Electronic Journal of Statistics
Vol. 4 (2010) 1300–1323
ISSN: 1935-7524
DOI: 10.1214/10-EJS585

Testing the type of a semi-martingale: Itō against multifractal

Laurent Duvernet

LAMA-Université Paris-Est Marne-la-Vallée and CMAP-École Polytechnique Paris
5 boulevard Descartes
77454 Marne-la-Vallée Cedex 2, France
e-mail: duvernet@cmap.polytechnique.fr

Christian Y. Robert

CREST-ENSAE
3 rue Pierre Larousse
92245 Malakoff Cedex, France
e-mail: chrobert@ensae.fr

and

Mathieu Rosenbaum

CMAP-École Polytechnique Paris
Route de Saclay
91128 Palaiseau Cedex, France
e-mail: mathieu.rosenbaum@polytechnique.edu

Abstract: We consider high frequency observations of a semi-martingale. From these data, we build simple test statistics allowing to distinguish between the two following situations: *i*) the data generating process is an Itō semi-martingale; *ii*) the data generating process is a Multifractal Random Walk. We also investigate the finite sample behavior of the test statistics on some simulated data.

AMS 2000 subject classifications: 60F05, 60G44, 62G10.
Keywords and phrases: Semi-martingales, multifractal processes, limit theorems, hypothesis testing.

Received February 2010.

1. Introduction

Semi-martingales are mathematically appealing quantities because they are stochastic processes which can be used as integrators in the general theory of stochastic integration. They are also natural modelling objects in various fields, especially in finance (for their link with the "no free lunch" assumption) and turbulence, see for example Delbaen and Schachermayer [11] and Barndorff-Nielsen and Schmiegel [7]. A semi-martingale is simply the sum of a local martingale and an adapted process with finite variation. Recall that any semi-martingale can be written as the sum of a predictable process of finite variation, a continuous local martingale, and a compensated pure jump process (the rigorous definition will be given below). The very widely used notion of *Itō semi-martingale* refers to the case where each of the following objects is absolutely continuous with respect to the Lebesgue measure: the finite variation process, the quadratic variation of the continuous local martingale, and the compensator of the jump measure.

A very large number of studies has been devoted to the statistical properties of Itō semi-martingales. Let us mention in particular a series of recent papers by Aït-Sahalia and Jacod [1, 2, 3] which will be of particular interest here. In these papers, the authors provide test statistics that address the following questions for Itō semi-martingales: Is the jump part of the semi-martingale equal to zero? Do the jumps have finite or infinite activity? Is the Brownian part equal to zero? A key element for the results of Aït-Sahalia and Jacod is the asymptotic behavior of the p-variation of the semi-martingale X, by which we mean the following quantity for $p > 0$ and some $n \in \mathbb{N}$ that goes to $+\infty$:

$$B(p, n^{-1}) = \sum_{i=1}^{n} |X_{i/n} - X_{(i-1)/n}|^p. \tag{1.1}$$

In this paper, our goal is to build test statistics aiming at answering questions that could be asked before the preceding ones. More precisely, we are looking for some statistical procedures allowing to say whether the data generating process is an Itō semi-martingale, against the alternative hypothesis that the data generating process belongs to a specific class of non-Itō semi-martingales, namely the Multifractal Random Walks of Bacry and Muzy [6] – and conversely. As explained below, the behavior of the p-variations will play a key role in our study.

This problem might appear surprising. Indeed, the class of Itō semi-martingales already yields a very large collection of models. However, in the past two decades, some authors have proposed a new class of models of non-Itō semi-martingales, namely multifractal processes. These processes have the nice feature of well reproducing most major stylized facts observed in finance or fully developed turbulence (in particular heavy-tail behavior, persistence and clustering of volatility, and intermittency of fluctuations), while remaining "simple" in the sense that they rely only on a small number of scalar parameters. For the introduction and pertinence of multifractal random models in turbulence and finance, we notably refer among others works to Frisch [14], Mandelbrot [24], Bouchaud and Potters [9], Bacry *et al.* [4], Calvet and Fisher [10].

In particular, Bacry *et al.* [5], Calvet and Fisher [10], and Duchon *et al.* [12] provide a thorough discussion of the multifractal approach to volatility modelling and pricing at various time scales. These authors notably show that multifractal models lead to quite superior volatility or VaR forecasts than the more usual methods based on GARCH, MS-GARCH and FIGARCH models – even when the former are calibrated out of sample and the latter are calibrated in sample. This indeed suggests that the multifractal setting should be of high interest for providing an accurate mathematical model of the dynamic of financial assets.

A distinctive property of these multifractal processes is the scaling behavior of their moments: for all p's in some real interval $I \supset [0,2]$ and $t \geq 0$,

$$\mathbb{E}[|X_{t+s} - X_t|^p] \sim \gamma(p) s^{\tau(p)+1} \text{ as } s \to 0, \tag{1.2}$$

where $p \mapsto \tau(p)$ is a *strictly concave* function and the $\gamma(p)$'s are some positive constants. The term multifractal, or multifractal scaling behavior, refers to the nonlinearity of the scaling exponent $\tau(\cdot)$. Therefore one would expect that the relation

$$n^{-1} B(p, n^{-1}) \approx \gamma(p) n^{-(\tau(p)+1)} \text{ for large } n \tag{1.3}$$

holds for this class of processes, where the p-variation $B(p, n^{-1})$ is as in (1.1). Note that if X is a continuous Itō semi-martingale, one would obtain a *linear* exponent $\tau(p) = p/2 - 1$ for all $p \geq 0$. Thus, when confronted to observations, it is natural to consider p-variations in order to assert whether the exponent τ is linear or not.

It should also be mentioned that the interest for the nonlinear nature $p \mapsto \tau(p)$ has rapidly grown since the seminal paper by Frisch and Parisi [15] who conjectured that this function $\tau(\cdot)$ in (1.3) characterizes the wildly varying pointwise Hölder regularity of the underlying function $t \mapsto X_t(\omega)$. Following this initial definition of the multifractal paradigm by Frisch and Parisi, the past two decades have then seen a large production of empirical studies in turbulence and finance, but also in DNA analysis or internet traffic among other fields, which base themselves notably on (1.3) to investigate the multifractal nature of the data – see for instance respectively for each of these four fields Gagne *et al.* [16], Ghashghaie *et al.* [17], Yu *et al.* [31] and Park and Willinger [27].

Nevertheless, it is important to note that these empirical works rarely rely on explicit random models; indeed, only a few research papers have directly addressed the issue of detecting the nonlinear nature of τ. This is notably the case of the works by Wendt, Abry and Jaffard [29, 30], who examine the performances on simulations of some specific algorithms that attempt to state whether some given signal is of "monofractal" or "multifractal" regularity. Note however that while it is a similar issue to the one we consider here, the classes of signal that these authors consider do not coincide with the ones that we study in the present work: for instance their "monofractal" regularity class contains fractional Brownian motions (which we do not consider here), but not all Itō type semi-martingales.

Some previous theoretical studies have already been devoted to the statistical properties of multifractal processes, see notably Ossiander and Waymire [26], Gloter and Hoffmann [18], Ludeña [22], Duvernet [13]. However, the elaboration of a probabilistic test of multifractal scaling behavior in the sense of (1.2) (assuming this informal relation is given a rigorous meaning) has apparently not been explicitly considered yet, with the exception of the studies in [29, 30] already mentioned. We address here this problem in the limited setting of multifractal processes that belong to the class of Multifractal Random Walks (MRW's for short) introduced by Bacry and Muzy [6]. These processes have the nice theoretical property that the scaling relation (1.2) is satisfied with an exact equality for all s in some real interval $[0, T]$. We give in Section 2 the proper definition of these MRW processes.

In this paper, we do not aim at being very accurate in term of statistical testing theory. We are just looking for simple quantities which have "opposite" behaviors when the data generating process X is an Itō semi-martingale or an MRW. More precisely, we are looking for two statistics, say T_1 and T_2, associated to the null assumption H_0 that X is an Itō semi-martingale ($X =$ Itō for short) and to the null assumption H_0 that X is an MRW ($X =$ MRW for short) such that, when $X =$ Itō (resp. $X =$ MRW), the asymptotic law of T_1 (resp. T_2) is non degenerate and known and T_2 (resp. T_1) goes to a degenerate limit.

The paper is organized as follows. In Section 2, we give a brief introduction to semi-martinagles and MRW's. We build our test statistics and give their asymptotic behaviors in Section 3. These statistics are based on suitably chosen p-variations of the process. An intensive simulation study can be found in Section 4. The proofs are relegated to Section 5. This paper is not technically very innovative and the results could probably be improved, for example using p-variations of higher orders. However, we believe it is a first step in order to solve this new problem.

2. Definitions

2.1. Semi-martingales and Itō semi-martingales

A real valued process X defined on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ is called a semi-martingale if it can be decomposed as $X = X_0 + M + A$ where X_0 is finite valued and $\mathcal{F}_0$-measurable, M is a local martingale on this space and A is an adapted process of finite variation. Any semi-martingale can be written as

$$X_t = X_0 + A_t^p + X_t^c + \int_0^t \int_{\mathbb{R}} \kappa(x)(\mu - \nu)(\mathrm{d}s, \mathrm{d}x) + \int_0^t \int_{\mathbb{R}} \kappa'(x)\mu(\mathrm{d}s, \mathrm{d}x),$$

where

- A^p is a predictable process of finite variation;
- X^c is a continuous local martingale with $X_0^c = 0$, called the "continuous martingale part" of X;
- μ is the "jump measure" of X;
- ν is the "compensator" of μ;
- κ is a continuous function with compact supports such that $\kappa(x) = x$ for all x in a neighborhood of 0 and $\kappa'(x) = x - \kappa(x)$.

With this notation, the decomposition is unique (up to null sets), but the process A^p depends on the choice of the truncation function κ. Let us denote by Σ^2 the quadratic variation of the "continuous martingale part" X^c. The triple (A^p, Σ^2, ν) is called the triple of characteristics of X because, in "good cases" (see [20]), it completely determines the law of X.

An Itō semi-martingale is a semi-martingale whose characteristics are absolutely continuous with respect to the Lebesgue measure in the following sense

$$A^p_t(\omega) = \int_0^t a_s(\omega)\mathrm{d}s, \quad \Sigma^2_t(\omega) = \int_0^t \sigma^2_s(\omega)\mathrm{d}s, \quad \nu(\omega, \mathrm{d}t, \mathrm{d}x) = \mathrm{d}t\ F_{\omega,t}(\mathrm{d}x),$$

where a, σ are optional and $F_t(C)$ is optional for all Borel subsets C of $\mathbb{R}$. Itō semi-martingales have a nice representation in terms of a Wiener process and a Poisson random measure

$$\begin{aligned} X_t = X_0 + \int_0^t a_s\mathrm{d}s + \int_0^t \sigma_s\mathrm{d}W_s + \int_0^t \int_{\mathbb{R}} \kappa \circ \delta(s,x)(\underline{\mu} - \underline{\nu})(\mathrm{d}s, \mathrm{d}x) \\ + \int_0^t \int_{\mathbb{R}} \kappa' \circ \delta(s,x)\underline{\mu}(\mathrm{d}s, \mathrm{d}x), \end{aligned} \tag{2.1}$$

where W denotes a $(\mathcal{F}_t)$-standard Wiener process and $\underline{\mu}$ is a $(\mathcal{F}_t)$-Poisson random measure on $(0,\infty) \times \mathbb{R}$ with intensity measure $\underline{\nu}(\mathrm{d}t,\mathrm{d}x) = \mathrm{d}t \otimes \lambda(\mathrm{d}x)$, where λ is a σ-finite and infinite measure without atom.

2.2. Presentation of the log-normal MRW

We first give an informal presentation of the MRW model in a specific Gaussian case. Let us define the process X as a Brownian motion in random time,

$$X_t = B_{\theta_t}, \quad t \geq 0,$$

where the process θ is increasing and independent of the standard $(\mathcal{F}_t)$-Wiener process B and can be for instance interpreted as an aggregated volatility. Let us moreover suppose that we are in a relatively simple case where we have

$$\theta_t = v^2 \int_0^t e^{w(u)}\mathrm{d}u, \quad t \geq 0,$$

for some stationary, $(\mathcal{F}_t)$-adapted, Gaussian process $\big(w(t)\big)_t$ such that $\mathbb{E}\big[e^{w(s)}\big] = 1$, $s \geq 0$, and for some constant $v^2 > 0$ that is simply an average level for θ: $\mathbb{E}\big[\theta_t\big] = v^2 t$. Then clearly, if we want the volatility of the process X to have some persistence property – as it would be the case on financial data, then the stationary Gaussian process w should have a slowly decaying autocovariance. Define a time window $[0,T]$ for some $T > 0$, the MRW setting then consists in specifying the following autocovariance function

$$\mathbb{C}ov[w(s), w(t)] = \lambda^2 \max\big(\log\big(T/|t-s|\big), 0\big) \tag{2.2}$$

for some constant $\lambda^2 > 0$. The parameter λ^2 can be interpreted as a "quantity" of multifractality (indeed, in the degenerate case $\lambda^2 = 0$ we obtain the very basic model $X_t = vB_t$), while the parameter $T > 0$ is a decorrelation scale, which is evaluated as a few months or a few years in the case of financial data.

Of course, the specification of autocovariance in (2.2) does not lead to a well defined random model since it implies that $\mathbb{V}[w(t)] = +\infty$. Nevertheless, this approach can be made rigorous by either defining w as a generalized Gaussian random process (see the presentation of the MRW model in Duchon *et al.* [12] which is based on the multiplicative chaos of Kahane in [21]), or by introducing a family of Gaussian processes $\big(w_l(t)\big)_{l,t}$ such that

$$\mathbb{C}\mathrm{ov}[w_l(s), w_l(t)] \uparrow \lambda^2 \max\big(\log\big(T/|t-s|\big), 0\big) \quad \text{as } l \to 0 \tag{2.3}$$

and defining θ as the limit $\theta_t = \lim_{l\to 0} v^2 \int_0^t e^{w_l(u)} \mathrm{d}u$. Note that in this approach, obtaining (1.2) for some even integer $p \geq 0$ follows from a simple application of Fubini's theorem and the monotone convergence theorem: one then finds $\tau(p) = p/2 - 1 + \lambda^2 p/4 - \lambda^2 p^2/8$. This construction of X through the limit $l \to 0$ has been proposed by Bacry and Muzy in [6] who also extended it to the framework of non Gaussian, infinitely divisible processes w. We present this construction in Section 2.3.

Let us also remark that in this construction, for all $u \geq 0$, we have that the random variable $e^{w_l(u)}$ goes to 0 in probability as $l \to 0$ while its moment of order $p > 1$ goes to $+\infty$. However, following Bacry and Muzy, we have that the limit $\theta_t = \lim_{l\to 0} v^2 \int_0^t e^{w_l(u)} \mathrm{d}u$ is valid and nondegenerate. It can more generally be shown that *the continuous, increasing process θ has actually no derivative with respect to the Lebesgue measure* (that is to say, almost surely, the corresponding random measure $\theta[s,t] = \theta_t - \theta_s$ for $0 \leq s \leq t$ is singular: it has no absolutely continuous nor discrete component.) It follows that X is a *non-Itō* continuous $(\mathcal{F}_t)$-martingale.

Finally, we mention that this log-normal MRW should provide a particularly parsimonious model that allows one to reproduce most of the well documented stylized facts observed on price fluctuations of assets. Bacry *et al.* [5] considered 29 of the largest French stocks of the Euronext market and estimated the three parameters of the log-normal MRW (λ^2, T, v^2) using various methods. At 5% confidence level, they find that all the considered stock return series are multifractal with a small intermittency coefficient $\lambda^2 \simeq 0.1$. The ability of the log-normal MRW model to forecast volatility and conditional Value at Risk is also studied (see also Duchon *et al.* [12]). It appears that the MRW-based estimation procedures outperform both GARCH or tGARCH models at any horizon and any time scale.

2.3. General construction and properties of the Multifractal Random Walks

Following Bacry and Muzy [6], we now present the construction of the class of MRW's. Fix an infinitely divisible distribution $\pi(\mathrm{d}x)$ on $\mathbb{R}$. Let ψ be the Laplace exponent of π

$$e^{\psi(q)} = \int_{\mathbb{R}} e^{qx} \pi(\mathrm{d}x)$$

for $q \geq 0$ (possibly $\psi(q) = \infty$). We assume the following on ψ

$$\psi(1) = 0, \quad \text{and} \quad \psi'(1) < 1. \tag{2.4}$$

Let μ be the measure on the open half-plane $\mathbb{R} \times (0, \infty)$ given by $\mu(\mathrm{d}t, \mathrm{d}l) = l^{-2}\mathrm{d}t \otimes \mathrm{d}l$. We now assume that we have an infinitely divisible, independently scattered random field P on $\mathbb{R} \times (0, \infty)$ with intensity μ and Laplace exponent ψ, that is

- for every Borel set A in $\mathbb{R} \times (0, \infty)$, $P(A)$ is an infinitely divisible random variable such that
$$\mathbb{E}[e^{qP(A)}] = e^{\mu(A)\psi(q)}$$
for every $q \geq 0$ such that $\psi(q) < \infty$,
- for every sequence $\{A_k\}_{k \in \mathbb{N}}$ of disjoint Borel sets in $\mathbb{R} \times (0, \infty)$, the variables $P(A_k)$ are independent and
$$P(\cup_{k \in \mathbb{N}} A_k) = \sum_{k \in \mathbb{N}} P(A_k) \text{ almost surely.}$$

Let $A_l(t)$ be the "cone" in $\mathbb{R} \times (0, \infty)$ defined by

$$A_l(t) = \left\{(t', l') \in \mathbb{R} \times (0, \infty),\, l \leq l' \text{ and } |t - t'| \leq \frac{1}{2}\min(l', T)\right\}$$

and let $A_l(t, s) = A_l(t) \cap A_l(s)$ the intersection of two cones. Then note that $\mu(A_l(t, s)) = \log\bigl(T/|t-s|\bigr)$ for $l \leq |t-s| \leq T$ and $\mu(A_l(t, s)) = 0$ for $|t-s| > T$.

We also assume that B and P are independent. Bacry and Muzy [6] then proved that the following process is nondegenerate

$$\theta_t = \lim_{l \to 0} v^2 \int_0^t e^{P(A_l(u))} \mathrm{d}u, \quad t \geq 0.$$

The process θ has continuous, positive and increasing sample paths, possesses stationary increments and satisfies $\mathbb{E}[\theta_t] = v^2 t$ for $t \geq 0$. We now define the MRW process as a subordinated Brownian motion

$$(X_t)_{t \geq 0} = (B_{\theta_t})_{t \geq 0}$$

where $v^2 > 0$ is the mean level of the volatility. When the random field P is a 2d-Gaussian white noise, we obtain the log-normal MRW described above.

The subordination of a Brownian process with a non decreasing process is not new and has been introduced by Mandelbrot and Taylor [25]. The MRW can also be understood as a Brownian motion in a "multifractal time" θ_t or as a limit of a stochastic integral since

$$(X_t)_{t \geq 0} \stackrel{d}{=} \lim_{l \to 0} \left(v \int_0^t e^{P(A_l(u))/2} \mathrm{d}B_s \right)_{t \geq 0}.$$

It is easy to see that it is a continuous martingale with respect to $\mathcal{F}$, but not an Itō semi-martingale since θ is not absolutely continuous with respect to the Lebesgue measure – see for instance [8].

Finally, we define p^* as $\sup\{p \geq 2, \psi(p/2) < p/2 - 1\}$. Since ψ is strictly convex (as a log-Laplace transform), (2.4) yields $p^* > 2$. Bacry and Muzy show that for $0 \leq p < p^*$ and $0 \leq t \leq T$,

$$\mathbb{E}\big[|X_t|^p\big] = \gamma(p) t^{p/2-\psi(p/2)}, \tag{2.5}$$

where $\gamma(p)$ is a positive constant. Thus, X satisfies (1.2) with

$$\tau(p) = p/2 - \psi(p/2) - 1.$$

We define $p_{max} = \sup\{p \geq 2, p\tau'(p) > \tau(p)\}$. Then it is easy to check the following facts from (2.4): $\tau(0) = -1$, $\tau(2) = 0$, $p_{max} > 2$ and $0 < \tau(4) < 1$ provided $p_{max} > 4$ (also, from basic convexity consideration, $p_{max} \leq p^*$). Let $\mu(p)$ be the absolute moment of order p of a centered standard Gaussian variable. The following lemma shows that the rates of convergence of the p-variations of X differ from those of an Itō semi-martingale.

Lemma 1. *(Duvernet, [13]) For $t > 0$ and $0 \leq p < p_{max}$, almost surely:*

$$2^{N\tau(p)} \sum_{i=1}^{\lfloor 2^N t \rfloor} |X_{i2^{-N}} - X_{(i-1)2^{-N}}|^p \to \mu(p)\theta_t^{(p)},$$

$$2^{N\tau(2p)} \sum_{i=1}^{\lfloor 2^N t \rfloor} |\theta_{i2^{-N}} - \theta_{(i-1)2^{-N}}|^p \to \theta_t^{(2p)},$$

as $N \to +\infty$, where $\theta_t^{(p)}$ and $\theta_t^{(2p)}$ are some positive random variables, independent of the Wiener process B.

Finally, we will speak of a *log-normal MRW* when the random field $P(\mathrm{d}t, \mathrm{d}l)$ is simply a 2d Gaussian white noise with expectation $-\lambda^2/2\mu(\mathrm{d}t, \mathrm{d}l)$ and variance $\lambda^2\mu(\mathrm{d}t, \mathrm{d}l)$ for some $\lambda^2 > 0$ (from (2.4) we also require $\lambda^2 < 2$). It is then straightforward to check that we then have $\tau(p) = p/2 - 1 - \lambda^2 p(p-2)/8$ and $p_{max} = 2\sqrt{2}/\lambda$, and that the autocovariance of the process $\big(P(\mathcal{A}_l(t))\big)_t$ is asymptotically given by (2.3) when $l \to 0$.

3. Statistical problem and results

As usual in the multifractal context (see also Lemma 1), we consider $n = 2^N$ and our asymptotic will be N goes to infinity. The semi-martingale is either an MRW as described in the preceding section or, following Aït-Sahalia and Jacod [1], an Itō semi-martingale of the form (2.1). The coefficients $a_t(\omega)$, $\sigma_t(w)$ and $\delta(\omega, t, x)$ are such that the various integrals in Equation (2.1) make sense and we will always assume that σ_t is also an Itō semi-martingale, of the form

$$\sigma_t = \sigma_0 + \int_0^t \tilde{a}_s \mathrm{d}s + \int_0^t \tilde{\sigma}_s \mathrm{d}W_s + \int_0^t \tilde{\sigma}'_s \mathrm{d}W'_s$$
$$+ \int_0^t \int_E \kappa \circ \tilde{\delta}(s,x)(\underline{\mu} - \underline{\nu})(\mathrm{d}s, \mathrm{d}x) + \int_0^t \int_E \kappa' \circ \tilde{\delta}(s,x)\underline{\mu}(\mathrm{d}s, \mathrm{d}x), \qquad (3.1)$$

where W' is another Wiener process independent of $(W, \underline{\mu})$. Let $\delta'_t(\omega) = \int_0^t \kappa \circ \delta(\omega, t, x)\lambda(\mathrm{d}x)$ if the integral makes sense and $+\infty$ otherwise. Finally, set $t_{inf} = \inf\{t, \Delta X_t \neq 0\}$. As in [1], we will systematically consider the following assumption for X when X is an Itō semi-martingale[1].

Assumption 1.

- *All paths $t \to a_t(\omega)$, $t \to \tilde{\sigma}_t(\omega)$, $t \to \tilde{\sigma}'_t(\omega)$, $t \to \delta(\omega, t, x)$, $t \to \tilde{\delta}(\omega, t, x)$ are left-continuous with right limits.*
- *All paths $t \to \tilde{a}_t(\omega)$, $t \to sup_{x \in E} \frac{|\delta(\omega,t,x)|}{\gamma(x)}$ and $t \to sup_{x \in E} \frac{|\tilde{\delta}(\omega,t,x)|}{\tilde{\gamma}(x)}$ are locally bounded, where γ and $\tilde{\gamma}$ are (non random) nonnegative functions satisfying $\int_E (\gamma(x)^2 \wedge 1)\lambda(dx) < \infty$, $\int_E (\tilde{\gamma}(x)^2 \wedge 1)\lambda(dx) < \infty$.*
- *All paths $t \to \delta'_t(\omega)$ are left-continuous with right limits on the semi open set $[0, t_{inf}(\omega))$.*
- *We have $\int_0^t \sigma_s^2 ds > 0$, a.s., for all $t > 0$.*

3.1. The case H_0: X = Itō

We are looking for a statistic whose behavior is different when X = Itō and when X = MRW. To build this statistic, we will use p-variations of the form (1.1). Before explaining why such quantities are natural in our problem, we need to define the two following sets:

$$\Omega^j = \{\omega, s \to X_s(w) \text{ is discontinuous on } [0,1]\},$$
$$\Omega^c = \{\omega, s \to X_s(w) \text{ is continuous on } [0,1]\}.$$

Remark that the sample path of an Itō semi-martingale can be in Ω^c even if this semi-martingale is not continuous (if no jump occurs before $t = 1$).

Let us consider $p > 2$. From Jacod [19], we know that if X =Itō, in restriction to Ω^j, then the jumps dominate and $B(p, 2^{-N})$ goes to $\sum_{t \leq 1} |\Delta X_t|^p$ as $N \to \infty$. In restriction to Ω^c, then

$$2^{N(p/2-1)} B(p, 2^{-N}) \xrightarrow{\mathbb{P}} \mu(p) \int_0^1 |\sigma_t|^p \mathrm{d}t, \qquad \text{as } N \to \infty.$$

On the other hand, if X = MRW,

$$2^{Nc(p)} B(p, 2^{-N}) \xrightarrow{\mathbb{P}} \mu(p)\theta_1^{(p)}.$$

[1]In order to make the paper self contained, we rewrite here the definitions and assumptions in [1].

Thus, in the spirit of Aït-Sahalia and Jacod [1] and Rosenbaum [28], we naturally consider for some $p > 2$ the ratio

$$\frac{B(p, 2^{-(N-1)})}{B(p, 2^{-N})}. \tag{3.2}$$

If $X =$ Itō, this tends to 1 in restriction to Ω^j and to $2^{p/2-1}$ in restriction to Ω^c. When X is a MRW, it goes to $2^{\tau(p)}$. Now, to have a feasible test, we need a central limit theorem (CLT) associated to this quantity. Before stating the results, we need to recall the definition of stable convergence in law. We say that a sequence T_n on $(\Omega, \mathcal{F}, \mathbb{P})$ converges stably in law to the law ϕ ($T_n \overset{\mathcal{L}_s}{\to} \phi$), in restriction to $A \in \mathcal{F}$, if for all bounded continuous functions f and all $\mathcal{F}-$measurable bounded variables Y vanishing outside A,

$$\mathbb{E}[f(T_n)Y] \to \mathbb{E}[Y]\mathbb{E}[f(U)], \text{ with } U \text{ a random variable with law } \phi.$$

Let us also define the constant $m(p)$ as

$$m(p) = \left(\frac{\mu^2(p)}{2^{p-2}(3\mu(2p) + \mu^2(p)) - 2^{p/2}\tilde{\mu}_p} \right)^{1/2},$$

with

$$\tilde{\mu}_p = \mathbb{E}\big[|U|^p |U + V|^p\big]$$

for U and V some independent standard $\mathcal{N}(0, 1)$ random variables.

Then from Aït-Sahalia and Jacod [1], if $X =$ Itō, in restriction to the set Ω^c, we have for $p > 2$

$$m(p) \frac{B(p, 2^{-N})}{(B(2p, 2^{-N}))^{1/2}} \big(\frac{B(p, 2^{-(N-1)})}{B(p, 2^{-N})} - 2^{p/2-1} \big) \overset{\mathcal{L}_s}{\to} \mathcal{N}(0, 1).$$

The term

$$m(p) \frac{B(p, 2^{-N})}{(B(2p, 2^{-N}))^{1/2}} 2^{-N/2}$$

corresponds to an estimator of the asymptotical standard deviation of the ratio in Theorem 3b) in [1].

However in restriction to the set Ω^j, we have the following convergence in probability

$$\begin{aligned} m(p) &\frac{B(p, 2^{-N})}{(B(2p, 2^{-N}))^{1/2}} \big(\frac{B(p, 2^{-(N-1)})}{B(p, 2^{-N})} - 2^{p/2-1} \big) \\ &\overset{\mathbb{P}}{\to} -m(p) \frac{\sum_{t \le 1} |\Delta X_t|^p}{(\sum_{t \le 1} |\Delta X_t|^{2p})^{1/2}} (2^{p/2-1} - 1). \end{aligned}$$

This result can not be used to build a convenient test statistic. So, we finally choose the following slightly modified test statistic:

$$T_1^N = m(p) 2^{(p/2-1)(\lfloor k_N N \rfloor - N)} \frac{B(p, 2^{-\lfloor k_N N \rfloor})}{(B(2p, 2^{-N}))^{1/2}} \big(\frac{B(p, 2^{-(N-1)})}{B(p, 2^{-N})} - 2^{p/2-1} \big),$$

where (k_N) is a positive sequence such that $k_N \leq 1$ for all N, $k_N \to 1$ and $(1-k_N)N \to \infty$ as $N \to \infty$. The following theorem shows that T_1^N is a suitable test statistic when the null hypothesis is H_0: $X =$ Itō. Its proof is just a direct application of the results of Aït-Sahalia and Jacod [1] for the Itō case. In the MRW case, it follows from Lemma 1 together with the fact that $\tau(2p)/2 > \tau(p)$ provided $p_{max} > 2p$.

Theorem 1. *Let $p > 2$.*

- *Assume X is an Itō semi-martingale such the above assumptions hold.*
  - *In restriction to the set Ω^c, $(T_1^N)^2 \xrightarrow{\mathcal{L}_s} \chi^2(1)$.*
  - *In restriction to the set Ω^j, $(T_1^N)^2 \xrightarrow{\mathbb{P}} 0$.*
- *Assume X is an MRW. If $p_{max} > 2p$, then $(T_1^N)^2 \xrightarrow{\mathbb{P}} +\infty$.*

Remark 1: If we restrict ourself to continuous Itō semi-martingales in H_0, we can choose $k_N = 1$.

Eventually, we can suggest the following rejection area for the test of asymptotic level α in the case where H_0 is $X =$ Itō: $\{(T_1^N)^2 \geq z_{1-\alpha}\}$, where $z_{1-\alpha}$ is the $1-\alpha$ quantile of a $\chi^2(1)$ distribution: $\mathbb{P}[X \leq z_{1-\alpha}] = 1-\alpha$ if X has a $\chi^2(1)$ distribution.

3.2. The case H_0: X = MRW

We now need to find a test statistic in the case H_0: $X =$ MRW. This statistic should satisfy a CLT under the MRW assumption and go to some degenerate limit under the Itō assumption. One idea would be to use once again quantities of the form (3.2) to estimate $\tau(p)$ for $p \neq 2$ (since for $p = 2$, the ratio goes to 1 under the Itō and the MRW assumption). However, to our knowledge, the available results on the asymptotic behavior of the p-variations under the MRW assumption, see [23], do not enable to obtain CLTs for quantities of the form

$$\frac{B(p, 2^{-(N-1)})}{B(p, 2^{-N})} - 2^{\tau(p)}.$$

We suggest another strategy which is based on quadratic variations. What we use is the difference between the rates of convergence of the quadratic variations under the MRW and the Itō assumption. We first consider the case where the null assumption is that $X =$ MRW with a given value for $\tau(4)$.

3.2.1. The case H_0: X = MRW and $\tau(4) = \tau_4$

Here we assume that under the null, X is an MRW where $\tau(4) = \tau_4$ is a given value smaller than 1. Let us first consider the following statistic

$$V^N = \frac{\sqrt{3}}{\sqrt{2(2^{\tau_4}-1)}} \frac{\{B(2, 2^{-N}) - B(2, 2^{-(N-1)})\}}{\sqrt{B(4, 2^{-N})}}.$$

The following proposition is proved in Section 5.1

Proposition 1. *If X is an MRW with $p_{max} > 4$, then $V^N \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$, as $N \to \infty$.*

Now, to assess if V^n is a suitable test statistic, we have to look at its behavior when X is an Itō semi-martingale. From Jacod [19] (see also [1]), we know that in restriction to the set Ω^c, V^N is of order 1 and in restriction to the set Ω^j, V^N goes to zero in probability. Therefore, the preceding statistic is suitable only if, under the alternative, the sample path has jumps on $[0,1]$. To solve this issue, we use an alternative estimate for the asymptotic variance in the CLT for the difference of quadratic variations (see Proposition 3 in Section 5.1). Indeed, this new estimator has different rates of convergence when $X = \text{MRW}$ and when $X = \text{Itō}$. More precisely, we estimate this variance using only $2^{\lfloor kN \rfloor}$ data instead of 2^N for some $k \in (0,1)$. Hence we consider the following statistic

$$T_2^N = \frac{\sqrt{3}}{\sqrt{2(2^{\tau_4}-1)}} 2^{(N-\lfloor kN \rfloor)\tau_4/2} \frac{\{B(2,2^{-N}) - B(2,2^{-(N-1)})\}}{\sqrt{B(4,2^{-\lfloor kN \rfloor})}}.$$

If X is an Itō semi-martingale, in restriction to the set Ω^c, the order of magnitude of T_2^N is $2^{(N-\lfloor kN \rfloor)(\tau_4-1)/2}$ and restriction to the set Ω^j, this order is $2^{(N-\lfloor kN \rfloor)(\tau_4-1)/2}2^{-\lfloor kN \rfloor/2}$. Thus, we finally get the following result, which is easily derived from Proposition 1.

Theorem 2.

- *If X is an MRW with $\tau(4) = \tau_4$ and $p_{max} > 4$, then $(T_2^N)^2 \xrightarrow{\mathcal{L}} \chi^2(1)$.*
- *For $k \in (0,1)$, if X is an Itō semi-martingale, then $(T_2^N)^2 \xrightarrow{\mathbb{P}} 0$.*

Remark 3: If we restrict ourself to sample paths with jumps in the alternative, we can take $k = 1$.

We can suggest the following rejection area for the test of asymptotic level α in the case where H_0 is $X = \text{MRW}$ and $\tau(4) = \tau_4$: $\{(T_2^N)^2 \leq z_\alpha\}$.

3.2.2. The case H_0: X = MRW with unknown $\tau(4)$

We now want to build a test statistic without assuming that $\tau(4)$ is known. When $X = \text{MRW}$, a natural convergent estimator of $\tau(4)$, providing an immediate equivalent for $2^{N\tau(4)}$, is given by

$$\hat{\tau}(4) = \frac{2}{N\log(2)}\big\{\log\big(B(4,2^{-\lfloor N/2 \rfloor})\big) - \log\big(B(4,2^{-N})\big)\big\}.$$

However, this estimator is not really convenient since it might tend to 0 or 1 if $X = \text{Itō}$. Thus we use the following modification of $\hat{\tau}(4)$

$$\tau^*(4) = (\hat{\tau}(4) \wedge (1 - v_N)) \vee u_N,$$

with u_N and v_N two positive sequences tending to 0 such that $2^N u_N \to +\infty$, Nu_N is bounded and $Nv_N \to +\infty$. Thanks to the sequences u_N and v_N, $\tau^*(4)$ can not tend to 0 or 1 too rapidly. We have the following proposition which is proved in Section 5.2.

Proposition 2. *If X is an MRW with $p_{max} > 4$, we have as $N \to +\infty$*

$$2^{N(\tau^*(4)-\tau(4))/2} \to 1.$$

If X is an Itō semi-martingale, we have as $N \to +\infty$

$$\frac{2^{N(\tau^*(4)-1)/2}}{\sqrt{2(2^{\tau^*(4)}-1)}} \to 0.$$

We now naturally define our last test statistic the following way:

$$\widetilde{T}_2^N = \frac{\sqrt{3}}{\sqrt{2(2^{\tau^*(4)}-1)}} 2^{(N-\lfloor kN \rfloor)\tau^*(4)/2} \frac{\{B(2,2^{-N}) - B(2,2^{-(N-1)})\}}{\sqrt{B(4,2^{-\lfloor kN \rfloor})}}.$$

We can now state our last result which is easily deduced from Proposition 1 and Proposition 2.

Theorem 3.

- *If X is an MRW with $p_{max} > 4$, then $(\widetilde{T}_2^N)^2 \xrightarrow{\mathcal{L}} \chi^2(1)$.*
- *For $k \in (0,1)$, if X is an Itō semi-martingale, then $(\widetilde{T}_2^N)^2 \xrightarrow{\mathbb{P}} 0$.*

Eventually, we can suggest the following rejection area for the test of asymptotic level α in the case where H_0 is $X =$ MRW: $\{(\widetilde{T}_2^N)^2 \leq z_\alpha\}$.

4. A simulation study

4.1. The setting

We begin here with some illustrations of our test procedures. For some integer $N \geq 1$, we simulated 100 times a sequence $X_{2^{-N}} - X_0, \dots, X_1 - X_{1-2^{-N}}$ where X is one of the following:

- A standard Brownian motion on $[0,1]$
- An Itō semi-martingale which is the sum of a standard Brownian motion and a compound Poisson process: $X_t = W_t + \sum_{k=0}^{N_t} A_k$, given that there is at least one jump. The Poisson process N has an intensity of 30 (N_t jumps 30 times on average for $0 \leq t \leq 1$) and the A_k's are uniformly distributed on the interval $[-1/2, 1/2]$. The values 30 and 1/2 have been chosen such that the sample path of the Itō semi-martingale seems at least visually hard to discern from the sample paths of the MRW, see Figures 1 and 3.
- An MRW as described in Section 2, with $\psi(p) = \lambda^2 p(p-1)/2$ for some $\lambda^2 \in (0,2)$ (thus the random field $P(\mathrm{d}t, \mathrm{d}l)$ is a 2D Gaussian white noise with expectation $-\lambda^2 l^{-2}/2\mathrm{d}t \otimes \mathrm{d}l$ and variance $\lambda^2 l^{-2}\mathrm{d}t \otimes \mathrm{d}l$). We consider

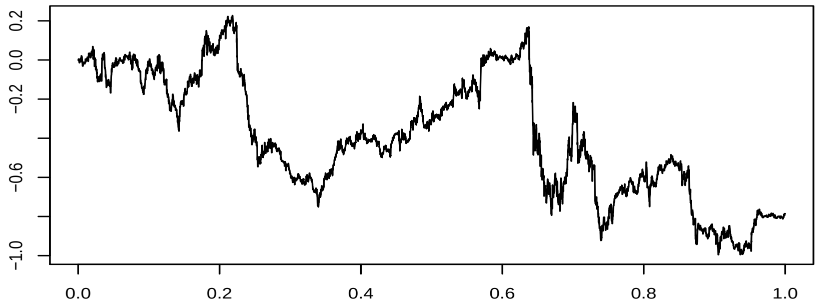

Fig 1. *A sample path of the log-normal MRW process on [0,1],* $\lambda^2 = 0.1$.

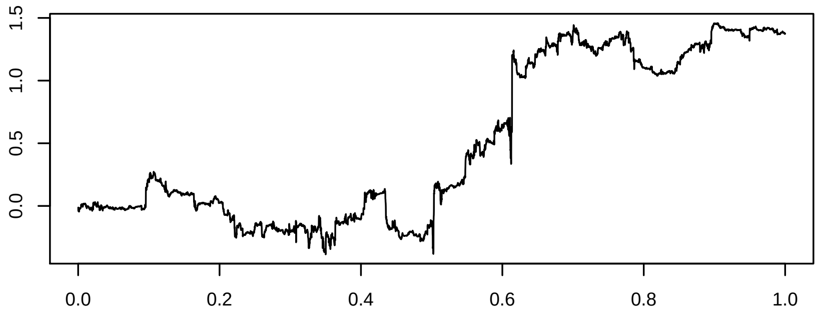

Fig 2. *A sample path of the log-normal MRW process on [0,1],* $\lambda^2 = 0.7$.

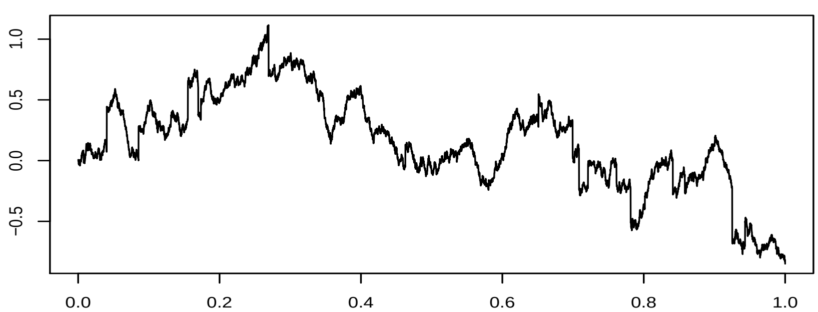

Fig 3. *A sample path of the Brownian motion with Poissonian jumps on [0,1].*

three possible values for λ^2: 0.02, 0.1, or 0.7. When modelling financial data, a common range for λ^2 would roughly be $[0.08, 0.20]$, see Bacry *et al.* [5]. The parameter T and v are both set to 1, which is of little consequence here, see again [5]. For this choice of MRW, we have $\tau(p) = p/2 - 1 - \lambda^2 p(p-2)/8$ and $p_{max} = 2\sqrt{2}/\lambda$. We refer to Bacry and Muzy [6] for the simulation procedure.

4.2. Case H_0: X = Itō

When using the test procedure in practice, it is of first importance to have some idea of its power, that is how fast the test statistic does manifest a degenerate behavior under the alternate hypothesis. Let us therefore give some orders of magnitude for the statistic T_1^N in the case where the null hypothesis is false, that is the data generating process X is an MRW.

From (2.5), we have that

$$\mathbb{E}\big[B(p, 2^{-\lfloor k_N N \rfloor})\big] = \gamma(p) 2^{\lfloor k_N N \rfloor (p/2 - \psi(p/2) - 1)}$$

so that

$$T_1^N = O_P\big(2^{N(1/2 + k_N \psi(p/2) - \psi(p)/2)}(2^{-\psi(p/2)} - 1)\big).$$

The statistic $(T_1^N)^2$ will therefore be large if $1/2 + k_N\psi(p/2) \gg \psi(p)/2$ and $\psi(p/2) \gg 0$. In the case of the log-normal MRW, we have $\psi(p) = \lambda^2 p(p-1)/2$. Hence, supposing that $k_N \approx 1$, we have

$$T_1^N = O_P\big(\lambda^2 2^{N(1/2 - \lambda^2 p^2/8)}\big) \quad \text{for small } \lambda^2.$$

Therefore, the value of $(T_1^N)^2$ may be small if either λ^2 is too small (the MRW process is "close" to a Brownian motion) or too large (indeed, Lemma 1 doesn't hold for low λ^2 or large p, and $B(2p, 2^{-N})$ becomes degenerate in such cases).

Tables 1 and 2 show the number of simulations for which H_0: X = Itō is rejected (out of 100 simulations of each process), that is the proportion of simulated sample paths for which the statistic T_1^N is above $z_{1-\alpha}$, where $z_{1-\alpha}$ is the $1-\alpha$ quantile of a $\chi^2(1)$ distribution and $\alpha = 10\%$, 5% or 1% is the asymptotic

Table 1

Number of rejections of H_0: X = Itō for 100 simulations of an Itō semi-martingale ($p = 3$, $k_N = 1$)

Simulated process	Itō with no jumps		Itō with jumps	
Number n of data	32 768	1 048 576	32 768	1 048 576
Level of the test				
10%	11	11	6	0
5%	3	5	2	0
1%	2	2	0	0

Table 2

Number of rejections of H_0: X = Itō for 100 simulations of a log-normal MRW ($p = 3$, $k_N = 1$)

Simulated process	MRW, $\lambda^2 = 0.02$		MRW, $\lambda^2 = 0.1$		MRW, $\lambda^2 = 0.7$	
Number n of data	32 768	1 048 576	32 768	1 048 576	32 768	1 048 576
Level of the test						
10%	12	13	15	66	7	7
5%	6	6	8	58	3	3
1%	1	2	1	30	0	1

level of the test. These simulations were obtained with $p = 3$ and $k_N = 1$. We see that for $\lambda^2 = 0.7$, the test statistic is very close to zero – indeed, Theorem 1 does not hold in this case. Also, we see that for the number of data we considered, our test statistic does not allow to recognize a log-normal MRW process from a Brownian motion if the value of λ^2 is too small. However, for a more reasonable value of λ^2 in the range of what can be estimated from financial data [5], we find that our test performs reasonably well, provided that the number of data is large enough.

4.3. Case H_0: X = MRW, $\tau(4)$ known

Tables 3 and 4 present the test results of H_0: X = MRW in the case where τ_4 is known. If Itō semi-martingales are simulated, we consider two configurations: either $\tau_4 = 0.9$ (that is, $\lambda^2 = 0.1$ in the case of a log-normal MRW), or $\tau_4 = 0.3$ ($\lambda^2 = 0.7$). Let us recall that if X is an Itō semi-martingale, then in restriction to the set Ω^c, the order of magnitude of T_2^N is $2^{(N-\lfloor kN\rfloor)(\tau_4-1)/2}$ and restriction to the set Ω^j, this order is $2^{(N-\lfloor kN\rfloor)(\tau_4-1)/2}2^{-\lfloor kN\rfloor/2}$, which yields a first approximation for the power of the test in this case.

When MRW's are simulated, the rejection rates are close to the theoretical rates 10%, 5%, 1%. This is in agreement with the Gaussian fit we obtain for T_2^N (see Figure 4). When Itō semi-martingales are simulated, we note that the test is not very powerful for low n: the probability of correctly rejecting H_0 is rather low. However, this probability becomes quite high for $N \geq 20$, especially in the case of a Brownian motion with Poissonian jumps.

Table 3
Number of rejections of H_0: X = MRW, $\tau(4)$ known, for 100 simulations of an Itō semi-martingale ($k = 1/2$)

Simulated process	Itō with no jumps				Itō with jumps			
Number n of data	32 768		1 048 576		32 768		1 048 576	
Value of τ_4	0.9	0.3	0.9	0.3	0.9	0.3	0.9	0.3
Level of the test								
10%	19	31	24	62	57	68	66	100
5%	4	16	5	26	19	34	40	89
1%	1	5	0	4	1	6	7	29

Table 4
Number of rejections of H_0: X = MRW, $\tau(4)$ known, for 100 simulations of a log-normal MRW ($k = 1/2$)

Simulated process	MRW, $\lambda^2 = 0.02$		MRW, $\lambda^2 = 0.1$		MRW, $\lambda^2 = 0.7$	
Number n of data	32 768	1 048 576	32 768	1 048 576	32 768	1 048 576
Level of the test						
10%	10	12	11	11	10	9
5%	5	6	8	5	3	3
1%	1	0	1	2	2	1

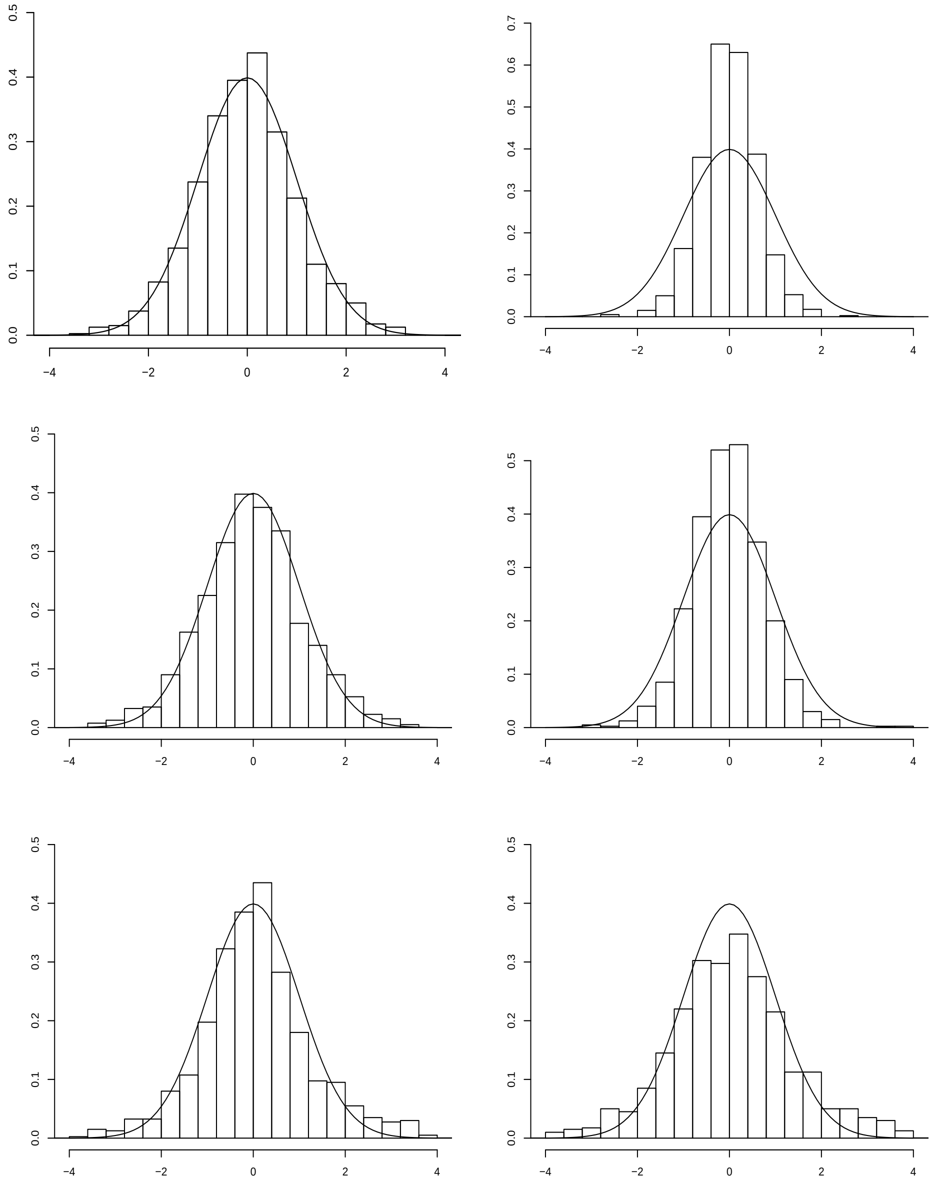

FIG 4. *Empirical distribution of T_2^N and $\widetilde{T}_2^N$ when MRW's are simulated, and fit with a standard Gaussian distribution ($n = 2^N$, $N = 15$, $k = 1/2$, $v_N = 1/\sqrt{N}$, $u_N = 1/N$). Top: $\lambda^2 = 0.02$, middle: $\lambda^2 = 0.1$, and bottom: $\lambda^2 = 0.7$. Left: T_2^N (case $\tau(4)$ known), right: $\widetilde{T}_2^N$ (case $\tau(4)$ unknown).*

Table 5
Number of rejections of H_0: X = MRW, $\tau(4)$ unknown, for 100 simulations of an Itō semi-martingale ($k = 1/2$, $v_N = 1/\sqrt{N}$, $u_N = 1/N$)

Simulated process	Itō with no jumps		Itō with jumps	
Number n of data	32 768	1 048 576	32 768	1 048 576
Level of the test				
10%	15	23	67	100
5%	6	13	35	90
1%	2	2	8	34

Table 6
Number of rejections of H_0: X = MRW, $\tau(4)$ unknown, for 100 simulations of a log-normal MRW ($k = 1/2$, $v_N = 1/\sqrt{N}$, $u_N = 1/N$)

Simulated process	MRW, $\lambda^2 = 0.02$		MRW, $\lambda^2 = 0.1$		MRW, $\lambda^2 = 0.7$	
Number n of data	32 768	1 048 576	32 768	1 048 576	32 768	1 048 576
Level of the test						
10%	17	20	11	18	8	5
5%	8	12	6	9	5	2
1%	1	2	1	2	1	2

4.4. Case H_0: X = MRW, $\tau(4)$ unknown

Next, we consider in Tables 5 and 6 the case where $\tau(4)$ is unknown. The results we find are very similar to the previous case. However, one can see that the Gaussian fit we obtain when MRW processes are simulated is somewhat less exact than in the case where $\tau(4)$ is known: the estimation of the variance of the Gaussian limit is less accurate. Hence, we find that the rejection rates are less close to the theoretical ones in this case.

In particular, this fit seems to be slightly worse for large $n = 2^{20}$ than for $n = 2^{15}$, which might appear as surprising. This comes from the fact that the estimator $\tau^*(4)$ achieves a very slow convergence rate on our simulations, so that the estimation

$$\frac{\sqrt{3}}{\sqrt{2(2^{\tau^*(4)} - 1)}} 2^{(N - \lfloor kN \rfloor)\tau^*(4)/2}$$

of the variance used in the statistic $\widetilde{T}_2^N$ is actually less accurate for $N = 20$ than for $N = 15$.

Finally, the rejection rates we obtain for Itō semi-martingales simulations are still satisfactory, especially for processes with jumps.

5. Proofs

5.1. Proof of Proposition 1

We in fact prove a slightly more general result, Proposition 3, from which Proposition 1 is easily deduced. We consider an MRW of the form $X_t = B_{\theta_t}$, such that $p_{max} > 4$, where B is a Brownian motion with respect to some filtration $\mathcal{F}'$. We fix here a path of θ (we will use afterwards the independence between B and θ). Note that this also defines $\theta_t^{(4)}$ in Lemma 1. We set $\mathcal{G}_t = \mathcal{F}'_{\theta_t}$ and define the $\mathcal{G}_{2i/n}$-measurable random vector $\xi_i^n = (\xi_i^{n,1}, \xi_i^{n,2})$ by

$$\begin{aligned}\xi_i^{n,1} &= n^{\tau(4)/2}\{(X_{2i/n} - X_{(2i-1)/n})^2 + (X_{(2i-1)/n} - X_{(2i-2)/n})^2 \\ &\quad - (\theta_{2i/n} - \theta_{(2i-2)/n})\} \\ \xi_i^{n,2} &= n^{\tau(4)/2}\{(X_{2i/n} - X_{(2i-2)/n})^2 - (\theta_{2i/n} - \theta_{(2i-2)/n})\}.\end{aligned}$$

For $t \in [0,1]$, let C_t be the 2×2 matrix defined by

$$C_t = 2\theta_t^{(4)} \begin{pmatrix} 1 & 1 \\ 1 & 2^{\tau(4)} \end{pmatrix}$$

and defined the process $M^n = \{M_t^n, t \in [0,1]\}$ by

$$M_t^n = \sum_{i=1}^{\lfloor 2^N t/2 \rfloor} \xi_i^n.$$

We have the following result.

Proposition 3. *If $p_{max} > 4$, then for a given path θ, the process M^n converges in law towards a continuous centered $\mathbb{R}^2$-valued Gaussian process Z, with independent increments such that $\mathbb{E}\Big[Z_t^j Z_t^k\Big] = C_t^{jk}$. This entirely characterizes the law of the limiting process.*

Remark 4: In Proposition 3, we retrieve in particular the result obtained by Ludeña [23] in the one dimensional case.

For the proof of Proposition 3, we will consider the four following lemmas:

Lemma 2. *We have*

$$\mathbb{E}_{\mathcal{G}_{(2i-2)/n}}[(\xi_i^{n,1})] = 0,\ \mathbb{E}_{\mathcal{G}_{(2i-2)/n}}[(\xi_i^{n,2})] = 0.$$

Lemma 3. *We have*

$$\sum_{i=1}^{\lfloor 2^N t/2 \rfloor} \mathbb{E}_{\mathcal{G}_{(2i-2)/n}}[(\xi_i^{n,1})^2] \xrightarrow{\mathbb{P}} 2\theta_t^{(4)}, \quad \sum_{i=1}^{\lfloor 2^N t/2 \rfloor} \mathbb{E}_{\mathcal{G}_{(2i-2)/n}}[(\xi_i^{n,2})^2] \xrightarrow{\mathbb{P}} 2^{1+\tau(4)}\theta_t^{(4)}.$$

Lemma 4. *We have*

$$\sum_{i=1}^{\lfloor 2^N t/2 \rfloor} \mathbb{E}_{\mathcal{G}_{(2i-2)/n}}[\xi_i^{n,1}\xi_i^{n,2}] \xrightarrow{\mathbb{P}} 2\theta_t^{(4)}.$$

Lemma 5. *For some $\varepsilon > 0$*

$$\sum_{i=1}^{n/2} \mathbb{E}_{\mathcal{G}_{(2i-2)/n}}[(\xi_i^{n,1})^{2+\varepsilon}] \xrightarrow{\mathbb{P}} 0, \quad \sum_{i=1}^{n/2} \mathbb{E}_{\mathcal{G}_{(2i-2)/n}}[(\xi_i^{n,2})^{2+\varepsilon}] \xrightarrow{\mathbb{P}} 0.$$

Since the time change θ_t is fixed, $\theta_t^{(4)}$ is deterministic. Thus, Proposition 3 follows from Lemmas 2–5 using a standard convergence result for triangular arrays of semi-martingales, see for example [20]. We now turn to the proofs of Lemmas 2–5.

Proof of Lemma 2

The result comes directly from the fact that

$$\begin{aligned}
&\mathbb{E}_{\mathcal{G}_{(2i-2)/n}}[(X_{2i/n} - X_{(2i-1)/n})^2] = \theta_{2i/n} - \theta_{(2i-1)/n}, \\
&\mathbb{E}_{\mathcal{G}_{(2i-2)/n}}[(X_{(2i-1)/n} - X_{(2i-2)/n})^2] = \theta_{(2i-1)/n} - \theta_{(2i-2)/n}, \\
&\mathbb{E}_{\mathcal{G}_{(2i-2)/n}}[(X_{2i/n} - X_{(2i-2)/n})^2] = \theta_{2i/n} - \theta_{(2i-2)/n}.
\end{aligned}$$

Proof of Lemma 3

For simplicity, we just give the proof for $\xi^{n,2}$, the result for $\xi^{n,1}$ being obviously deduced. We easily get

$$\mathbb{E}_{\mathcal{G}_{(2i-2)/n}}[(\xi_i^{n,2})^2] = 2^{1+\tau(4)}(n/2)^{\tau(4)}(\theta_{2i/n} - \theta_{(2i-2)/n})^2.$$

We conclude using Lemma 1.

Proof of Lemma 4

Conditional on $\mathcal{G}_{(2i-2)/n}$, $(X_{2i/n} - X_{(2i-1)/n}, X_{2i/n} - X_{(2i-2)/n})$ is a centered Gaussian vector with variance-covariance matrix equal to

$$\begin{pmatrix} \theta_{2i/n} - \theta_{(2i-1)/n} & \theta_{2i/n} - \theta_{(2i-1)/n} \\ \theta_{2i/n} - \theta_{(2i-1)/n} & \theta_{2i/n} - \theta_{(2i-2)/n} \end{pmatrix}.$$

Thus, it has the same law as

$$\big((\theta_{2i/n} - \theta_{(2i-1)/n})^{1/2} Z_1, (\theta_{2i/n} - \theta_{(2i-2)/n})^{1/2} Z_2\big),$$

where (Z_1, Z_2) is a centered Gaussian vector with variance-covariance matrix equal to

$$\begin{pmatrix} 1 & \big(\frac{\theta_{2i/n} - \theta_{(2i-1)/n}}{\theta_{2i/n} - \theta_{(2i-2)/n}}\big)^{1/2} \\ \big(\frac{\theta_{2i/n} - \theta_{(2i-1)/n}}{\theta_{2i/n} - \theta_{(2i-2)/n}}\big)^{1/2} & 1 \end{pmatrix}.$$

Then, note that

$$\mathbb{E}_{\mathcal{G}_{(2i-2)/n}}\big[\big\{(X_{2i/n}-X_{(2i-1)/n})^2-(\theta_{2i/n}-\theta_{(2i-1)/n})\big\} \\ \times\big\{(X_{2i/n}-X_{(2i-2)/n})^2-(\theta_{2i/n}-\theta_{(2i-2)/n})\big\}\big]$$

is equal to

$$(\theta_{2i/n}-\theta_{(2i-1)/n})(\theta_{2i/n}-\theta_{(2i-2)/n})\mathbb{E}_{\mathcal{G}_{(2i-2)/n}}[(Z_1^2-1)(Z_2^2-1)].$$

Using Mehler's formula, the preceding conditional expectation is finally equal to

$$2(\theta_{2i/n}-\theta_{(2i-1)/n})^2.$$

In the same way, we get

$$n^{-\tau(4)/2}\mathbb{E}_{\mathcal{G}_{(2i-2)/n}}[\xi_i^{n,2}\{(X_{(2i-1)/n}-X_{(2i-2)/n})^2-(\theta_{(2i-1)/n}-\theta_{(2i-2)/n})\}] \\ =2(\theta_{(2i-1)/n}-\theta_{(2i-2)/n})^2.$$

Eventually, we obtain

$$\mathbb{E}_{\mathcal{G}_{(2i-2)/n}}[\xi_i^{n,1}\xi_i^{n,2}]=2n^{\tau(4)}\{(\theta_{2i/n}-\theta_{(2i-1)/n})^2+(\theta_{(2i-2)/n}-\theta_{(2i-2)/n})^2\}.$$

We conclude using Lemma 1.

Proof of Lemma 5

Here again, we just give the proof for $\xi^{n,2}$. It is clear that conditional on $\mathcal{G}_{(2i-2)/n}$ the law of $(\xi_i^{n,2})$ is the same as the law of $n^{\tau(4)/2}(\theta_{2i/n}-\theta_{(2i-2)/n})(Z^2-1)$, with Z a standard centered Gaussian variable. Since we are in the case $p_{max}>4$, some basic concavity consideration and the fact that $\tau(\cdot)$ is necessarily a right-continuous function show that we can choose some $\varepsilon>0$ such that $4<4(1+\varepsilon)<p_{max}$ and $\tau(4(1+\varepsilon))>(1+\varepsilon)\tau(4)$. Thus,

$$\mathbb{E}_{\mathcal{G}_{(2i-2)/n}}[(\xi_i^{n,2})^{2(1+\varepsilon)}]\le cn^{(1+\varepsilon)\tau(4)}(\theta_{2i/n}-\theta_{(2i-2)/n})^{2(1+\varepsilon)} \\ \le cn^{(1+\varepsilon)\tau(4)-\tau(4(1+\varepsilon))}n^{\tau(4(1+\varepsilon))}(\theta_{2i/n}-\theta_{(2i-2)/n})^{2(1+\varepsilon)}.$$

We conclude using Lemma 1.

5.2. Proof of Proposition 2

Assume first that X is an MRW. We have

$$N(\hat{\tau}(4)-\tau(4))=\frac{2}{\log(2)}\big\{\log\big(2^{N\tau(4)/2}B(4,2^{-\lfloor N/2\rfloor})\big)-\log\big(2^{N\tau(4)}B(4,2^{-N})\big)\big\}.$$

Using Lemma 1, we get that

$$\frac{2^{\lfloor N/2\rfloor\tau(4)}B(4,2^{-\lfloor N/2\rfloor})}{2^{N\tau(4)}B(4,2^{-N})}\to 1.$$

Since $B(4, 2^{-\lfloor N/2 \rfloor})$ tends to zero, we can replace $\lfloor N/2 \rfloor$ by $N/2$ and so we obtain that $N(\hat{\tau}(4) - \tau(4))$ tends to 0 almost surely. Since $0 < \tau(4) < 1$, the first assertion of Proposition 2 follows.

We now turn to the second assertion. In restriction to Ω^c, we get the result using that $\tau^*(4)$ tends to 1 together with the inequality

$$2^{N(\tau^*(4)-1)} \leq 2^{-Nv_N} + 2^{N(u_N - 1)}.$$

In restriction to Ω^j, we use the inequality

$$2^{N(\tau^*(4)-1)} \leq 2^{N(\hat{\tau}(4)-1)} + 2^{N(u_N-1)},$$

and the facts that $N\hat{\tau}(4)$ goes to zero and that $2^{\tau^*(4)} - 1$ is of the same order than $\tau^*(4)$ which is greater than u_N.

Acknowledgements

We are very grateful to the anonymous referee and to the associate editor whose relevant comments have significantly improved a former version of the paper.

References

[1] Aït-Sahalia, Y. and J. Jacod (2009). Testing for jumps in a discretely observed process. *The Annals of Statistics*, **37** 184–222. MR2488349

[2] Aït-Sahalia, Y. and J. Jacod (2009). Is Brownian motion necessary to model high frequency data? *The Annals of Statistics*, forthcoming.

[3] Aït-Sahalia, Y. and J. Jacod (2008). Testing whether jumps have finite or infinite activity. Working paper, Princeton University and Université Paris VI.

[4] Bacry, E., Kozhemyak, A. and J.F. Muzy (2008). Continuous cascade models for asset returns. *Journal of Economic Dynamics and Control,* **32** 156–199. MR2381693

[5] Bacry, E., Kozhemyak, A. and J.F. Muzy (2008). Log-Normal continuous cascades: aggregation properties and estimation. Application to financial time-series. *Quantitative Finance*, to appear. MR2381693

[6] Bacry, E. and J.F. Muzy (2003). Log-infinitely divisible multifractal process. *Comm. in Math. Phys.,* **236** 449–475. MR2021198

[7] Barndorff-Nielsen, O.E. and J. Schmiegel (2008). Time change, volatility and turbulence. In *Proc. of the Workshop on Math. Control Theory and Finance 2007, Springer* 29–53. MR2484103

[8] Barral, J. and B.B. Mandelbrot (2002). Random multiplicative multifractal measures, part I II and III. *Proceedings of the Symposium on Pure Mathematics*, **72** 1–90.

[9] Bouchaud, J.P. and M. Potters (2003). *Theory of financial risks and derivative pricing.* Second Edition, Cambridge University Press. MR1787145

[10] Calvet, L., and A. Fisher (2008). *Multifractal volatility: theory, forecasting and pricing*. Academic Press Advanced Finance Series.

[11] Delbaen, F. and W. Schachermayer (1994). A general version of the fundamental theorem of asset pricing. *Mathematische Annalen*, **300** 463–520. MR1304434

[12] Duchon, J., Robert, R. and V. Vargas (2010). Forecasting volatility with the multifractal random walk model. To appear in *Mathematical Finance*. MR2642887

[13] Duvernet, L. (2009). Convergence of the structure function of a multifractal random walk in a mixed asymptotic setting. *Stochastic Analysis and Applications*. **28**(5) 763–792.

[14] Frisch, U. (1995). *Turbulence: the legacy of A. N. Kolmogorov*. Cambridge University Press. MR1428905

[15] Frisch, U. and G. Parisi (1985). Fully developed turbulence and intermittency. *Proc. of Int. Summer school Phys. Enrico Fermi.*

[16] Gagne, Y., Marchand, M., and B. Castaing (1994). Conditional velocity pdf in 3D turbulence. *Journal de Physique II*,**4** 1–8.

[17] Ghashghaie, S., Breymann, W., Peinke, J., Talkner, P., and Y. Dodge (1996). Turbulence cascades in foreign exchange markets. *Nature,* **381** 767–770.

[18] Gloter, A., and M. Hoffmann (2010). Nonparametric reconstruction of a multifractal function from noisy data. *Probability Theory and Related Fields,* **146** 155–187. MR2550361

[19] Jacod, J. (2008). Asymptotic properties of realized power variations and related functionals of semi-martingales. *Stochastic Processes and their Applications,* **118** 517–559. MR2394762

[20] Jacod, J., and A.N. Shiryaev (2003). *Limit Theorems for Stochastic Processes*. Second Edition, Springer-Verlag. MR1943877

[21] Kahane, J. P. (1985). Sur le chaos multiplicatif. *Ann. Sci. Math. Quebec* **9**(2) 105–150. MR0829798

[22] Ludeña, C. (2008). L^p-variations for multifractal fractional random walks. *Annals of Applied Probability,* **18** 1138–1163. MR2418240

[23] Ludeña, C. (2009). Confidence intervals for the scaling function of multifractal random walks. *Statistics and Probability Letters,* **79** 1186–1193. MR2519001

[24] Mandelbrot, B.B. (1997). *Fractals and scaling in finance*. Springer, New York. MR1475217

[25] Mandelbrot, B.B. and H.M. Taylor (1967). On the distribution of stock price differences. *Operations Research*, **15** 1057-1062.

[26] Ossiander, M. and E.C. Waymire (2000). Statistical estimation for multiplicative cascades. *The Annals of Statistics*, **28** 1533–1560. MR1835030

[27] Parks, K., and W. Willinger (1999). Self-similar network traffic: an overview. In Parks, K., and W. Willinger (eds) *Self-Similar Network Traffic and Performance Evaluation*, Wiley.

[28] Rosenbaum, M. (2007). A new microstructure noise index. *Quantitative Finance*, to appear.

[29] Wendt, H., and P. Abry (2007). Multifractality Tests using Bootstrapped Wavelet Leaders. *IEEE Transactions on Signal Processing* **55**(10) 4811-4820.

[30] Wendt, H., P. Abry and S. Jaffard (2007). Bootstrap for Empirical Multifractal Analysis with Application to Hydrodynamic Turbulence. *IEEE Signal Processing Magazine*, **24**(4) 38–48.

[31] Yu, Z.G., Anh, V., and K. S. Lau (2001). Multifractal characterization of length sequences of coding and noncoding segments in a complete genom. *Physica A*, **301** 351–361.