative, and the hypotheses to be tested will be, with the real $p > 2$ being fixed (recall (2) and the convention $0/0 = 0$):

$$\begin{gathered} \Omega^1_{\text{Itô}} = \{H_1 = 0\} \cap \{C_1 > 0\}, \\ \Omega(p)^1_{\text{No-Itô}} = \{H_1 > 0\} \cap \Omega(p) \cap \Omega(2p) \cap \Omega(p)' \\ \text{where} \ \ \Omega(p) = \Big\{ \tfrac{V(C,p)_n}{V(C,p/2)^2_n} \to 0 \Big\}, \qquad \Omega(p)' = \Big\{ \limsup_n \tfrac{\overline{V}(C,p/2)_{n-1}}{\overline{V}(C,p/2)_n} < 1 \Big\}. \end{gathered} \tag{24}$$

In the definition of $\Omega(p)$ we could replace the process C by the process H (an easy thing to check), and the set $\Omega(p)'$ is also the set where $\limsup_n \frac{V(C,p/2)_{n-1}}{V(C,p/2)_n} < a^{p/2-1}$. We also have more simply $\Omega(p)^1_{\text{No-Itô}} = \Omega(p) \cap \Omega(2p) \cap \Omega(p)'$, since necessarily $H_1 > 0$ on $\Omega(p)'$, but we think that the definition (24) is clearer because it emphasizes the condition $H_1 > 0$ which expresses the non-Itô property.

When X satisfies (4) and (7) (hence (H) as well), the sets (24) are either empty or equal to Ω itself, and $\Omega^1_{\text{Itô}}$ (resp. $\Omega(p)^1_{\text{No-Itô}}$) equals Ω if and only if the null in (i) (resp. the alternative in (ii)) of Theorem 3 are in force (except for the Lipschitz property of c_t, though). So $\Omega^1_{\text{Itô}}$ and $\Omega(p)^1_{\text{No-Itô}}$ constitute, when X is a continuous semimartingale, the natural counterpart of the null and alternative hypotheses in Theorem 3.

At this stage, we can describe the test, with the null hypothesis being $\Omega^1_{\text{Itô}}$. We consider the pth power variations $V(X,p)_n$ of (2), and as in (13) the test statistic will be

$$S(p)_n \ = \ \frac{V(X,p)_{n-1}}{V(X,p)_n}. \tag{25}$$

We associate the same critical (rejection) region as in (14), that is

$$\mathcal{C}(p)_n \ = \ \left\{ S(p)_n < a^{p/2-1} - z_\alpha \sqrt{\frac{\gamma_p \, V(X,2p;u_n)_n}{m_{2p}\, V(X,p)^2_n}} \right\}. \tag{26}$$

Theorem 4. *Assuming (H) and $p > 2$ and letting $u_n \in (0, \infty]$ be such that $u_n a^{n\varpi} \to \infty$ for some $\varpi \in (0, 1/2)$, the tests with critical regions $\mathcal{C}(p)_n$ have the strong asymptotic size α for the null hypothesis $\Omega^1_{It\hat{o}}$, and are consistent for the alternative $\Omega(p)^1_{No\text{-}It\hat{o}}$.*

The two hypotheses here are random sets, so we recall that alternative-consistency means $\mathbb{P}((\mathcal{C}(p)_n)^c \cap \Omega(p)^1_{\text{No-Itô}}) \to 0$ as $n \to \infty$, and having the strong asymptotic size α means that, for any subset $A \subset \Omega^1_{\text{Itô}}$ with positive probability, we have $\mathbb{P}(\mathcal{C}(p)_n \mid A) \to \alpha$ as $n \to \infty$.

Remark 3. *In practice, we are given the smallest possible value Δ_n. To compute $S(p)_n$ we use this Δ_n, and another time lag (labeled Δ_{n-1} above) which is a multiple $a\Delta_n$ of Δ_n, where the integer a is also the one occurring in the definition of the critical region $\mathcal{C}(p)_n$. Typically one may take $a = 2$. In principle, doing this necessitates that the actual number of observed returns is a multiple of a, for example is even if we want to choose $a = 2$. This is not necessarily the case, but upon dropping the last return if needed one can use exactly the formulation (25).*

Remark 4. *The critical region $\mathcal{C}(p)_n$ depends on the tuning parameter u_n. This can be dispensed with by taking $u_n = \infty$, in contrast with the case where the statistic $S(p)_n$ is used for testing for jumps, as discussed in Section 10.3.3 of Aït-Sahalia and Jacod (2014). However, for the same reason as for jump testing, it is advisable to use a truly truncated version since truncating is helpful for increasing the power of the test. One can choose u_n "as small as possible", up to the condition $u_n a^{n\varpi} \to \infty$ for some $\varpi < 1/2$ (this is necessary for the consistency of the estimation of the asymptotic variance of $S(p)_n$ under the null). In practice, the choice of u_n should be of the order of magnitude of 3 to 5 times the average standard deviation of the individual returns.*

The previous test does not test the hypotheses (23), but rather those in (24): this is not an issue for the null hypothesis, but for the alternative $\Omega(p)^1_{\text{No-Itô}}$ is smaller than $\Omega_{\text{No-Itô}}$. The same issue already arose in the PII case, and there seems to be no way to circumvent it. A further issue is the following. The above test works if the continuity assumption of the underlying X is satisfied, but when X is an Itô semimartingale with jumps the probability of rejection goes to the probability that there is at least one jump, which is not a very desirable property because in this case X is Itô by assumption. Indeed, the statistic $S(p)_n$ is the one called $S^{J-PV2}(p, 2, \Delta_n)$ in Aït-Sahalia and Jacod (2014) (when $a = 2$), and the critical region $\mathcal{C}(p)_n$ is the one used to test the null that X is continuous against the alternative that X has jumps.

In other words, the above procedure is indeed a test for the null "X is a continuous Itô semimartingale", against the composite alternative "X is non Itô, or has jumps". This is why, in the next section we propose another procedure which provides a test for a null which includes the possibility that X has jumps.

5. Discontinuous semimartingales. As mentioned above, we want to relax the continuity assumption on X. We will assume that he jumps have finite activity. Apart from this, the jumps are totally arbitrary, for their dependence with the other ingredients occurring in X, and also for the laws of the jump times: in particular these can be predictable, or even non random, in contrast with the typical assumptions in the literature, according to which the jump times are related to a Poisson or modulated Poisson process.

More specifically, the assumption is as follows:

Assumption (H'). *The process X has the structure*

$$X = X_0 + X' + X'' + J, \qquad J_t = \sum_{n\geq 1} Z_n \, 1_{\{T_n \leq t\}}, \tag{27}$$

where $X' + X''$ satisfies Assumption (H), and (T_n) is a strictly increasing sequence of stopping times with $\lim_n T_n = \infty$ and the jump sizes Z_n are $\mathcal{F}_{T_n}$-measurable.

Our aim is still to test the null hypothesis $\Omega^1_{\text{Itô}}$ of (24): note that this null is concerned with the Itô property of the continuous part $X'+X''$ of X, and not about the Itô property of the jump part J which, as already mentioned, can simply not be tested at all. On the other hand, the alternative will be slightly more restrictive than $\Omega(p)^1_{\text{No-Itô}}$, and goes as follows:

$$\begin{gathered} \Omega(p)^2_{\text{No-Itô}} = \Omega(p)^1_{\text{No-Itô}} \cap \Omega(p)'' \cap \Omega(p/2)'', \quad \text{where} \\ \Omega(p)'' = \Big\{ \tfrac{n \, \sup_{1\leq i \leq a^n} \, |\Delta_i^n H|^{p-\varepsilon}}{V(H,p)_n} \to 0 \ \text{ for some } \varepsilon > 0 \Big\}. \end{gathered} \tag{28}$$

The idea for testing the null $\Omega^1_{\text{Itô}}$ under Assumption (H') is to use a simple modification of the test previously constructed in the continuous case, which allows us to get rid of the jumps, and this works because there are only finitely many of them. At stage n we observe the a^n returns $\Delta_i^n X$, which we re-order in an increasing way:

$$|\Delta^n_{i_1} X| \leq |\Delta^n_{i_2} X| \leq \cdots \leq |\Delta^n_{i_q} X| \leq \cdots \leq |\Delta^n_{i_{a^n}} X|. \tag{29}$$

We take a sequence k_n of integers and another sequence $u_n \in (0,\infty]$, and we set

$$\begin{gathered} V'(X,p)_n = \sum_{j=1}^{a^n - k_n} |\Delta^n_{i_j} X|^p, \quad V'(X,p;u_n)_n = \sum_{j=1}^{a^n-k_n} |\Delta^n_{i_j} X|^p \, 1_{\{|\Delta^n_{i_j} X| \leq u_n\}}, \\ S'(p)_n = \tfrac{V'(X,p)_{n-1}}{V'(X,p)_n}. \end{gathered} \tag{30}$$

With z_α and γ_p as in (26), we then define the critical region at stage n to be

$$\mathcal{C}'(p)_n = \left\{ S'(p)_n < a^{p/2-1} - z_\alpha \sqrt{\frac{\gamma_p \, V'(X,2p;u_n)_n}{m_{2p} \, V'(X,p)^2_n}} \right\}, \tag{31}$$

where $u_n \in (0,\infty]$ is as described below. We then have:

Theorem 5. *Under Assumption (H'), let $p > 2$ and k_n going to ∞ and smaller than Kn for some constant K and also $u_n \in (0,\infty]$ with $u_n a^{n\varpi} \to \infty$ for some $\varpi \in (0,1/2)$. Then, the tests with critical regions $\mathcal{C}'(p)_n$ have the strong asymptotic size α for the null hypothesis $\Omega^1_{Itô}$, and are consistent for the alternative $\Omega^2_{No\text{-}Itô}$.*

The time lag between successive observation being $\Delta_n = 1/a^n$, the upper bound condition on k_n might be written in the equivalent, but perhaps more concrete, form $k_n \leq K\log(1/\Delta_n)$.

Remark 5. *The construction of the variables $V'(X,p)_n$ in (30) is somewhat non standard, in the sense that in contrast with (2) for example the summands $|\Delta^n_{i_j}X|^p$ are "anticipating", since each one depends on all observed increments $\Delta^n_i X$ (through the sorting procedure (29)). This, however, does not impair the results, because k_n is "small enough" for $V'(X,p)_n$ to be essentially equal to $V(X,p)_n$ minus the sum of the pth powers of the finitely many jumps.*

Remark 6. *By our choice of k_n the previous method obviously* cannot *accommodate the case when the jumps of X have infinite activity. However, one could wonder whether truncated power variations $V(X,p;u_n)_n$, as defined in (15), could be used instead of $V'(X,p)_n$, so that the previous theorem holds when in (27) the process J has infinite activity. The answer is basically* no*, for the following reason.*

What happens when jumps have infinite activity? Suppose for instance that in (27) the process J is a symmetric stable process with index $\beta \in (0,2)$ and that $X_0 + X' + X''$ satisfy (H) with $r'_t = r''_t = 0$ and $c'_t = 1$ and F_C is the Cantor function (see Example 1), and also that either $c''_t = 0$ (so X is Itô), or $c''_t = 1$ (so X is not Itô). For simplicity, consider the case $a = 3$ to match the setting of that example, and $u_n = 3^{-\varpi n}$. In this scenario, the normalized variables $3^{\gamma n}V(X,p,u_n)_n$ can be shown to converge in probability to a non-vanishing constant, if (with the notation $\kappa = \log 2/\log 3$)

$$\gamma = \begin{cases} \big((p-\beta)\varpi\big) \bigwedge \frac{p-2}{2} & \text{if } c''_t \equiv 0,\ \varpi < \frac{1}{2}, \ \text{ or if } c''_t \equiv 1,\ \varpi > \frac{\kappa}{2} \\ \big((p-\beta)\varpi\big) \bigwedge \kappa\frac{p-2}{2} & \text{if } c''_t \equiv 1,\ \varpi < \frac{\kappa}{2}. \end{cases} \tag{32}$$

For a test based on these truncated power variation, we need the value of γ to be different in the two cases $c''_t \equiv 0$ and $c''_t \equiv 1$, that is, we need to take

$$\varpi \in \Big(\frac{\kappa}{2}\,\frac{p-2}{p-\beta}, \frac{\kappa}{2}\Big). \tag{33}$$

In other words, the irregular (non-Itô) component of the model and the infinite activity jump component both generate large increments that in finite samples can be close to each other in magnitude; the truncation needs to be finely tuned to eliminate the increments due to the jumps while keeping intact those due to the non-Itô component, and of course those due to the Brownian component as well.

Asymptotically, this separation is theoretically possible based on (33), but in order to implement a test based on truncated variations, one would need to know

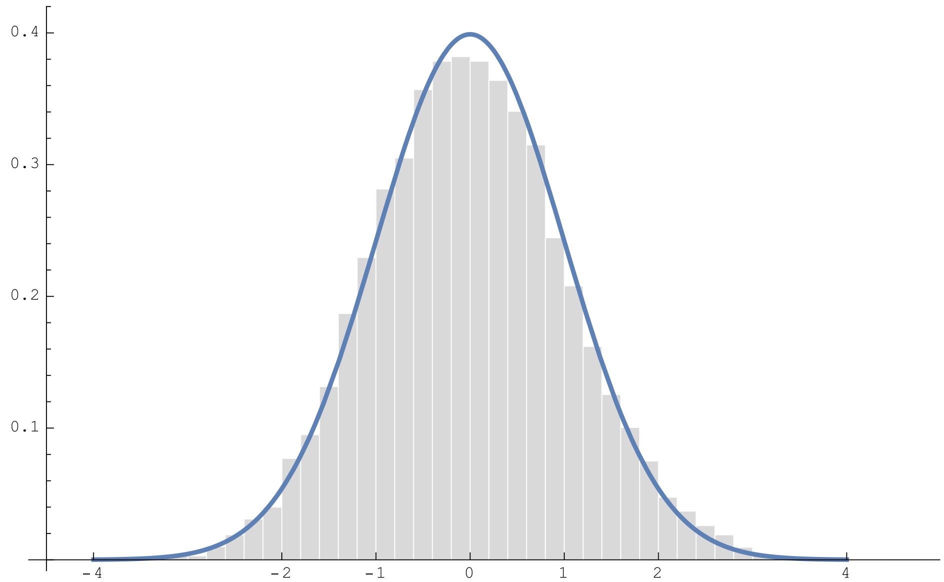

FIG. 1. *Small sample and asymptotic distribution of the test statistic under the null hypothesis*

(or at least have bounds on) the index β, and also on κ: for the Cantor function κ is given by $\log 2/\log 3$, but for another singular function H, κ should be replaced by the Hölder exponent of the function H, and this exponent is *a priori* arbitrary in $(0,1)$, and is of course unknown to the statistician who at this stage does not even know whether such a component is present in the data (which is the purpose of the test). Since the range of admissible values of ϖ given by (33) is narrow (especially when β is close to 2), it would be rather unrealistic to hope for a sound practical statistical procedure based on truncated power variations to solve the problem in this case. And since the proximity of the increments themselves is the root cause of the difficulty, a different statistic based on the same increments would likely face similar issues.

6. Simulations. In this section, we examine the performance of the test in Monte Carlo simulations. We consider as base models for X a straight Brownian motion and a Brownian motion time-changed by the Cantor function, F_C, given in Example 1. More specifically, we simulate from a Brownian motion time-changed from i/a^n to the following instants

$$A_{\frac{i}{a^n}} = \eta F_C\left(\frac{i}{a^n}\right) + (1-\eta)\,\frac{i}{a^n} \tag{34}$$

with $a = 3$, $\eta \in [0,1]$ a parameter controlling the mixing of the two base models, and F_C computed based on the recursion (17). The test statistic uses the increments

$$\Delta_i^n X = W_{A_{i/a^n}} - W_{A_{(i-1)/a^n}}. \tag{35}$$

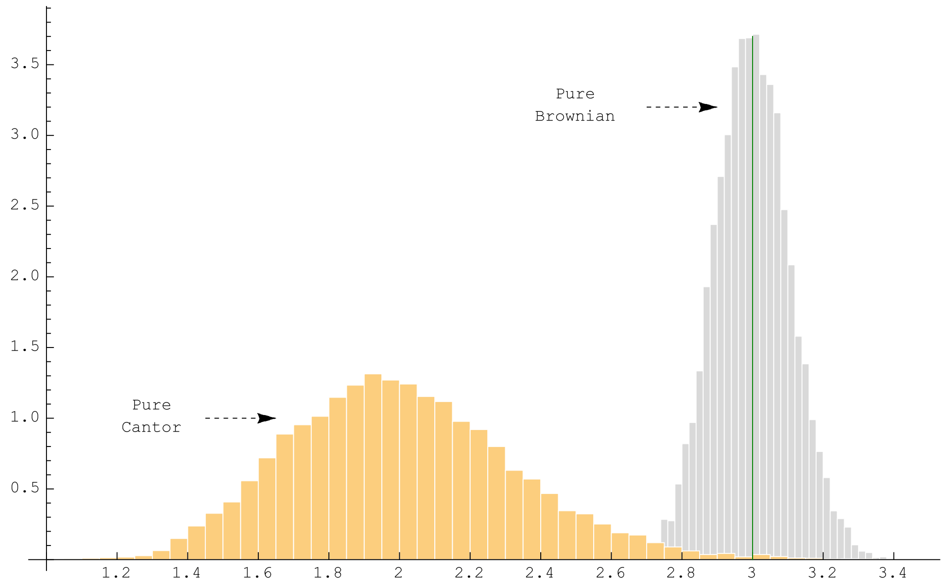

FIG. 2. *Small sample distributions of the test statistic under the null (pure Brownian) and alternative (pure Cantor) hypotheses*

As η varies, we examine the power of the test by mixing these two base models in varying quantities, thereby varying the irregularity of process. Models with $\eta > 0$ correspond to a non-Itô semimartingale, which in an intuitive sense get more and more irregular as η increases. The results below are obtained with $n = 9$, so the sample size is $3^9 = 19,683$ observations (corresponding in financial data to slightly less than one trading day of one second observations) and the test is implemented with $p = 4$. We run $M = 10,000$ simulated paths.

We start by reporting in Figure 1 a histogram of the small sample distribution of the standardized test statistic under the null hypothesis (the model with $\eta = 0$) and the theoretical $\mathcal{N}(0,1)$ asymptotic distribution as expressed in (26). Figure 2 reports the non-standardized values of the test statistic (25) for the two models $\eta = 0$ (pure Brownian) and $\eta = 1$ (pure Cantor). Under the null hypothesis, the theoretical probability limit of the test statistic is $a^{p/2-1} = 3$. Under the alternative hypothesis, recall from (12) that the test statistic takes values lower than 3. Finally, we validate in Figure 3 the result of Theorem 4, reporting the size of the test when $\alpha = 0.05$ and 0.10 (seen in the figure at $\eta = 0$) and its power as η increases from 0 to 1.

7. Concluding remarks. Although theoretical developments involving semimartingales do not typically require the Itô property, most semimartingales used in practice or in statistical applications turn out to satisfy this property. This paper studies what can be said with and without the Itô property regarding the identifi-

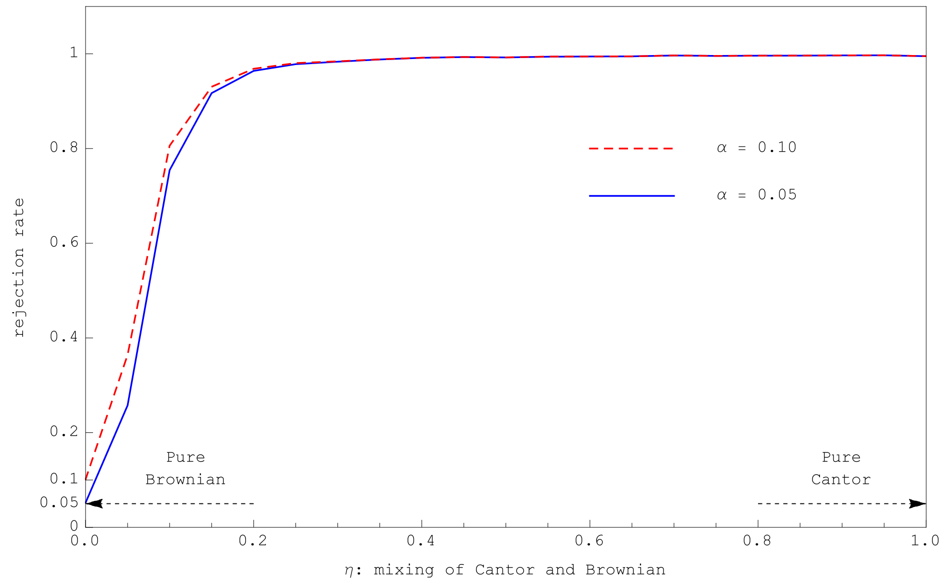

FIG. 3. *Size and power of the test*

ability and estimation of the characteristics of the process, and how to test the Itô property against the alternative of a non-Itô semimartingale. For example, we find that in some cases the drift of the process becomes identifiable, when it was not before, and the Itô property changes radically the rate of convergence of estimators of the integrated volatility of the process when the sampling scheme is regular.

8. Acknowledgements. We are grateful to the editor, associate editor and two referees for very constructive and helpful comments.

9. Proofs.

Proof of Lemma 1. We divide the situation into two cases: Case 1, where $\overline{V}(Y,Z,p)_n < \infty$ for all n and Case 2, where $\overline{V}(Y,Z,p)_n = \infty$ for at least one n.

We begin with Case 1. We equip the set $[0,1]$ with its Borel σ-field $\mathcal{G}$, the probability measure λ such that $\lambda([0,t]) = Z_t/Z_1$, and for each n the σ-field $\mathcal{G}_n$ generated by the intervals $I(i,n) = ((i-1)\Delta_n, i\Delta_n]$ for $i = 1, \cdots, a^n$. We set

$$M_n(t) = \sum_{i=1}^{a^n} \frac{\Delta_i^n Y}{\Delta_i^n Z}\, 1_{I(n,i)}(t), \qquad \text{hence } \ \mathbb{E}_\lambda(|M_n|^p) = \frac{1}{Z_1}\, \overline{V}(Y,Z,p)_n < \infty.$$

Since $\mathcal{G}_n \subset \mathcal{G}_{n+1}$, the sequence M_n is a martingale under λ, and relative to the filtration $(\mathcal{G}_n)$. Thus, by Jensen's inequality, $|M_n|^p$ is a submartingale (recall $p > 1$) and therefore $\overline{V}(Y,Z,p)_n$ is non-decreasing, with some limit denoted as $\overline{V}(Y,Z,p)_\infty$.

In Case 2, we let $N = \inf(n : \overline{V}(Y,Z,p)_n = \infty)$. First, we can define M_n and apply the same argument as previous, as long as $n < N$, so in particular the sequence $(\overline{V}(Y,Z,p)_n)_{1\leq n<N}$ is non-decreasing (and finite-valued). We also have $N < \infty$ and $\overline{V}(Y,Z,p)_N = \infty$ since we are in Case 2, and this implies that there is some integer j with $\Delta_j^N Y \neq 0$ and $\Delta_j^N Z = 0$ (otherwise each summand in $\overline{V}(Y,Z,p)_N$ would be finite). Looking at stage $n > N$, we deduce (since Z is non-decreasing) that $\Delta_i^n Z = 0$ for all i such that $I(n,i) \subset I(N,j)$, whereas $\Delta_i^n Y \neq 0$ for at least one such j. Therefore $\overline{V}(Y,Z,p)_n = \infty$, so indeed the whole sequence $(\overline{V}(Y,Z,p)_n)_{n\geq 1}$ is non-decreasing.

Now we turn to the sufficient condition in the second claim, thus assuming $Y_t = \int_0^t y_s\, dZ_s$ with $\int_0^1 |y_s|^p\, dZ_s < \infty$. Hölder's inequality yields $|\Delta_i^n Y|^p \leq (\Delta_i^n Z)^{p-1}\, \Delta_i^n H$, where $H_t = \int_0^t |y_s|^p\, dZ_s$. Then $\overline{V}(Y,Z,p)_n \leq H_1 < \infty$ and the sufficient condition holds.

Conversely, assume $\overline{V}(Y,Z,p)_\infty < \infty$. The martingale M_n is bounded in $L^p(\lambda)$, hence converges λ-a.s. and in $L^p(\lambda)$ to a limit M_∞ and $M_n = \mathbb{E}_\lambda(M_\infty \mid \mathcal{G}_n)$. Since M_∞ is a λ-integrable function on $[0,1]$ we can set $Y'_t = \int_0^t M_\infty(s)\, dZ_s$, and a simple computation shows that $M_n(t) = \Delta_i^n Y'/\Delta_i^n Z$ when $t \in I(n,i)$. Thus $\Delta_i^n Y' = \Delta_i^n Y$ for all i,n, clearly implying $Y' = Y$, which gives the necessary condition and also the final claim (by the above L^p-convergence). □

Proof of Theorem 1. We will repeatedly use the property $N_n \to \infty$, which comes from the fact that H is continuous and $H_1 > 0$.

(i) By hypothesis, for any $n \geq n_0$ there is an index i with $\Delta_i^n G \neq 0$ and $\Delta_i^n H = 0$, and all such i's are known to the statistician because G and H are known. Then if we take $\widehat{b}_n = \Delta_i^n X/\Delta_i^n G$ for such an i we have $\widehat{b}_n = b$ a.s. and the first part of the claim is proved.

For the second part, and assuming $n \geq n_0$ again, the observation of all $\Delta_i^n X$ amounts to the observation of all $\Delta_i^n Y$, where $Y = M$, because $B = bG$ is (almost surely) known. The variables $\Delta_i^n Y$ when i varies are independent and $\mathcal{N}(0, c\Delta_i^n H)$-distributed. Hence the log-likelihood for the Gaussian vector $(\Delta_i^n Y : i \in I_n)$ and the MLE for c are (recall $c > 0$):

$$\ell_n(c) = -\frac{1}{2} \sum_{i\in I_n} \left(\frac{1}{c}\, \frac{|\Delta_i^n Y|^2}{\Delta_i^n H} + \log(2\pi\, c\Delta_i^n H) \right), \qquad \widehat{c}_n = \frac{1}{N_n} \sum_{i\in I_n} \frac{|\Delta_i^n Y|^2}{\Delta_i^n H}.$$

Clearly, $\widehat{c}_n$ has the same law as $\frac{1}{N_n}\sum_{i\in I_n} (\zeta_i)^2$, where the ζ_i's are i.i.d. $\mathcal{N}(0,c)$, and the result readily follows.

(ii) Now we have $\Delta_i^n G = 0$ and $\Delta_i^n X = 0$ a.s., whenever $\Delta_i^n H = 0$. The variables $\Delta_i^n X$ when i varies are independent and $\mathcal{N}(b\Delta_i^n G, c\Delta_i^n H)$-distributed. Hence the log-likelihood for the Gaussian vector $(\Delta_i^n X : i \in I_n)$ is

$$\ell_n(b,c) = -\frac{1}{2} \sum_{i\in I_n} \left(\frac{1}{c}\, \frac{|\Delta_i^n X - b\Delta_i^n G|^2}{\Delta_i^n H} - \log(2\pi\, c\Delta_i^n H) \right)$$

and the MLE for b, c are (with $0/0 = 0$):

$$\widehat{b}_n = \frac{\sum_{i\in I_n} \Delta_i^n X \Delta_i^n G/\Delta_i^n H}{\sum_{i\in I_n} (\Delta_i^n G)^2/\Delta_i^n H} \qquad \text{and} \qquad \widehat{c}_n = \frac{1}{N_n}\sum_{i\in I_n} \frac{(\Delta_i^n X - \widehat{b}_n\, \Delta_i^n G)^2}{\Delta_i^n H}.$$

With ζ_i as in (i) above, the pair $(\widehat{b}_n - b, \widehat{c}_n - c)$ has the same law as $(\overline{b}_n, \overline{c}_n)$, where

$$\overline{b}_n = \frac{1}{v_n^2}\sum_{i\in I_n} \frac{\Delta_i^n G}{\sqrt{\Delta_i^n H}}\,\zeta_i^n, \qquad \overline{c}_n = \alpha_n - \alpha'_n,$$

$$\alpha_n = \frac{1}{N_n}\sum_{i\in I_n} \left((\zeta_i^n)^2 - c\right), \qquad \alpha'_n = \frac{(\overline{b}_n)^2\, v_n^2}{N_n}.$$

The variable $\overline{b}_n$ is $\mathcal{N}(0, c/v_n^2)$ distributed, and the ordinary two-dimensional CLT gives us that $(v_n\overline{b}_n, \sqrt{N_n}\,\alpha_n)$ converges to a centered Gaussian vector having uncorrelated components with respective variances c and $2c^2$. On the other hand, $\mathbb{E}(|\alpha'_n|) = c/N_n$, so the result follows.

(iii) The first claim follows from Girsanov's Theorem, which under (7) asserts that the laws of X (as a process) when b varies are all equivalent. With I_n denoting the set of all $i = 1, \ldots, a^n$ such that $\Delta_i^n H > 0$, the variable

$$\widehat{c}_n = \frac{1}{N_n}\sum_{i\in I_n} \frac{(\Delta_i^n X)^2}{\Delta_i^n H}$$

is the MLE for c when $b = 0$. When b is arbitrary, and since the variables $\zeta_i^n = \Delta_i^n M/\sqrt{c\Delta_i^n H}$ for $i \in I_n$ are i.i.d. $\mathcal{N}(0,1)$, we have $\widehat{c}_n - c = U_n + U'_n + U''_n$, where

$$U_n = \frac{c}{N_n}\sum_{i\in I_n}((\zeta_i^n)^2 - 1), \qquad U'_n = \frac{2b\sqrt{c}}{N_n}\sum_{i\in I_n} \frac{\Delta_i^n G}{\sqrt{\Delta_i^n H}}\,\zeta_i^n, \qquad U''_n = \frac{b^2}{N_n}\sum_{i\in I_n} \frac{(\Delta_i^n G)^2}{\Delta_i^n H}.$$

The classical CLT gives us $\sqrt{N_n}\,U_n \xrightarrow{\mathcal{L}} \mathcal{N}(0, 2c^2)$. Next, U'_n is a centered Gaussian variable with variance $\frac{4b^2c}{N_n^2}\,\overline{V}(G,H,2)_n$, and also $U''_n = \frac{b^2}{N_n}\,\overline{V}(G,H,2)_n$. Since the sequence $\overline{V}(G,H,2)_n$ is bounded, because of (5) and Lemma 1, we deduce $\sqrt{N_n}\,U'_n \xrightarrow{\mathcal{L}} 0$ and $\sqrt{N_n}\,U''_n \to 0$ and the second claim follows. □

Proof of Theorem 2. The proof is somewhat similar to the previous one. Namely, the variable $\widehat{C}_n - C_1$ has the same law as $U_n + U'_n + U''_n$, where

$$U_n = \sum_{i=1}^{a^n} \Delta_i^n C(\zeta_i^2 - 1), \qquad U'_n = 2\sum_{i=1}^{a^n} \Delta_i^n B\,\sqrt{\Delta_i^n C}\,\zeta_i, \qquad U''_n = V(B,2)_n,$$

and the ζ_i's are i.i.d. $\mathcal{N}(0,1)$. Note also that $V(C,2)_n \to 0$ because C is increasing continuous, hence $v_n \to \infty$.

First, $Z_t = \int_0^t r_s^2\,dC_s$ is a continuous non-decreasing function with $Z_1 < \infty$, hence it satisfies $V(Z,2)_n \to 0$ and $Z_n^* := \sup_{1\leq i\leq a^n} \Delta_i^n Z \to 0$. The Cauchy-Schwarz inequality yields $(\Delta_i^n B)^2 \leq \Delta_i^n C \Delta_i^n Z$ and thus $V(B,2)_n \leq \sqrt{V(C,2)_n V(Z,2)_n}$, hence $v_n U_n'' \leq \sqrt{V(Z,2)_n/2} \to 0$. Observe also that $v_n U_n'$ is a centered Gaussian variable with variance $4v_n^2 \sum_{i=1}^{a^n} (\Delta_i^n B)^2 \Delta_i^n C$, which by the same argument is not bigger than $2Z_n^*$, hence $v_n U_n' \stackrel{\mathbb{P}}{\to} 0$.

Second, if $\xi_i^n = v_n \Delta_i^n C(\zeta_i^2 - 1)$, we have for any $q > 2$ and some constant K_q:

$$\mathbb{E}(\xi_i^n) = 0, \quad \mathbb{E}((\xi_i^n)^2) = 2(v_n \Delta_i^n C)^2, \quad \mathbb{E}(|\xi_i^n|^q) \leq K_q (v_n \Delta_i^n C)^q,$$

and in particular $\mathbb{E}((v_n U_n)^2) = \sum_{i=1}^{a^n} \mathbb{E}((\xi_i^n)^2) = 1$. The tightness of the sequence $v_n U_n$ follows. Moreover, by a classical convergence criterion for rowwise independent triangular arrays, the convergence $v_n U_n \to \mathcal{N}(0,1)$ holds as soon as $v_n^p V(C,p)_n \to 0$ for some $p > 2$, and this is (9). □

(4) plus (7) imply Assumption (H), so Lemma 2 and Theorem 3 are special cases of Lemma 3 below and Theorem 4. Hence we now turn to "general" semimartingales. By a standard localization procedure one can replace Assumption (H) by the following strengthened assumption:

Assumption (SH)**.** *We have Assumption (H), and furthermore r_t', σ_t', r_t'', σ_t'', $1/\sigma_t''$ and H_1 are uniformly bounded. Moreover, c_t' is a bounded semimartingale with bounded spot characteristics.*

In all the sequel, K is a constant which changes from line to line and may depend on the powers p or q involved. Set $I(n,i) = ((i-1)\Delta_n, i\Delta_n]$ and

$$\eta_i^n = \sup_{t\in I(n,i)} \big(|c_t' - c_{(i-1)\Delta_n}'| + |\sigma_t' - \sigma_{(i-1)\Delta_n}'| + |c_t'' - c_{(i-1)\Delta_n}''| + |\sigma_t'' - \sigma_{(i-1)\Delta_n}''|\big).$$

By Lemma 1, for all $q \geq 1$ we have the pointwise convergence (i.e., for each ω).

$$\Delta_n^{1-q}\, V(C',q)_n \to \int_0^1 {c_s'}^q\, ds, \tag{36}$$

and thus $\frac{V(C',p)_n}{V(C',p/2)_n^2} \to 0$. On the other hand, X', as any continuous Itô semimartingale with locally bounded spot characteristics, satisfies

$$\Delta_n^{1-q/2}\, V(X',q)_n \stackrel{\mathbb{P}}{\longrightarrow} m_q \int_0^1 {c_s'}^{q/2}\, ds,$$

which, joint with (36), yields (11) in restriction to the set $\{C_1' > 0\}$ for the pair (X', C'). The next lemma tells us that (11) for (X,C) also holds, although X is not Itô. However, this strongly relies upon the structural assumption (ii) in (H), and whether it holds for *all* continuous semimartingales (with locally bounded spot characteristics) is an open question.

Lemma 3. *If $p \geq 2$ we have*

$$\frac{V(X,p)_n}{V(C,p/2)_n} \stackrel{\mathbb{P}}{\longrightarrow} m_p \qquad \text{on the set} \quad \{C_1 > 0\} \cap \Omega(p). \tag{37}$$

Proof. a) We can write

$$\begin{gathered} \Delta_i^n X = \xi_i^n + \widehat{\xi}(1)_i^n + \widehat{\xi}(2)_i^n + \widehat{\xi}(3)_i^n \\ \xi_i^n = \sigma'_{(i-1)\Delta_n} \Delta_i^n W' + \sigma''_{(i-1)\Delta_n} \Delta_i^n M, \qquad \widehat{\xi}(1)_i^n = \Delta_i^n B \\ \widehat{\xi}(2)_i^n = \int_{I(n,i)} (\sigma'_t - \sigma'_{(i-1)\Delta_n})\, dW'_t, \quad \widehat{\xi}(3)_i^n = \int_{I(n,i)} (\sigma''_t - \sigma''_{(i-1)\Delta_n})\, dM_t. \end{gathered}$$

Since $\big||x+y|^p - |x|^p\big| \leq K(|y|^p + |x|^{p-1}|y|)$ for all $x, y \in \mathbb{R}$, it is enough to show that

$$R_n \stackrel{\mathbb{P}}{\to} m_p, \qquad R(j)'_n \stackrel{\mathbb{P}}{\to} 0, \qquad R(j)''_n \stackrel{\mathbb{P}}{\to} 0, \qquad \text{on } \{C_1 > 0\} \cap \Omega(p). \tag{38}$$

for $j = 1, 2, 3$, where

$$R_n := \frac{\sum_{i=1}^{a^n} |\xi_i^n|^p}{V(C,p/2)_n}, \qquad R(j)'_n := \frac{\sum_{i=1}^{a^n} |\xi_i^n|^{p-1} |\widehat{\xi}(j)_i^n|}{V(C,p/2)_n}, \qquad R(j)''_n := \frac{\sum_{i=1}^{a^n} |\widehat{\xi}(j)_i^n|^p}{V(C,p/2)_n}.$$

b) Define a continuous non-decreasing process $\overline{C}_t = H_t + t\, 1_{\{C'_1 > 0\}}$. We will show that

$$V(C,q)_n \leq K_q V(\overline{C},q)_n, \qquad n \geq n_0 \implies V(\overline{C},q)_n \leq A_q V(C,q)_n \tag{39}$$

for any $q > 0$, a constant K_q, a (finite) random variable A_q, and an integer-valued random variable n_0. The first inequality above follows from the boundedness of c'_t and c''_t. The second one reduces to $V(H,q)_n \leq K_q V(C,q)_n$, which holds because $1/c''_t$ is bounded, plus $\Delta_n^{q-1} \leq A_q V(C',q)_n$ if $C'_1 > 0$ and $n \geq n_0$. For the latter observe that $C'_1(\omega) > 0$ implies $c'(\omega)_t \geq \gamma(\omega) > 0$ for all t in a subinterval of $[0,1]$ with length $L(\omega) > 0$, hence $V(C',q)_n(\omega) \geq \gamma(\omega)^q \Delta_n^{q-1} L(\omega)/2$ as soon as $L(\omega) \geq 4\Delta_n$, hence the claim with $A_q = 2/\gamma(\omega)^q L(\omega)$ and $n_0(\omega) = \inf(n : L(\omega) \geq 4\Delta_n)$. A consequence of (39) is

$$\{C_1 > 0\} \cap \Omega(p) = \{\overline{C}_1 > 0\} \cap \{\frac{V(\overline{C},p)_n}{V(\overline{C},p/2)_n^2} \to 0\} \tag{40}$$

c) Here we prove $R_n \stackrel{\mathbb{P}}{\to} m_p$. Let $\mathcal{H}_t$ be the σ-field generated by $\mathcal{F}_t$ and the whole process H. Assumption (H) implies that the variable ξ_i^n is, conditionally on $\mathcal{H}_{(i-1)\Delta_n}$, centered Gaussian with variance $v_i^n = c'_{(i-1)\Delta_n} \Delta_n + c''_{(i-1)\Delta_n} \Delta_i^n H$, hence $|\xi_i^n|^p - m_p (v_i^n)^{p/2}$ is a martingale increment for $(\mathcal{H}_{i\Delta_n})_{i \geq 0}$, with conditional variance $(m_{2p} - m_p^2)(v_i^n)^p$. Since further $V(C,p/2)_n > 0$ when $C_1 > 0$, the first part of (38) will follow from the next two (pointwise) convergences:

$$x_n := \frac{\sum_{i=1}^{a^n} (v_i^n)^{p/2}}{V(C,p/2)_n} \to 1 \quad \text{and} \quad y_n := \frac{\sum_{i=1}^{a^n} (v_i^n)^p}{V(C,p/2)_n^2} \to 0, \qquad \text{on } \{C_1 > 0\} \cap \Omega(p).$$

We have $|v_i^n - \Delta_i^n C| \leq K\eta_i^n \Delta_i^n \overline{C}$ and also v_i^n, $\Delta_i^n C \leq K\Delta_i^n \overline{C}$ and $(\eta_i^n)^q \leq K\eta_i^n$, hence $|(v_i^n)^q - (\Delta_i^n C)^q| \leq K\eta_i^n (\Delta_i^n \overline{C})^q$ and $(v_i^n)^q \leq K(\Delta_i^n \overline{C})^q$ for any $q \geq 1$. Then, if

$$z_n = \frac{\sum_{i=1}^{a^n} \eta_i^n (\Delta_i^n \overline{C})^{p/2}}{V(\overline{C}, p/2)_n},$$

upon using (39) we obtain on the set $\{C_1 > 0\} \cap \{n \geq n_0\}$:

$$|x_n - 1| \leq K \frac{\sum_{i=1}^{a^n} \eta_i^n (\Delta_i^n \overline{C})^q}{V(C, p/2)_n} \leq K A_{p/2}\, z_n, \qquad y_n \leq K \frac{V(\overline{C}, p)_n}{V(\overline{C}, p/2)_n^2}.$$

(40) yields $y_n \to 0$ on the set $\{C_1 > 0\} \cap \Omega(p)$. Next, fix $\varepsilon > 0$ and denote by $S_1, S_2, \cdots$ the successive jump times of the two-dimensional process (c', c''), with size bigger than ε. There are $N = N(\varepsilon, \omega)$ such jumps in $[0,1]$ and each S_j belongs to some interval $I(n, i(n,j))$, and for n large enough (depending on ω) we have $\eta_i^n \leq 2\varepsilon$ for all $i = 1, \cdots, a^n$ different from $i(n,j)$ for $j = 1, \cdots, N$. In this case, since further $\big(\sum_{j=1}^N (\Delta_{i(n,j)}^n \overline{C})^{p/2}\big)^2 \leq NV(\overline{C}, p)_n$, we deduce

$$z_n \leq 2\varepsilon + \Big(N \frac{V(\overline{C}, p)_n}{V(\overline{C}, p/2)_n^2}\Big)^{1/2},$$

which in view of (40) gives us $z_n \to 0$ (let first n go to ∞ and then ε to 0), hence $x_n \to 1$, on the set $\{C_1 > 0\} \cap \Omega(p)$: we conclude the first part of (38).

d) Hölder's inequality yields $R(j)'_n \leq (R_n)^{\frac{p-1}{p}} (R(j)''_n)^{\frac{1}{p}}$, so the second part of (38) readily follows from the first and third parts, and we are left to prove the last part.

In the case $j = 1$ we fix ω. Since $|\Delta_i^n B| \leq K\, \Delta_i^n \overline{C}$ we deduce from (39) that $R(1)''_n \leq K A_{p/2} \frac{V(\overline{C}, p)_n}{V(\overline{C}, p/2)_n}$, which goes to 0 on $\{C_1 > 0\} \cap \Omega(p)$ by (40).

In the case $j = 2$ and since $\widehat{\xi}(2)_i^n = 0$ for all i if $C'_1 = 0$, by (36) it is enough to show that $U_n = \Delta_n^{1-p/2} \sum_{i=1}^{a^n} |\widehat{\xi}(2)_i^n|^p \stackrel{\mathbb{P}}{\to} 0$. Using the Burkholder-Gundy inequality and $|\eta_i^n| \leq K$, we see that $\mathbb{E}(U_n) \leq K\mathbb{E}(Z_n)$, where $Z_n = \Delta_n \sum_{i=1}^{a^n} \eta_i^n$. Now, obviously $Z_n \leq K$ and by the same argument as in the previous step for evaluating z_n we obtain $Z_n \to 0$ pointwise. Hence the last part of (38) holds when $j = 2$.

Finally $\mathbb{E}(|\widehat{\xi}(3)_i^n|^p \mid \mathcal{H}_0) \leq K(\Delta_i^n H)^{p/2}\, \mathbb{E}(\eta_i^n \mid \mathcal{H}_0)$, whereas $V(H, p/2)_n \leq KV(C, p/2)_n$ and $V(H, p/2)_n$ is $\mathcal{H}_0$-measurable. Thus $\mathbb{E}(R(3)''_n \mid \mathcal{H}_0) \leq K\mathbb{E}(z(p/2)''_n \mid \mathcal{H}_0)$, where $z(q)''_n$ is defined as $z(q)_n$, with H instead of $\overline{C}$. As seen above, $z(p/2)''_n$ is bounded and goes pointwise to 0 on the set $\{H_1 > 0\}$. Then $\mathbb{E}(R(3)''_n \mid \mathcal{H}_0) \to 0$ on this set, and $R(3)''_n = 0$ on the complement $\{H_1 = 0\}$. Thus $R(3)''_n \stackrel{\mathbb{P}}{\to} 0$, and the proof of the lemma is complete. □

Proof of Theorem 4. On the null set $\Omega^1_{\text{Itô}}$ we have $X_t = X_0 + X'_t$ for all $t \leq 1$, so the statistic $S(p)_n$ is the same as if it were constructed on the basis of observing X'. Hence the claim about the null hypothesis is Theorem 10.15 of Aït-Sahalia and Jacod (2014) under $\Omega_1^{(cW)}$ (the restriction $p > 3$ in that theorem is needed for the alternative consistency, whereas $p>0$ is enough for the behavior under the null, by virtue of Theorem A.2-(a) of the same reference).

For the alternative consistency, recalling that the critical region (26) has the form $\mathcal{C}(p)_n = \{S(p)_n < a^{p/2-1} - \rho\sqrt{S''(p)_n}\}$, where ρ is a positive constant and $S''(p)_n = V(X,2p;u_n)_n/V(X,p)_n^2$, it is obviously enough to show that,

$$\begin{array}{l} \lim_{\varepsilon\downarrow 0}\limsup_n \mathbb{P}\big(\{S(p)_n > a^{p/2-1}-\varepsilon\}\cap\Omega(p)^1_{\text{No-Itô}}\big)\to 0, \\ \mathbb{P}\big(\{S''(p)_n>\varepsilon\}\cap\Omega(p)^1_{\text{No-Itô}}\big)\to 0 \quad \text{for all } \varepsilon>0. \end{array} \tag{41}$$

Note that $S''(p)_n \le S'(p)_n = V(X,2p)_n/V(X,p)_n^2$. Two applications of (37) yield that the variables $S(p)_n/\Gamma_n$ and $S'(p)_n/\Gamma'_n$ converge in probability to 1 on the set $\{C_1>0\}\cap\Omega(p)\cap\Omega(2p)$, where $\Gamma_n = V(C,p/2)_{n-1}/V(C,p/2)_n$ and $\Gamma'_n = m_{2p}V(C,p)_n/(m_pV(C,p/2)_n)^2$. Since $\limsup_n \Gamma_n < a^{p/2-1}$ on $\Omega(p)'$ and $\Gamma'_n\to 0$ on $\Omega(p)$ by definition of these sets, we deduce (41). □

Proof of Theorem 5. Here we write $\overline{X} = X'+X''$. Up to modifying the stopping times T_n, it is no restriction to assume that each variable Z_n is non vanishing, hence $\underline{Z} = \inf_{n\ge 1:\,T_n\le 1}|Z_n|)$ (with the convention $\underline{Z}=1$ when X has no jump on $[0,1]$) is a positive variable.

The idea of the proof is to compare the two variables $V'(X,p)_n$ and $V(\overline{X},p)_n$. Let $i(n,q)$ be the only integer i such that $(i-1)\Delta_n < T_q \le i\Delta_n$ and set

$$\begin{array}{rl} \overline{\Omega}_n = \big\{ & \text{each interval } I(n,i) \text{ for } i=1,\cdots,a^n \text{ contains at most one jump} \\ & \text{of } J, \text{ and } T_{k_n}>1, \text{ and further } \sup_{1\le i\le a^n}|\Delta_i^n\overline{X}|\le (1/4)\,\underline{Z}\big\} \end{array}$$

Since $\overline{X}$ is continuous, we clearly have $\mathbb{P}(\overline{\Omega}_n)\to 1$.

On the set $\overline{\Omega}_n$, any $i(n,q)$ not bigger than a^n is necessarily an index i_j (recall (29)) for some $j>a^n-k_n$. It follows that, on $\overline{\Omega}_n$ and for $v_n=u_n$ or $v_n=\infty$ and with $A_n=\sup_{1\le i\le a^n}|\Delta_i^n\overline{X}|$,

$$V(\overline{X},p;v_n)_n - k_nA_n^p \le V'(X,p;v_n)_n \le V(\overline{X},p;v_n)_n. \tag{42}$$

First we consider what happens under the null. Without loss of generality we can thus assume $H_1=0$, and a classical result using Assumption (SH) yields that, for any $\varepsilon>0$, we have

$$\frac{A_n}{\Delta_n^{1/2-\varepsilon}} \stackrel{\mathbb{P}}{\to} 0. \tag{43}$$

Then (36) and (11) for (X',C') yield that

$$k_n\Delta_n^{1/2-p\varepsilon}\to 0 \quad\Rightarrow\quad \frac{k_nA_n^p}{\sqrt{\Delta_n}\;V(\overline{X},p)_n} \stackrel{\mathbb{P}}{\to} 0. \tag{44}$$

Combining this with (42) we see that, as soon as $k_n\Delta_n^\rho\to 0$ for some $\rho<\frac12$ and upon taking ε above small enough, the claim for $\mathcal{C}'(p)_n$ under the null amounts to the claim for $\mathcal{C}(p)_n$ in Theorem 4, for $\overline{X}$ instead of X, and under the null again.

Next we consider what happens under the alternative. The claim will follow by exactly the same argument as in Theorem 4, provided we prove the following version of (11), for $q=p$ and $q=2p$,

$$\frac{V'(X,q)_n}{V(C,q/2)_n} \stackrel{\mathbb{P}}{\to} m_q \quad \text{on the set } \{C_1>0\}\cap\Omega(q/2)''. \tag{45}$$

In view of (42) and of the fact that (11) holds with $\overline{X}$ instead of X, this amounts to proving

$$\frac{k_n A_n^q}{V(C,q/2)_n} \stackrel{\mathbb{P}}{\to} 0 \quad \text{on the set } \{C_1>0\}\cap\Omega(q/2)''. \tag{46}$$

We have $A_n \le A'_n + A''_n$, with $A'_n = \sup_{1\le i\le a^n} |\Delta_i^n X'|$ and $A''_n = \sup_{1\le i\le a^n} |\Delta_i^n X''|$, so it is sufficient to show (46) with A'_n and A''_n instead of A_n. For A'_n, either $C'_1=0$ and then $A'_n=0$ (recall $B'_t=0$ when $C'_t=0$), in which case there is nothing to prove, or $C'_1>0$ and (43) for A'_n holds, as well as (36): upon taking again ε small enough for, we get (46) for A'_n.

Finally, since $H_1>0$ under the alternative, arguing conditionally on the process H and since σ'' and r'' are bounded under Assumption (SH), we see that, instead of (43) we have for all $\varepsilon>0$,

$$\frac{A''_n}{\overline{H}_n^{1/2-\varepsilon}} \stackrel{\mathbb{P}}{\to} 0, \qquad \text{where } \overline{H}_n = \sup_{1\le i\le a^n} \Delta_i^n H.$$

It is thus enough to prove that, provided ε is small enough, we have

$$\frac{k_n \overline{H}_n^{q/2-q\varepsilon}}{V(H,q/2)_n} \stackrel{\mathbb{P}}{\to} 0 \qquad \text{on } \{H_1>0\}\cap\Omega(q/2)'',$$

which is obvious from $k_n\le Kn$ and from the definition of $\Omega(q/2)''$. □

REFERENCES

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

Aït-Sahalia, Y. and Jacod, J. (2014). *High Frequency Financial Econometrics.* Princeton University Press.

Dovgoshey, O., Martio, O., Ryazanov, V. and Vuorinen, M. (2006). The Cantor function. *Expositiones Mathematicae* **24**.

Duvernet, L., Robert, C. Y. and Rosenbaum, M. (2010). Testing the type of a semi-martingale: Itô against multifractal. *Electronic Journal of Statistics* **4** 1300–1323.

Itô, K. (1951). *On Stochastic Differential Equations.* No. 4 in Memoirs of the American Mathematical Society, AMS, Providence, Rhode Island.

Department of Economics
Princeton University and NBER
Princeton, NJ 08544-1021
E-mail: yacine@princeton.edu

Institut de Mathématiques de Jussieu
CNRS UMR 7586
Université P. et M. Curie (Paris-6)
75252 Paris Cédex 05
E-mail: jean.jacod@upmc.fr