Nonparametric tests for pathwise properties of semimartingales

Rama CONT* & Cecilia MANCINI†

Jan 2010‡

Abstract

We propose two nonparametric tests for investigating the pathwise properties of a signal modeled as the sum of a Lévy process and a Brownian semimartingale. Using a nonparametric threshold estimator for the continuous component of the quadratic variation, we design a test for the presence of a continuous martingale component in the process and a test for establishing whether the jumps have finite or infinite variation, based on observations on a discrete time grid. We evaluate the performance of our tests using simulations of various stochastic models and use the tests to investigate the fine structure of the DM/USD exchange rate fluctuations and SPX futures prices. In both cases, our tests reveal the presence of a non-zero Brownian component and a finite variation jump component.

Continuous-time stochastic models based on *discontinuous semimartingales* have been increasingly used in many applications, such as financial econometrics, option pricing and stochastic control. Some of these models are constructed by adding IID jumps to a continuous process driven by Brownian motion [22, 16], while others are based on purely discontinuous processes which move only through jumps [18, 8]. Even within the class of purely discontinuous models, one finds a variety of models with different path properties– finite/infinite jump intensity, finite/infinite variation– which turn to have an importance in applications such as optimal stopping [5] and the asymptotic behavior of option prices [9, 10]. It is therefore of interest to investigate which class of models – diffusion, jump-diffusion or pure jump– is the most appropriate for a given data set. Nonparametric procedures have been recently proposed for investigating the presence of jumps [6, 2, 17] and studying some fine properties of the jumps [3, 4, 25, 26] in a signal. We address here related, but different, issues: for a semimartingale whose jump component is a Lévy process, we propose a test for the presence of a continuous martingale component in the price process, which allows to discriminate between pure-jump and jump-diffusion models, and a test for determining whether the jump component has finite or infinite variation. Our tests are based on a nonparametric threshold estimator [20] for the integrated variance -defined as the continuous component of the quadratic variation- based on observations on a discrete time grid. Without imposing restrictive assumptions on the continuous martingale component, we obtain a central limit theorem for this threshold estimator (Section 2) and use it to design our tests (Section 3).

Using simulations of stochastic models commonly used in finance, we check the performance of our tests for realistic sample sizes (section 4). Applied to time series of the DM/USD exchange rate and SPX futures prices (section 5), our tests reveal in both cases the presence of a non-zero Brownian component, combined with a finite variation jump component. These results suggest that these asset prices may be modeled as the sum of a Brownian martingale and a jump component of finite variation.

*IEOR Dept., Columbia University, New York and Laboratoire de Probabilités et Modèles Aléatoires, CNRS-Université Paris VI, France. Email: `Rama.Cont@columbia.edu`.

†Dipartimento di Matematica per le Decisioni, Università di Firenze, Italy. Email: `cecilia.mancini@dmd.unifi.it`.

‡A previous version of this working paper appeared as: "Nonparametric test for analyzing the fine structure of price fluctuations", Columbia Financial Engineering Report 2007-13. This work has benefited from support by the European Science Foundation program *Advanced Mathematical methods in Finance*, by Istituto Nazionale di Alta Matematica and by MIUR grants n.206132713-001 and n.2004011204-002. We thank Jean Jacod and Suzanne Lee for important comments.

1 Definitions and notations

We consider a semimartingale $(X_t)_{t\in[0,T]}$, defined on a (filtered) probability space $(\Omega, (\mathcal{F}_t)_{t\in[0,T]}, \mathcal{F},\ P)$ with paths in $D([0,T],\mathbb{R})$, driven by a (standard) Brownian motion W and a pure jump Lévy process L:

$$X_t = x_0 + \int_0^t a_s ds + \int_0^t \sigma_s\ dW_s + L_t,\ t \in]0,T], \tag{1}$$

where a, σ are adapted processes with right continuous paths with left limits (càdlàg processes) such that (1) admits a unique strong solution X on $[0,T]$ which is adapted and càdlàg [11]. L has Lévy measure ν and may be decomposed as $L_t = J_t + M_t$, where

$$J_t := \int_0^t \int_{|x|>1} x\mu(dx\ ds) = \sum_{\ell=1}^{N_t} \gamma_\ell, \quad M_t := \int_0^t \int_{|x|\le 1} x[\mu(dx\ ds) - \nu(dx)dt]. \tag{2}$$

J is a compound Poisson process representing the "large" jumps of X, μ is a Poisson random measure on $[0,T]\times\mathbb{R}$ with intensity measure $\nu(dx)dt$, N is a Poisson process with intensity $\nu(\{x, |x|>1\}) < \infty$, γ_ℓ are IID and independent of N and the martingale M is the compensated sum of small jumps of L. We will denote $\mu(dx,dt) - \nu(dx)dt =: \tilde{\mu}(dx,dt)$ the compensated Poisson random measure associated to μ. We allow for the *infinite activity* (IA) case $\nu(\mathbb{R}) = \infty$, where small jumps of L occur infinitely often. For a semimartingale Z we denote $\Delta_i Z = Z_{t_i} - Z_{t_{i-1}}$ its increments and $\Delta Z_t = Z_t - Z_{t-}$ its jump at time t. The *Blumenthal-Getoor (BG) index* of L, defined as

$$\alpha := \inf\{\delta \ge 0,\ \int_{|x|\le 1} |x|^\delta \nu(dx) < +\infty\} \le 2,$$

measures the degree of *activity* of small jumps. A compound Poisson process has $\alpha = 0$, while an α-stable process has BG index equal to $\alpha \in]0,2[$. The Gamma process and the Variance Gamma (VG) process are examples of infinite activity Lévy processes with $\alpha = 0$. A pure jump Lévy process with BG index $\alpha < 1$ has paths with *finite variation*, while for $\alpha > 1$ then the sample paths have *infinite variation* almost-surely. When $\alpha = 1$ the paths may have either finite or infinite variation [7]. The Normal Inverse Gaussian process (NIG) and the Generalized Hyperbolic Lévy motion (GHL) have infinite variation and $\alpha = 1$. Tempered stable processes [8, 10] allow for $\alpha \in [0,2[$. We call $IV = \int_0^T \sigma_u^2 du$ the *integrated variance* of X and $IQ = \int_0^T \sigma_u^4 du$ the *integrated quarticity* of X and denote

$$X_{0t} = \int_0^t a_s ds + \int_0^t \sigma_s dW_s, \qquad X_{1t} = X_{0t} + J_t$$

We will use the following assumption:

$$\textbf{Assumption A1} \qquad \exists \alpha \in [0,2]: \qquad \int_{|x|\le\varepsilon} x^2 \nu(dx) \sim \varepsilon^{2-\alpha},\ \text{as } \varepsilon \to 0, \tag{3}$$

where $f(h) \sim g(h)$ means that $f(h) = O(g(h))$ and $g(h) = O(f(h))$ as $h \to 0$. This assumption implies that α is the BG index of L. **A1** is satisfied if for instance ν has a density which behaves as $\frac{K_\pm}{|x|^{1+\alpha}}$ when $x \to 0\pm$, where $K_\pm > 0$. In particular **A1** holds for all Lévy processes commonly used in finance [10]: NIG, Variance Gamma, tempered stable processes or Generalized Hyperbolic processes.

Typically, we observe X_t in form of a discrete record $\{x_0, X_{t_1} .. X_{t_{n-1}}, X_{t_n}\}$ on a time grid $t_i = ih$ with $h = T/n$. Our goal is to provide, given such a discrete observations, nonparametric tests for

- detecting the presence of a continuous martingale component in the price process
- analyzing the qualitative nature of the jump component i.e. whether it has finite or infinite variation.

2 CLT for a threshold estimator of integrated variance

The "realized variance" $\sum_{i=1}^{n}(\Delta_i X)^2$ of the semimartingale X converges in probability [24] to

$$[X]_T := \int_0^T \sigma_t^2 dt + \int_0^T \int_{\mathbb{R}-\{0\}} x^2 \mu(dx, ds).$$

A *threshold estimator* [19, 20] of the integrated variance $IV = \int_0^T \sigma_t^2 dt$ is based on the idea of summing only some of the squared increments of X, those whose absolute value is smaller than some *threshold* r_h:

$$\hat{IV}_h := \sum_{i=1}^{n} (\Delta_i X)^2 I_{\{(\Delta_i X)^2 \le r_h\}}. \tag{4}$$

The term $\int_0^T \int_{\mathbb{R}-\{0\}} x^2 \mu(dx,\ ds)$ due to jumps vanishes as $h \to 0$ for an appropriate choice of the threshold. Paul Lévy's law for the modulus of continuity of the Brownian paths implies

$$P\left(\lim_{h\to 0} \sup_{i\in\{1..n\}} \frac{|\Delta_i W|}{\sqrt{2h \ln \frac{1}{h}}} \le 1 \right) = 1$$

and allows to choose such a threshold. It is shown in [20, Cor 2, Thm 4] that, under the above assumptions, if we choose a deterministic threshold r_h such that

$$\lim_{h\to 0} r_h = 0 \text{ and } \lim_{h\to 0} \frac{h \ln h}{r_h} = 0 \tag{5}$$

then $\hat{IV}_h \stackrel{P}{\to} IV$, as $h \to 0$. If the jumps have finite intensity the thresholding procedure allows, as $h \to 0$, to detect a jump in $]t_{i-1}, t_i]$. In fact since a and σ are càdlàg (or càglàd), their paths are a.s. bounded on $[0, T]$ so

$$\limsup_{h\to 0} \frac{\sup_i |\int_{t_{i-1}}^{t_i} a_s(\omega) ds|}{h} \le A(\omega) < \infty \text{ and } \limsup_{h\to 0} \frac{\sup_i |\int_{t_{i-1}}^{t_i} \sigma_s^2(\omega) ds|}{h} \le \Sigma(\omega) < \infty \quad a.s. \tag{6}$$

It follows from [20] that

$$a.s. \qquad \sup_i \frac{|\int_{t_{i-1}}^{t_i} a_s ds + \int_{t_{i-1}}^{t_i} \sigma_s dW_s|}{\sqrt{2h \log \frac{1}{h}}} \le A\sqrt{h} + \sqrt{\Sigma} + 1 := \Lambda. \tag{7}$$

Since realistic values of σ for asset prices belong to $[0.1, 0.8]$ (in annual units), we have that for small h the r.v. Λ has order of magnitude of 1, thus in the finite jump intensity case, a.s. for sufficiently small h, $(\Delta_i X)^2 > r_h > 2h \log \frac{1}{h}$ indicates the presence of jumps in $]t_{i-1}, t_i]$.
When L has infinite activity, $\sum_{i=1}^{n}(\Delta_i X)^2 I_{\{(\Delta_i X)^2 \le r_h\}}$ behaves like $\sum_{i=1}^{n}(\Delta_i X)^2\ I_{\{\Delta_i N = 0, |\Delta_i M| \le 2\sqrt{r_h}\}}$ for small h (lemma 6.2). Moreover for any $\delta > 0$ the jumps contributing to the increments $\Delta_i X$ such that $(\Delta_i X)^2 \le r_h$ for small h have size smaller than $c\sqrt{r_h + \delta}$ [20, Lemma 1] so their contribution vanishes when $h \to 0$. Note that $r_h = ch^\beta$ satisfy condition (5) for any $\beta \in]0, 1[$ and any constant c. Since $\sqrt{2}\sigma \simeq 1$ in most applications, we use $c = 1$. Denote

$$\eta^2(\varepsilon) := \int_{|x|\le\varepsilon} x^2 \nu(dx), \qquad d(\varepsilon) := \int_{\varepsilon<|x|\le 1} x \nu(dx) \tag{8}$$

Let us remark that if $\lim_{h\to 0} r_h = 0$ then by **A1** we have, as $h \to 0$,

$$\eta^2\left(2\sqrt{r_h}\right) = \int_{|x|\le 2\sqrt{r_h}} x^2 \nu(dx) \sim r_h^{1-\frac{\alpha}{2}}, \qquad \int_{|x|\le 2\sqrt{r_h}} x^k \nu(dx) \sim r_h^{\frac{k-\alpha}{2}},\ k = 3, 4$$
$$\int_{2\sqrt{r_h}<|x|\le 1} x\nu(dx) \sim \left[c + r_h^{\frac{1-\alpha}{2}}\right] I_{\{\alpha\ne 1\}} + \left[\ln \frac{1}{2\sqrt{r_h}}\right] I_{\{\alpha=1\}}, \qquad \int_{2\sqrt{r_h}<|x|\le 1} \nu(dx) \sim r_h^{-\alpha/2}, \tag{9}$$

where α is the BG index of L. The following lemma, proved in the appendix, states that under (5), each increment $\Delta_i M$ such that $|\Delta_i M| \le 2\sqrt{r_h}$ only contains jumps of magnitude less than $2\sqrt{r_h}$ if $\alpha \le 1$, or smaller than $2h^{\frac{1}{2\alpha}} \log^{\frac{1}{2\alpha}} \frac{1}{h}$ if $\alpha > 1$.

Lemma 2.1. *Define, for $h > 0$, $v_h := h^{\frac{1}{2\alpha}} \log^{\frac{1}{2\alpha}} \frac{1}{h}$. Under (5) there exists a sequence $h_k = T/n_k$ tending to zero as $k \to \infty$ such that, for k_0 sufficiently large and $h \in \{h_k, k \geq k_0\}$,*
i) if $\alpha \leq 1$, then for all $i = 1..n$

$$\Delta_i M \; I_{\{(\Delta_i M)^2 \leq 4r_h\}} = \Big(\int_{t_{i-1}}^{t_i} \int_{|x| \leq 2\sqrt{r_h}} x\tilde{\mu}(dx, dt) - \int_{t_{i-1}}^{t_i} \int_{2\sqrt{r_h} < |x| \leq 1} x\nu(dx)dt \Big) I_{\{(\Delta_i M)^2 \leq 4r_h\}} \quad a.s.$$

ii) if $\alpha > 1$, then for all $i = 1..n$ we have

$$\Delta_i M \; I_{\{(\Delta_i M)^2 \leq 4r_h\}} = \Big(\int_{t_{i-1}}^{t_i} \int_{|x| \leq 2v_h} x\tilde{\mu}(dx, dt) - \int_{t_{i-1}}^{t_i} \int_{2v_h < |x| \leq 1} x\nu(dx)dt \Big) I_{\{(\Delta_i M)^2 \leq 4r_h\}} \quad a.s.$$

Remark 2.2. *Note that $v_h \leq r_h^{1/4}$ so that in the case ii) above ($\alpha > 1$), for all $i = 1..n$ the jumps of M on $\{(\Delta_i M)^2 \leq 4r_h\}$ are bounded by $r_h^{1/4}$.*

Definition. Denote

$$L_t^{(h)} := \int_0^t \int_{|x| \leq \; 2\sqrt[4]{r_h}} x \; \tilde{\mu}(dx, dt) - \int_0^t \int_{2\sqrt[4]{r_h} < |x| \leq 1} x\nu(dx)dt, \quad \Delta_i M^{(h)} := \int_{t_{i-1}}^{t_i} \int_{|x| \leq 2\sqrt[4]{r_h}} x\tilde{\mu}(dx, dt), \tag{10}$$

By lemma 2.1, on a subsequence a.s. for sufficiently small h, $\forall i = 1..n$, on $\{(\Delta_i M)^2 \leq 4r_h\}$ we have

$$\Delta_i M = \Delta_i L^{(h)} = \Delta_i M^{(h)} - hd(2\sqrt[4]{r_h}) : \tag{11}$$

$\Delta_i M^{(h)}$ is the compensated sum of jumps smaller in absolute value than $2\sqrt[4]{r_h}$, while $hd(2\sqrt[4]{r_h})$ is the compensator of the (missing) jumps larger than $2\sqrt[4]{r_h}$.

In [20] a CLT for $I\hat{V}_h$ was shown in the case of finite intensity jumps and càdlàg adapted σ. Theorem 2.5 extends this to the case of infinite activity without extra assumptions on σ. In particular when $\alpha < 1$ the error $I\hat{V}_h - IV$ has the same rate of convergence and asymptotic variance as in the case of finite intensity jumps. The following proposition gives the asymptotic variance of $(I\hat{V}_h - IV)/\sqrt{2h}$ when $\alpha < 1$.

Proposition 2.3. *If $r_h = h^\beta$ with $1 > \beta > \frac{1}{2-\alpha/2} \in [1/2, 1[$ then as $h \to 0$*

$$I\hat{Q}_h := \frac{\sum_i (\Delta_i X)^4 I_{\{(\Delta_i X)^2 \leq r_h\}}}{3h} \stackrel{P}{\to} IQ = \int_0^T \sigma_t^4 dt.$$

The following result will be used to prove Theorem 2.5:

Theorem 2.4. *Under assumption **A1**, as $h \to 0$*

$$\frac{\sum\limits_{i=1}^{n} \Big(\int_{t_{i-1}}^{t_i} \int_{|x| \leq \varepsilon} x\tilde{\mu}(dx, dt) - \int_{t_{i-1}}^{t_i} \int_{|x| \in]\varepsilon, 1]} x\nu(dx)dt \Big)^2 - T\ell_{2,h}\varepsilon^{2-\alpha} - T\ell_{1,h}^2 h\varepsilon^{2-2\alpha} I_{\{\alpha \neq 1\}}}{\sqrt{T}\sqrt{\ell_{4,h}}\; \varepsilon^{2-\frac{\alpha}{2}}} \stackrel{d}{\to} N(0, 1) \tag{12}$$

where $\varepsilon = h^u$, $0 < u \leq 1/2$, $\ell_{j,h} = \int_{|x| \leq \varepsilon} x^j \nu(dx)/\varepsilon^{j-\alpha}$ for $j = 2, 4$, and $\ell_{1,h} = \int_{\varepsilon < |x| \leq 1} x\nu(dx)/\Big[(c + \varepsilon^{1-\alpha}) I_{\{\alpha \neq 1\}} + \ln \frac{1}{2\varepsilon} I_{\{\alpha = 1\}}\Big]$ tend to non-zero constants depending on ν.

We are now ready to state our central limit theorem for the estimator $I\hat{V}_h$. A sequence (X_n) is said to converge stably in law to a random variable X (defined on an extension $(\Omega', \mathcal{F}', P')$ of the original probability space) if $\lim E[Uf(X_n)] = E'[Uf(X)]$ for every bounded continuous function $f : \mathbb{R} \to \mathbb{R}$ and all bounded random variables U. This is obviously stronger than convergence in law [15].

Theorem 2.5. *Assume* **A1** *and $\sigma \not\equiv 0$; choose $r_h = h^\beta$ with $\beta > \frac{1}{2-\cancel{\alpha/2}} \in [1/2, 1[$. Then*
a) if $\alpha < 1$ we have, with $\stackrel{st}{\to}$ denoting stable convergence in law,

$$\frac{I\hat{V}_h - IV}{\sqrt{2h I\hat{Q}_h}} \stackrel{st}{\to} N(0, 1) \, ; \tag{13}$$

$$b)\quad \text{if} \quad \alpha \geq 1, \qquad \frac{\hat{IV}_h - IV}{\sqrt{2h\hat{IQ}_h}} \overset{a.s.}{\to} +\infty.$$

Remark For $\alpha < 1$ Jacod [13, Thm 2.10, i)], has shown a related central limit result for the threshold estimator of IV where L is a semimartingale but under the additional assumption that σ is an Ito semimartingale. The proof of Theorem 2.5 in the case $\alpha < 1$ does not rely on Theorem 2.10 i) of [13]. An alternative proof under the Ito semimartingale assumption for σ could combine the results [20] with [13, Thm 2.10, i)], in that

$$\frac{\hat{IV} - IV}{\sqrt{h}} = \frac{\hat{IV}(X_1) - IV}{\sqrt{h}} + \frac{\hat{IV}(M)}{\sqrt{h}} + \frac{\sum_{i=1}^n (\Delta_i X_1)^2 \big(I_{\{(\Delta_i X)^2 \leq r_h\}} - I_{\{(\Delta_i X_1)^2 \leq r_h\}}\big)}{\sqrt{h}} +$$

$$\frac{\sum_{i=1}^n (\Delta_i M)^2 \big(I_{\{(\Delta_i X)^2 \leq r_h\}} - I_{\{(\Delta_i M)^2 \leq r_h\}}\big)}{\sqrt{h}} + 2\frac{\sum_{i=1}^n \Delta_i X_1 \Delta_i M I_{\{(\Delta_i X)^2 \leq r_h\}}}{\sqrt{h}},$$

$$\text{where} \quad \hat{IV}(X_1) \doteq \sum_{i=1}^n (\Delta_i X_1)^2 I_{\{(\Delta_i X_1)^2 \leq r_h\}}, \quad \hat{IV}(M) \doteq \sum_{i=1}^n (\Delta_i M)^2 I_{\{(\Delta_i M)^2 \leq r_h\}}.$$

The first term converges stably in law by [20], the second one converges stably to zero by theorem 2.10 i) of [13]. That the remaining terms are negligible needs some work (see proof of Theorem 2.5).

3 Statistical tests

3.1 Test for the presence of a continuous martingale component

We now use the above results to design a test to detect the presence of a continuous martingale component $\int_0^t \sigma_t dW_t$ given discretely recorded observations. Our test is feasible in the case when L has BG index $\alpha < 1$ i.e. the jumps are of finite variation (see Section 3.2). The test proceeds as follows. First, we choose a coefficient $\beta \in [1/2, 1[$ close to 1. If we have an estimate $\hat{\alpha}$ of the BG index [26, 3, 25] we may choose $\beta > \frac{1}{2-\hat{\alpha}}$ (recall that $\frac{1}{2-\alpha} \in [1/2, 1[$). We choose a threshold $r_h = h^\beta$ and use the estimator $\hat{IQ}_h$ of the integrated quarticity defined in Proposition 2.3. We have shown in theorem 2.5 that, when $\sigma \not\equiv 0$ in the case $\alpha < 1$, the estimator $\hat{IV}_h$ is asymptotically Gaussian as $h \to 0$. However if $\sigma \equiv 0$ then both the numerator and the denominator of (13) tend to zero. To handle this case we add an IID noise term:

$$\Delta_i X^v := \Delta_i X + v\sqrt{h} Z_i \qquad Z_i \overset{\text{IID}}{\sim} N(0,1).$$

$$\text{As } h \to 0, \quad \sum_{i=1}^n (\Delta_i X^v)^2 \overset{P}{\to} [X^v]_T = \int_0^T \sigma_s^2 ds + v^2 T + T \int_{\mathbb{R}-\{0\}} x^2 \mu(dx, ds),$$

and $I_{\{(\Delta_i X^v)^2 \leq r_h\}}$ removes the jumps of X^v, so that under the assumptions of theorem 2.5, as $h \to 0$

$$\hat{IV}_h^v := \sum_{i=1}^n (\Delta_i X^v)^2 I_{\{(\Delta_i X^v)^2 \leq r_h\}} \overset{P}{\to} \int_0^T \sigma_s^2 ds + v^2 T.$$

Under the null hypothesis $\sigma \equiv 0$ we have $\hat{IV}_h^v \overset{P}{\to} v^2 T$, $\hat{IQ}_h^v := \sum_i (\Delta_i X^v)^4 I_{\{(\Delta_i X^v)^2 \leq r_h\}}/(3h) \overset{P}{\to} v^4 T$ and

$$U_h := \frac{\hat{IV}_h^v - v^2 T}{\sqrt{2h\hat{IQ}_h^v}} \overset{st}{\to} \mathcal{N}. \tag{14}$$

Note that if on the contrary $\sigma \not\equiv 0$ we have that the limit in probability of $\hat{IV}_h^v$ is strictly larger than $v^2 T$ and, by lemma 6.2, passing to a subsequence, a.s.

$$\lim_{h \to 0} h\ \hat{IQ}_h^v = \frac{1}{3} \lim_{h \to 0} \sum_i (\Delta_i X^v)^4 I_{\{(\Delta_i X^v)^2 \leq r_h\}} = \frac{1}{3} \lim_{h \to 0} \sum_i (\Delta_i X^v)^4 I_{\{\Delta_i N = 0, (\Delta_i M)^2 \leq 2r_h\}}$$

$$\leq \frac{1}{3}\lim_{h\to 0}\sum_i \left(\Delta_i X_0 + \Delta_i M + v\sqrt{h}Z_i\right)^4 I_{\{(\Delta_i M)^2\leq 2r_h\}}$$

$$\leq \frac{c}{3}\lim_{h\to 0}\sum_i(\Delta_i X_0)^4 + \frac{c}{3}\lim_{h\to 0}\sum_i(\Delta_i M)^4 I_{\{(\Delta_i M)^2\leq 2r_h\}} + \frac{c}{3}\lim_{h\to 0}\sum_i(v\sqrt{h}Z_i)^4.$$

Using that $\lim_{h\to 0}\sum_i(\Delta_i M)^4 I_{\{(\Delta_i M)^2\leq 2r_h\}} \leq \lim_{h\to 0} 2r_h\sum_i(\Delta_i M)^2 I_{\{(\Delta_i M)^2\leq 2r_h\}} = 0$ by (44), $\sum_i(\Delta_i X_0)^4/h \overset{P}{\to} c\int_0^T \sigma_s^4 ds$ and $\sum_i(v\sqrt{h}Z_i)^4/h \overset{a.s.}{\to} cv^4$, we have, as $h\to 0$, $h\ \hat{IQ}_h^v \overset{P}{\to} 0$. Therefore under the alternative (H_1) $\sigma\not\equiv 0$, $U_h\to+\infty$ and $P\{|U_h|>1.96\}\to 1$ so the test is consistent.

Local power of the test. To investigate the local power of test U_h, we consider a sequence of alternatives $H_1^h)$ $\sigma=\sigma^h$ where $\sigma^h\downarrow 0$. We denote by $\hat{IQ}_{\sigma^h}^v, U_{\sigma^h}$ the statistics analogous to $\hat{IQ}_h^v, U_h$ but constructed from $X_t^h = x_0 + \int_0^t a_s ds + \int_0^t \sigma_s^h\, dW_s + L_t,\ t\in]0,T]$. In the case of constant σ and σ^h and finite jump intensity using standard results on convergence of sums of a triangular array [14, Lemmas 4.1. and 4.3.],

$$\hat{IQ}_{\sigma^h}^v \overset{ucp}{\to} v^4 T,\quad U_{\sigma^h}\overset{d}{\to}\lim_{h\to 0}\frac{(\sigma^h)^2}{\sqrt{h}}T+\sqrt{2}v^2 Z_T,$$

where $\overset{ucp}{\to}$ denotes uniform convergence in probability on compacts subsets of $[0,T]$ [24] and Z is a standard Brownian motion. So either U_{σ^h} tends in distribution to $c+\sqrt{2}v^2 Z_T$, if $\sigma^h = O(h^{1/4})$, or $U_{\sigma^h}\to\infty$, if $h^{1/4}=o(\sigma^h)$. So if c is a (possibly zero) constant we have

$$\text{if } \tfrac{\sigma^h}{h^{1/4}}\to c \text{ then } P\{U_{\sigma^h}>1.64|H_1^h\}\to P\{Z_1>\tfrac{1.64-c^2T}{\sqrt{2T}v^2}\}$$

$$\text{if } \tfrac{\sigma^h}{h^{1/4}}\to +\infty \text{ then } P\{U_{\sigma^h}>1.64|H_1^h\}\to 1.$$

For values of v in section 4 we have $1.64/\sqrt{2T}v^2 = O(10^8)$ thus the local power of the test is small if $\sigma^h=O(h^{1/4})$.

3.2 Testing whether the jump component has finite variation

To construct a test for discriminating $\alpha<1$ from $\alpha\geq 1$, Theorem 2.5 suggests to use $(\hat{IV}_h - IV)/\sqrt{2h\hat{IQ}_h}$ but this requires knowing the process σ to compute IV. We propose a feasible alternative. Consider instead the estimator

$$\hat{H}_h := \sum_{i=1}^n \Delta_i X I_{\{(\Delta_i X)^2>r_h\}} = X_T - \sum_{i=1}^n \Delta_i X I_{\{(\Delta_i X)^2\leq r_h\}}.$$

Proposition 3.1. *When $\alpha<1$, $\hat{H}_h$ is a consistent estimator of J_T+mT, $m:=\int_{-1}^1 x\nu(dx)$.*

Consider $Z_i=\Delta_i W^v$, where W^v is a Wiener process independent from W, L and define

$$\Delta_i\hat{H}^v := \Delta_i X I_{\{(\Delta_i X)^2>r_h\}} + v\sqrt{h}Z_i \quad\text{and}\quad H_T^v := J_T+mT+vW_T^v.$$

Under the null hypothesis $\alpha<1$,

$$\hat{IV}_h^{H^v} := \sum_i(\Delta_i\hat{H}^v)^2 I_{\{(\Delta_i\hat{H}^v)^2\leq r_h\}}$$

is an estimator of the integrated variance v^2T of H^v, so under the null hypothesis (H_0) $\alpha<1$ we can find $\beta>\frac{1}{2-\alpha}\in]\frac{1}{2},1[$ such that

$$U_h^{(\alpha)} := \frac{\hat{IV}_h^{H^v}-v^2T}{\sqrt{2h\hat{IQ}_h^{H^v}}}\overset{d}{\to} N(0,1), \tag{15}$$

where $\hat{IQ}_h^{H^v} := \frac{1}{3h}\sum_i(\Delta_i\hat{H}^v)^4 I_{\{(\Delta_i\hat{H}^v)^2\leq r_h\}}$ and $r_h=h^\beta$. In particular $P\left\{|U_h^{(\alpha)}|>1.96\right\}\to 5\%$.

If on the contrary $\alpha\geq 1$, then reasoning as in theorem 2.5, for any $\beta\in]0,1[$ we have $U_h^{(\alpha)}\overset{P}{\to}+\infty$, so the test is consistent. If $|U_h^{(\alpha)}|>1.96$, we reject (H_0) $\alpha<1$ at 95% confidence level.

Remark. To apply this test we first need to decide whether $\alpha<1$, using the previously described test.

4 Numerical experiments

4.1 Testing the finite variation of the jump component

We simulate n increments $\Delta_i X$ of a process $X = \sigma W + L$, where L is a symmetric $\alpha-$stable Lévy process, $\sigma = 0.2$. We generate 1000 independent samples containing n increments each, and compute $U_h^{(\alpha)}$ as in (15) for a range of values of v, h (1 minute, 5 minutes, 1 hour, 1 day) and number of observations n. The table below reports the percentage *pct* of outcomes where $|U_{h(j)}^{(\alpha)}| > 1.96$, $j = 1..1000$, for threshold exponent $\beta = 0.999$. Note that with $n = 1000$ and h equal to five minutes ($h = 1/(252 \times 84)$) we have $T < 1$ year; for $\alpha = 0.6$ the lower bound for β is $\frac{1}{2-\alpha} = 0.71$; when $n = 1000, h = 1/(84 \times 252)$ and the BG index of L is 0.6 (resp. 1.6) the ratio of $v = 10^{-4}$ to the standard deviation of the increments $\Delta_i X$ is 0.074 (resp. 0.022).

n	h	v	α	*pct*	α	*pct*
1000	5 min	0.000001	0.6	0.067	1.6	0.439
1000	5 min	0.0001	0.6	0.056	1.6	0.407
1000	5 min	0.01	0.6	0.047	1.6	0.250
1000	5 min	0.1	0.6	0.053	1.6	0.726
1000	1 min	0.0001	0.6	0.049	1.6	0.241
1000	1 hour	0.0001	0.6	0.051	1.6	0.875
1000	1 day	0.0001	0.6	0.066	1.6	0.984
100	5 min	0.0001	0.6	0.065	1.6	0.137
10000	5 min	0.0001	0.6	0.065	1.6	0.928

The test results are observed to be reliable if we use $n = 10000$ observations, a time resolution of five minutes and $v = 10^{-4}$. In fact when the data generating process has BG index 0.6 the test leads us to accept the hypothesis (H_0) $\alpha < 1$ in about 94 cases out of 100. On the contrary when the process has BG index 1.6 the test tells us to reject (H_0) in 92 cases out of 100.

4.2 Test for the presence of a Brownian component.

We simulate 1000 independent paths of a process $X_t = \int_0^t \sigma_u dW_u + L$, for different Lévy processes L and constant or stochastic σ, on a time grid with n steps. We take threshold $r_h = h^{0.999}$. For each trial $j = 1..1000$ we compute $U_{h(j)}$ given in (14) and report the percentage *pct* of cases where $|U_{h(j)}| > 1.96$.

Example 4.1. (Brownian motion plus compound Poisson process, BG index $\alpha = 0$). We consider here constant σ and $L = \sum_{i=1}^{N_t} B_i$, a compound Poisson process with IID $N(0, 0.6^2)$ sizes of jump and jump intensity $\lambda = 5$ (as in [1]). The table below illustrates the performance of our test for various time steps h, number of observations n and noise level v:

n	h	v	σ	*pct*	σ	*pct*
1000	5 min	0.000001	0	0.043	0.2	1
1000	5 min	0.0001	0	0.048	0.2	1
1000	5 min	0.01	0	0.054	0.2	1
1000	5 min	0.1	0	0.041	0.2	1
1000	1 min	0.0001	0	0.047	0.2	1
1000	1 hour	0.0001	0	0.054	0.2	1
1000	1 day	0.0001	0	0.082	0.2	1
100	5 min	0.0001	0	0.065	0.2	1
10000	5 min	0.0001	0	0.049	0.2	1

Note that when $\sigma = 0$ (resp. 0.2), $n = 1000 and h = 1/(84 \times 252)$ the ratio of $v = 10^{-4}$ on the standard deviation of the returns $\Delta_i X$ equals 0.007 (resp. 0.052).

We find that the test is reliable for values $n = 1000, h = 5$ minutes and $v = 10^{-4}$, since it correctly accepts (H_0) in 95 cases out of 100 and rejects (H_0) in all cases when it is false.

Example 4.2. (Brownian motion + α-stable jumps: $\alpha \in]0, 2[$). Here L is a symmetric α-stable Lévy process and σ is constant. The following results confirm the satisfactory performance of the test when $\alpha = 0.3 < 1$ for $n = 1000, h = 5$ minutes and $v = 10^{-4}$ chosen as above.

n	h	v	σ	*pct*	σ	*pct*
1000	5 min	0.000001	0	0.042	0.2	1
1000	5 min	0.0001	0	0.026	0.2	1
1000	5 min	0.01	0	0.054	0.2	1
1000	5 min	0.1	0	0.053	0.2	1
1000	1 min	0.0001	0	0.046	0.2	1
1000	1 hour	0.0001	0	0.140	0.2	1
1000	1 day	0.0001	0	0.805	0.2	1
100	5 min	0.0001	0	0.056	0.2	1
10000	5 min	0.0001	0	0.165	0.2	1

The next table, for the case $\alpha = 1.2 > 1$, confirms that we cannot rely on the test results in this case: even when $\sigma \equiv 0$ the statistic U_h diverges if $\alpha \geq 1$.

n	h	v	σ	*pct*	σ	*pct*
1000	5 min	0.000001	0	1	0.2	1
1000	5 min	0.0001	0	1	0.2	1
1000	5 min	0.01	0	1	0.2	1
1000	5 min	0.1	0	1	0.2	1
1000	1 min	0.0001	0	1	0.2	1
1000	1 hour	0.0001	0	1	0.2	1
1000	1 day	0.0001	0	1	0.2	1
100	5 min	0.0001	0	0.994	0.2	1
10000	5 min	0.0001	0	1	0.2	1

The main point here is that we may use a model-free choice of threshold.

Example 4.3. (Stochastic volatility plus Variance Gamma jumps: $\alpha = 0$). Let us now consider a model X with stochastic volatility σ_t, correlated with the Brownian motion driving X, and with jumps given by an independent Variance Gamma process:

$$dX_t = (\mu - \sigma_t^2/2)dt + \sigma_t dW_t^{(1)} + dL_t,$$

where

$$\sigma_t = e^{K_t}, \quad dK_t = -k(K_t - \bar{K})dt + \varsigma dW_t^{(2)}, \quad d < W^{(1)}, W^{(2)} >_t = \rho dt, \tag{16}$$

$W^{(\ell)}$ are standard Brownian motions, $\ell = 1, 2, 3$ and $L_t = cG_t + \eta W_{G_t}^{(3)}$ is an independent Variance Gamma process, a pure jump Lévy process with BG index $\alpha = 0$ [18]: G is a Gamma subordinator independent from $W^{(3)}$ with $G_h \sim \Gamma(h/b, b)$. For σ we choose $K_0 = \ln(0.3)$, $k = 0.09$, $\bar{K} = \ln(0.25)$, $\varsigma = 0.05$ to ensure that σ fluctuates in the range 0.2–0.4. As for the jump part of X we use $Var(G_1) = b = 0.23$, $\eta = 0.2$, $c = -0.2$, estimated from the SP500 index in [18]. The remaining parameters are $\rho = -0.7$ and $\mu = 0$. The following results confirm the reliability of the test for the presence of a Brownian component with $n = 1000, h = 5$ minutes and $v = 10^{-4}$.

n	h	v	σ	pct	σ	pct
1000	5 min	0.000001	0	0.032	stoch	1
1000	5 min	0.0001	0	0.017	stoch	1
1000	5 min	0.01	0	0.027	stoch	1
1000	5 min	0.1	0	0.054	stoch	1
1000	1 min	0.0001	0	0.034	stoch	1
1000	1 hour	0.0001	0	0.918	stoch	1
1000	1 day	0.0001	0	1.000	stoch	1
100	5 min	0.0001	0	0.049	stoch	1
10000	5 min	0.0001	0	0.912	stoch	1

Remark. In [21] a variable threshold function is used to estimate the volatility, in order to account for heteroskedasticity and volatility clustering, with results very similar to the ones obtained with a constant threshold. This is justified by the fact that in most applications values of σ are within the range [0.1, 0.8], thus the order of magnitude of Λ in (7) is of 1.

5 Applications to financial time series

We apply our tests to explore the fine structure of price fluctuations in two financial time series. We consider the DM/USD exchange rate from 1-10-1991 to 29-11-1994 and the SPX futures prices from 3-1-1994 to 18-12-1997. From high-frequency time series, we build 5-minute log-returns (excluding, in the case of SPX futures, overnight log-returns). This sampling frequency avoids many microstructure effects seen at shorter time scales (e.g. seconds) while leaving us with a relatively large sample.

5.1 Deutschemark/USD exchange rate

The DM/\$ exchange rate time series was compiled by Olsen & Associates. We consider the series of 64284 equally spaced 5 minutes log-returns, with $h = \frac{1}{252\times 84} \approx 4.7 \times 10^{-5}$, displayed in Figure 5.2.

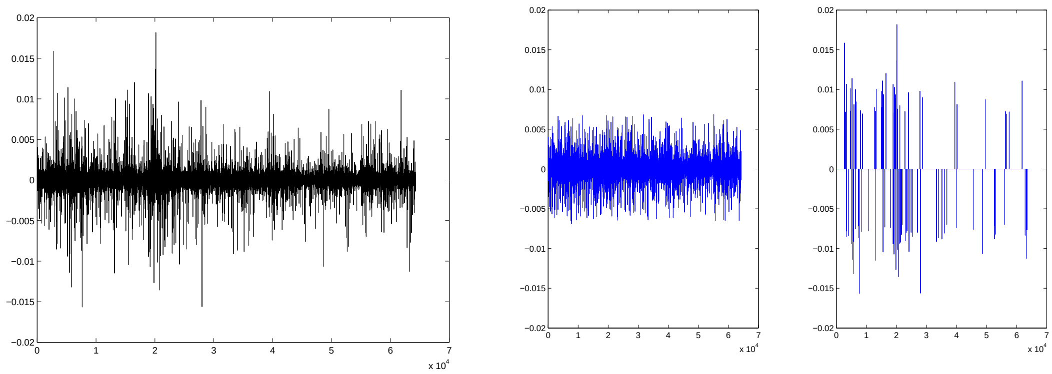

Figure 5.1. Left: DEM/USD five minutes log-returns, October 1991 to November 1994. Center: plot of $\Delta_i X I_{\{(\Delta_i X)^2 \le r_h\}}$, $i = 1..n$. Right: increments with jumps $\Delta_i X I_{\{(\Delta_i X)^2 > r_h\}}$, $i = 1..n$.

Barndorff-Nielsen and Shephard [6] provide evidence for the presence of jumps in this series using non-parametric methods. Using as threshold $r_h = h^{0.999}$, we apply the test of section 4.1 to the degree of activity of the jump component. As in the simulation study, we divide the data into 64 non-overlapping batches of $n = 1000$ observations each and compute for each batch the statistic $U^{(\alpha)}_{h(j)}$, $j = 1..64$ with $v = 10^{-4}$. Only 4.7% of the values observed are outside the interval $[-1.96, 1.96]$, hence we cannot reject the assumption (H_0) $\alpha < 1$.

Given this result, we can now use the test in Section 4.2 for the presence of a Brownian component in the price process. Computation of the statistic U_h shows values much larger than 1.96 for all batches: we reject (H_0) $\sigma \equiv 0$.

These results indicate, for instance, that a Variance Gamma model, with no Brownian component, would be inadequate for the DM/$ time series.

5.2 SPX index

We consider the series of 78497 non-overlapping 5 minute log-returns displayed in Figure 5.2. Using as threshold $r_h = h^{0.999}$, we decompose the series into periods displaying jumps and other periods, as displayed in Figure 5.2 (central and right panels).

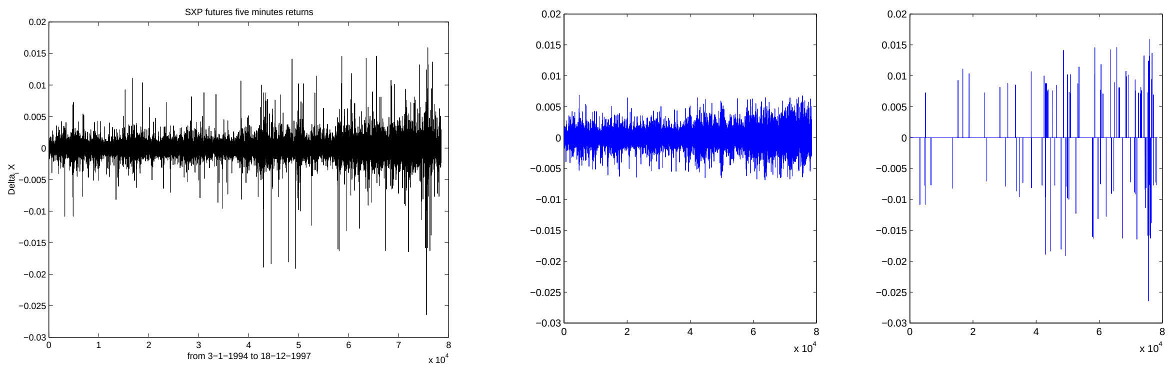

Figure 5.2. Left: SPX five minutes log-returns, January 1994 to December 1997. Center: plot of $\Delta_i X I_{\{(\Delta_i X)^2 \leq r_h\}}$, $i = 1..n$. Right: increments with jumps $\Delta_i X I_{\{(\Delta_i X)^2 > r_h\}}$, $i = 1..n$.

We divide the data into 78 non-overlapping batches of $n = 1000$ observations each and compute for each batch the statistic $U_{h(j)}^{(\alpha)}$, $j = 1..64$ with $v = 10^{-4}$. 5.1% of the values observed are outside the interval $[-1.96, 1.96]$: for this period we cannot reject the assumption (H_0) $\alpha < 1$. Given this result, we can use the test for the presence of a Brownian component in the price process. Computation of the statistic U_h shows values much larger than 1.96 for all batches: we reject (H_0) $\sigma \equiv 0$. The test thus indicates the presence of a Brownian martingale component.

We note that our findings contradict the conclusion of Carr et al. [8] who model the (log) SPX index from 1994 to 1998 as a tempered stable Lévy process plus a Brownian motion and conclude towards a pure jump model using a parametric estimation method. Under less restrictive assumptions on the structure of the process and using our non-parametric test, we find evidence for a non–zero Brownian component in the index.

6 Appendix: technical results and proofs

Proof of lemma 2.1. By [23, Theorem 25.1] there exists a sequence (n_k) such that

$$\sup_{t_j \in \Pi^{(n_k)}} \left|(\Delta_j M)^2 - \sum_{s \in]t_{j-1}, t_j]} (\Delta M_s)^2\right| \stackrel{a.s}{\to} 0, \tag{17}$$

where $\Pi^{(n_k)}$ is the partition of $[0, T]$ on which the increments $(\Delta_i M)^2$ are constructed. Let us rename n_k by n. Using Ito's formula we have

$$(\Delta_i M)^2 - \sum_{s \in]t_{i-1}, t_i]} (\Delta M_s)^2 = 2 \int_{t_{i-1}}^{t_i} (M_{s_-} - M_{t_{i-1}}) dM_s.$$

i) For $\alpha < 1$ our statement is proved in [21], lemma A.2, where it is used that the speed of convergence to 0 of $\sum_{i=1}^n |\int_{t_{i-1}}^{t_i} (M_{s_-} - M_{t_{i-1}})dM_s|$ is shown in [12] to be $u_n = n$. For $\alpha = 1$ the same reasoning can be repeated since $u_n = n/(\log n)^2$ does not change the conclusion.
ii) If $\alpha > 1$ we have $u_n = (n/\log n)^{1/\alpha}$, and we can only conclude that a.s. for small h

$$\sup_i \Big| \int_{t_{i-1}}^{t_i} (M_{s_-} - M_{t_{i-1}})dM_s \Big| \leq c u_n^{-1}$$

with $c > 0$, so that a.s. for small h we have

$$\sup_i \Big(\sum_{s\in]t_{i-1},t_i]} (\Delta M_s)^2 \Big) I_{\{(\Delta_i M)^2 \leq 4r_h\}} \leq \sup_i \Big| (\Delta_i M)^2 - \sum_{s\in]t_{i-1},t_i]} (\Delta M_s)^2 \Big| + \sup_i |(\Delta_i M)^2|$$

$$\leq c u_n^{-1} + 4r_h = O(\delta^{\frac{1}{\alpha}} \log^{\frac{1}{\alpha}} \frac{1}{h}).$$

□

Lemma 6.1. *Under (5)*
i) there exists a strictly positive variable $\bar{h}$ such that for all $i = 1..n$,

$$I_{\{h\leq \bar{h}\}}\ I_{\{(\Delta_i X_0)^2 > r_h\}} = 0 \quad a.s., \tag{18}$$

$$ii) \quad \forall c > 0, \quad nP\{\Delta_i N \neq 0, (\Delta_i M)^2 > cr_h\} \overset{h\to 0}{\to} 0 \tag{19}$$

iii) In the case $r_h = h^\beta$, $\beta \in]0,1[$, we have

$$\limsup_{h\to 0} h^{\frac{\alpha\beta}{2}} \sum_{i=1}^n P\{(\Delta_i X)^2 > r_h\} \leq c \tag{20}$$

Proof. Equality (18) is a consequence of (7), while (19) is a consequence of the fact that N and M are independent and of the Chebyshev inequality: as $h \to 0$

$$nP\{\Delta_i N \neq 0, (\Delta_i M)^2 > cr_h\} \leq nO(h) \cdot \frac{E[(\Delta_i M)^2]}{cr_h} = O\left(\frac{h}{r_h}\right).$$

The proof of (20) can be done as in [3, Lemma 6] but we give a simpler proof under our assumptions. It is sufficient to show that

$$P\{(\Delta_i X)^2 > r_h\} \leq ch^{1-\frac{\alpha\beta}{2}}. \tag{21}$$

First we show that

$$P\{|\Delta_i X| > \sqrt{r_h}\} = P\{|\Delta_i M| > \sqrt{r_h}/4\} + O(h^{1-\alpha\beta/2}) \tag{22}$$

so that for (21) it is sufficient to prove that

$$P\{|\Delta_i M| > \sqrt{r_h}/4\} \leq ch^{1-\frac{\alpha\beta}{2}}. \tag{23}$$

To show (22) note that if $|\Delta_i X| > \sqrt{r_h}$ either $\Delta_i J \neq 0$ or $|\Delta_i M| > \sqrt{r_h}/4$, since for small h,

$$\sqrt{r_h} < |\Delta_i X| \leq |\Delta_i X_0| + |\Delta_i J| + |\Delta_i M| \leq \sqrt{r_h}/2 + |\Delta_i J| + |\Delta_i M| \quad a.s. \tag{24}$$

$$\text{Thus} \quad P\{|\Delta_i X| > \sqrt{r_h}\} \leq P\{\Delta_i J \neq 0\} + P\{|\Delta_i M| > \sqrt{r_h}/4\},$$

and since $P\{\Delta_i J \neq 0\} = O(h) = o(h^{1-\alpha\beta/2})$, (22) is verified.
In order to verify (23), define $\tilde{N}_t := \sum_{s\leq t} I_{\{|\Delta M_s| > \sqrt{r_h}/4\}}$ and write

$$P\{|\Delta_i M| > \sqrt{r_h}/4\} = P\{\Delta_i \tilde{N} = 0, |\Delta_i M| > \sqrt{r_h}/4\} + P\{\Delta_i \tilde{N} \geq 1, |\Delta_i M| > \sqrt{r_h}/4\}$$

$$\leq P\{\Delta_i \tilde{N} \geq 1\} + P\{\Delta_i \tilde{N} = 0, |\Delta_i M| > \sqrt{r_h}/4\}. \tag{25}$$

Note that $\tilde{N}_t = \int_0^t \int_{|x|>\sqrt{r_h}/4} \mu(dx,dt)$ is a compound Poisson process with intensity $\nu\{|x| > \sqrt{r_h}/4\} = O(r_h^{-\alpha/2})$, so $P\{\Delta_i \tilde{N} \geq 1\} = O(h\nu\{|x| > \sqrt{r_h}/4\}) = O(h^{1-\alpha\beta/2})$ and thus the first term above is dominated by $h^{1-\alpha\beta/2}$ as we need. Finally on $\{\Delta_i\tilde{N} = 0\}$, M does not have jumps bigger than $\sqrt{r_h}/4$ on the interval $]t_{i-1}, t_i]$, so

$$\Delta_i M = \int_{t_{i-1}}^{t_i} \int_{|x|\leq\sqrt{r_h}/4} x\tilde{\mu}(dx,dt) - h\int_{\sqrt{r_h}/4<|x|\leq 1} x\nu(dx),$$

therefore

$$P\{\Delta_i\tilde{N} = 0, |\Delta_i M| > \sqrt{r_h}/4\} \leq P\{\ |\Delta_i M| > \sqrt{r_h}/4, |\Delta M_s| \leq \sqrt{r_h}/4 \text{ for all } s \in]t_{i-1}, t_i]\}$$

$$\leq 4\frac{E\left[(\Delta_i M)^2 I_{\{|\Delta M_s|\leq\sqrt{r_h}/4 \text{ for all } s\in]t_{i-1},t_i]\}}\right]}{r_h} = O\Big(\frac{h\eta^2(\frac{\sqrt{r_h}}{4})}{r_h}\Big) = O(h^{1-\alpha\beta/2}),$$

and (23) is verified. □

Proof of proposition 2.3.

$$\frac{\sum_i(\Delta_i X)^4 I_{\{(\Delta_i X)^2\leq r_h\}}}{3h} = \frac{\sum_i(\Delta_i X_1)^4 I_{\{(\Delta_i X_1)^2\leq 4r_h\}}}{3h} +$$

$$\frac{1}{3h}\sum_i(\Delta_i X_1)^4(I_{\{(\Delta_i X)^2\leq r_h\}} - I_{\{(\Delta_i X_1)^2\leq 4r_h\}}) + \sum_{k=1}^4 \binom{4}{k}\frac{\sum_i(\Delta_i X_1)^{4-k}(\Delta_i M)^k I_{\{(\Delta_i X)^2\leq r_h\}}}{3h} := \sum_{j=1}^3 I_j(h)$$

By proposition 1 in [20], $I_1(h)$ tends to $\int_0^T \sigma_t^4 dt$ in probability. We show here that the other terms tends to zero in probability. Let us consider $I_2(h) := \frac{1}{3h}\sum_i(\Delta_i X_1)^4(I_{\{(\Delta_i X)^2\leq r_h\}} - I_{\{(\Delta_i X_1)^2\leq 4r_h\}})$: on $\{(\Delta_i X)^2 \leq r_h, (\Delta_i X_1)^2 > 4r_h\}$ we have

$$\sqrt{r_h} \geq |\Delta_i X| > |\Delta_i X_1| - |\Delta_i M| > 2\sqrt{r_h} - |\Delta_i M| \tag{26}$$

so $|\Delta_i M| > \sqrt{r_h}$. Moreover if $|\Delta_i X_1| > 2\sqrt{r_h}$ we necessarily have $\Delta_i N \neq 0$, since

$$|\Delta_i X_0| + |\Delta_i J| \geq |\Delta_i X_1| > 2\sqrt{r_h} \tag{27}$$

and by (18), a.s. for sufficiently small h, for all $i = 1..n$, $|\Delta_i X_0| \leq \sqrt{r_h}$ thus $|\Delta_i J| > 2\sqrt{r_h} - |\Delta_i X_0| \geq \sqrt{r_h}$. It follows that

$$P\Big\{\frac{1}{h}\sum_i(\Delta_i X_1)^4 I_{\{(\Delta_i X)^2\leq r_h, (\Delta_i X_1)^2>4r_h\}} \neq 0\Big\} \leq nP\{|\Delta_i M| > \sqrt{r_h}, \Delta_i N \neq 0\} \to 0$$

by lemma 6.1. On the other hand, for all $i = 1..n$ on $\{(\Delta_i X_1)^2 \leq 4r_h\}$ we have, for sufficiently small h, $\Delta_i N = 0$, because

$$|\Delta_i J| - |\Delta_i X_0| \leq |\Delta_i X_1| \leq 2\sqrt{r_h} \tag{28}$$

so if $\Delta_i N \neq 0$ then a.s. for small h in fact $\Delta_i N = 1$ and $\Delta J_s \geq 1$ by definition of J, therefore if $\Delta_i N \neq 0$ we would have $1 \leq |\Delta_i J| \leq 2\sqrt{r_h} + \sqrt{r_h} = 3\sqrt{r_h}$, which is impossible for small h. It follows that

$$\{(\Delta_i X)^2 > r_h, (\Delta_i X_1)^2 \leq 4r_h\} \subset \{(\Delta_i X_0 + \Delta_i M)^2 > r_h\} \subset \{(\Delta_i X_0)^2 > \frac{r_h}{4}\} \cup \{(\Delta_i M)^2 > \frac{r_h}{4}\}.$$

This implies by (18) and (23) that a.s. as $h \to 0$

$$\frac{1}{h}\sum_i(\Delta_i X_1)^4 I_{\{(\Delta_i X)^2>r_h, (\Delta_i X_1)^2\leq 4r_h\}} \leq \frac{\sum_i(\Delta_i X_0)^4 I_{\{(\Delta_i M)^2>r_h/4\}}}{h}$$

$$\leq \Lambda^4 h \ln^2 \frac{1}{h}\sum_i I_{\{(\Delta_i M)^2>r_h/4\}} \xrightarrow{P} 0,$$

We can conclude that $I_2(h) \xrightarrow{P} 0$, as $h \to 0$. Now consider $I_3(h) := \sum_{k=1}^4 \binom{4}{k} I_{3,k}(h)$, where

$$I_{3,k}(h) := \frac{1}{3h} \sum_i (\Delta_i X_1)^{4-k} (\Delta_i M)^k I_{\{(\Delta_i X)^2 \le r_h\}}, \quad k = 1..4$$

is decomposable as

$$\frac{1}{3h} \sum_i (\Delta_i X_1)^{4-k} (\Delta_i M)^k I_{\{(\Delta_i X)^2 \le r_h, (\Delta_i M)^2 \le 4r_h\}} + \frac{1}{3h} \sum_i (\Delta_i X_1)^{4-k} (\Delta_i M)^k I_{\{(\Delta_i X)^2 \le r_h, (\Delta_i M)^2 > 4r_h\}} \tag{29}$$

We have a.s. for small h, $\forall i$ on $\{(\Delta_i X)^2 \le r_h, (\Delta_i M)^2 > 4r_h\}$ then $\Delta_i N \neq 0$, since

$$2\sqrt{r_h} - |\Delta_i X_1| < |\Delta_i M| - |\Delta_i X_1| \le |\Delta_i X| \le \sqrt{r_h}$$

and then $|\Delta_i X_1| > \sqrt{r_h}$ and, similarly as in (27), $|\Delta_i J| > 3\sqrt{r_h}/4$. So the probability that the second term of (29) differs from zero is bounded by (19) and tends to zero. As for the first term, a.s. for sufficiently small h, $\forall i$ on $\{(\Delta_i X)^2 \le r_h, (\Delta_i M)^2 \le 4r_h\}$ we have $\Delta_i N = 0$, because

$$|\Delta_i X_1| - |\Delta_i M| \le |\Delta_i X| \le \sqrt{r_h}$$

thus $|\Delta_i X_1| < 3\sqrt{r_h}$ and we proceed as in (28). So the first term in (29) is a.s. dominated by

$$\frac{\sum_i |\Delta_i X_1|^{4-k} |\Delta_i M|^k I_{\{\Delta_i N = 0, (\Delta_i M)^2 \le 4r_h\}}}{3h} \le \frac{\sum_i |\Delta_i X_0|^{4-k} |\Delta_i M|^k I_{\{(\Delta_i M)^2 \le 4r_h\}}}{3h}.$$

Now for $k = 4$ we apply to M property (C.19) in [4, Lemma 5], with β there being α here, $u_n = \sqrt{r_h} = h^{\frac{\beta}{2}}$, $p = 4$, $v_h = h^\phi$ for a proper exponent ϕ we specify below, $\beta' = 0$. Result (C.19) of [4] then implies

$$\frac{1}{h} E\Big[\Big| \sum_{i=1}^n (\Delta_i M)^4 I_{\{|\Delta_i M| \le 2\sqrt{r_h}\}} - \sum_{v \le T} |\Delta M_v|^4 I_{\{|\Delta M_v| \le 2\sqrt{r_h}\}} \Big|\Big] \le c h^{\frac{\beta}{2}(4-\alpha)-1} \cdot \eta_{4,n},$$

where $\eta_{4,n} = h(h^{\frac{\beta}{2}} v_h)^{-\alpha} + h^2 h^{\frac{\alpha\beta}{2}} (h^{\frac{\beta}{2}} v_h)^{-3\alpha} + h h^{\frac{\alpha\beta}{2}} (h^{\frac{\beta}{2}})^{-2\alpha} + (2h^{\frac{\beta}{2}})^\alpha + h^{\frac{1}{4}} h^{-\frac{4-\alpha}{4}\frac{\beta}{2}} + v_h^{\frac{4-\alpha}{4}}$. As soon as $\beta > 1/(2 - \alpha/2)$ and we choose $\phi \in]0, \frac{1-\beta}{3}[$, so that for all $\alpha \in]0, 2[$ we have $\phi < (2/\alpha - \beta)/3$, it is guaranteed both that $h^{\frac{\beta}{2}(4-\alpha)-1} \to 0$ and that $h^{\frac{\beta}{2}(4-\alpha)-1} \cdot \eta_{4,n} \to 0$. Thus

$$\lim_h \frac{\sum_i |\Delta_i M|^4 I_{\{(\Delta_i M)^2 \le 4r_h\}}}{3h} = \lim_h \frac{\sum_i \int_{t_{i-1}}^{t_i} \int_{|x| \le 2\sqrt{r_h}} |x|^4 \mu(dx, dt)}{3h},$$

and $E[\sum_i \int_{t_{i-1}}^{t_i} \int_{|x| \le 2\sqrt{r_h}} |x|^4 \mu(dx, dt)/3h] = O\Big(\int_{|x| \le 2\sqrt{r_h}} |x|^4 \nu(dx)/h\Big) = O(h^{\frac{\beta}{2}(4-\alpha)-1}) \to 0$, given that $\beta > 1/(2 - \alpha/2)$.

To show that further the terms $\frac{\sum_i |\Delta_i X_0|^{4-k} |\Delta_i M|^k I_{\{(\Delta_i M)^2 \le 4r_h\}}}{3h}$ tend to zero in probability for $k = 1, 2, 3$ we use that, by (11), each term is dominated by (recall the notation in (10))

$$c\, \frac{\sum_i |\Delta_i X_0|^{4-k} |\Delta_i M^{(h)}|^k}{3h} + c\, \frac{\sum_i |\Delta_i X_0|^{4-k} |hd(2\sqrt[4]{r_h})|^k}{3h}$$

Now a.s.

$$\frac{\sum_i |\Delta_i X_0|^{4-k} |hd(2\sqrt[4]{r_h})|^k}{3h} \le \big(h \ln \frac{1}{h}\big)^{\frac{4-k}{2}} n h^{k-1} \Big[|c + r_h^{\frac{1-\alpha}{4}}|^k I_{\{\alpha \neq 1\}} + \ln^k \frac{1}{r_h^{1/4}} I_{\{\alpha = 1\}} \Big] \le$$

$$c h^{k/2} \big(\ln \frac{1}{h}\big)^{\frac{4-k}{2}} + c h^{k/2} \big(\ln \frac{1}{h}\big)^{\frac{4-k}{2}} \; r_h^{k\frac{1-\alpha}{4}} + h^{\frac{h}{2}} \ln^{2-k/2} \frac{1}{h r_h^{1/4}} = o(1) + c h^{k[\frac{1}{2} + \beta\frac{1-\alpha}{4}]} \log^{\frac{4-k}{2}} \frac{1}{h} \to 0$$

$\forall\ k=1,2,3$ as $r_h=h^\beta$, $\beta\in]0,1[$. As for

$$\frac{\sum_i|\Delta_i X_0|^{4-k}|\Delta_i M^{(h)}|^k}{3h}, \tag{30}$$

we need to deal separately with each $k=1,2,3$. Note that since a and σ are locally bounded on $\Omega\times[0,T]$, we can assume they are bounded without loss of generality, so $E[(\int_{t_{i-1}}^{t_i}\sigma_s dW_s)^{2k}]=O(h^k)$ for each $k=1,2,3$, using e.g. the Burkholder inequality [24, p. 226], and a.s. $(\int_{t_{i-1}}^{t_i}a_s ds)^{2k}=o(h^k)$. Therefore $E[(\Delta_i X_0)^{2k}]=O(h^k)$ for each $k=1,2,3$. For $k=1$ the expected value of (30) is bounded by $\frac{n}{3h}\sqrt{E[(\Delta_i X_0)^6]}\sqrt{E(\Delta_i M^{(h)})^2}=O(r_h^{\frac14(1-\frac{\alpha}{2})})$ and thus it tends to zero as $h\to 0$. As for $k=2$,

$$\frac{\sum_i(\Delta_i X_0)^2(\Delta_i M^{(h)})^2}{h}\le h\ln\frac1h\,\frac{\sum_i(\Delta_i M^{(h)})^2}{h}, \tag{31}$$

whose expected value is given by

$$\ln\frac1h\ \eta^2(2r_h^{\frac14})\to 0,$$

as $h\to 0$, since $r_h=h^\beta$, with $\beta>0$. Concerning $k=3$, we have

$$\frac{\sum_i|\Delta_i X_0||\Delta_i M^{(h)}|^3}{h}\le\frac{c}{h}\sum_i(\Delta_i X_0)^2(\Delta_i M^{(h)})^2+\frac{c}{h}\sum_i(\Delta_i M^{(h)})^4,$$

so that this step is reduced to the steps with $k=2,4$ we dealt with previously. □

Proof of theorem 2.4. Let us define $K_{ni}:=\left(\int_{t_{i-1}}^{t_i}\int_{|x|\le\varepsilon}x\tilde\mu(dx,dt)-h\int_{\varepsilon<|x|\le 1}x\nu(dx)\right)^2$. We apply the Lindeberg-Feller theorem to the double array sequence H_{ni} given by the normalized versions of the variables K_{ni}, $i=1,..,n$ and $n=T/h$. Using relations (9) we have

$$E\left[K_{ni}\right]=h\ell_{2,h}\varepsilon^{2-\alpha}+\left(h\int_{\varepsilon<|x|\le1}x\nu(dx)\right)^2=h\ell_{2,h}\varepsilon^{2-\alpha}+\ell_{1,h}^2h^2\Big[(c+\varepsilon^{1-\alpha})^2I_{\{\alpha\neq1\}}+\big(\ln^2\frac1\varepsilon\big)I_{\{\alpha=1\}}\Big]. \tag{32}$$

Taking $\varepsilon=h^u$, any $u\in]0,1/2]$, we obtain that

$$v_{ni}^2:=var\left[K_{ni}\right]=E\left[\left(\int_{t_{i-1}}^{t_i}\int_{|x|\le\varepsilon}x\tilde\mu(dx,dt)-h\int_{\varepsilon<|x|\le1}x\nu(dx)\right)^4\right]-E_{ni}^2\sim h\int_{|x|\le\varepsilon}x^4\nu(dx)=h\ell_{4,h}\varepsilon^{4-\alpha},$$

as $h\to 0$. Consider then

$$H_{ni}:=\frac{K_{ni}-E[K_{ni}]}{\sqrt{n}\ v_{ni}}\sim\frac{K_{ni}-h\ell_{2,h}\varepsilon^{2-\alpha}-\ell_{1,h}^2h^2\Big[(c+\varepsilon^{1-\alpha})^2I_{\{\alpha\neq1\}}+\big(\ln^2\frac1\varepsilon\big)I_{\{\alpha=1\}}\Big]}{\sqrt{T}\sqrt{\ell_{4,h}}\varepsilon^{2-\alpha/2}}$$

We now show that for any $\delta>0$, there exists a $q>1$ such that

$$\sum_{i=1}^n E[H_{ni}^2 I_{\{|H_{ni}|>\delta\}}]\le c\varepsilon^{\frac{\alpha}{2q}}\to 0, \tag{33}$$

as $h\to 0$, so the Lindeberg condition is satisfied and implies that

$$\sum_{i=1}^n H_{ni}\xrightarrow{d}N(0,1). \tag{34}$$

Noting that $h/\varepsilon^{2-\alpha/2}$ and $(h\varepsilon^{1-\alpha})/(\varepsilon^{2-\alpha/2})I_{\{\alpha\neq1\}}+(h\ln^2(1/\varepsilon))/(\varepsilon^{2-\alpha/2})I_{\{\alpha=1\}}$ tend to zero as $h\to 0$, (34) leads to (12). To show inequality (33), consider

$$nE[H_{n1}^2I_{\{|H_{n1}|>\delta\}}]\le nE^{\frac1p}[H_{n1}^{2p}]P^{\frac1q}\{|H_{n1}|>\delta\}: \tag{35}$$

as for the last factor above we note that $|H_{n1}| > \delta$ iff either

$$K_{n1} < h\ell_{2,h}\varepsilon^{2-\alpha} + \ell_{1,h}^2 h^2 \Big[(c+\varepsilon^{1-\alpha})^2 I_{\{\alpha\neq 1\}} + \big(\ln^2 \frac{1}{\varepsilon}\big) I_{\{\alpha=1\}}\Big] - \delta\sqrt{T\ell_{4,h}}\varepsilon^{2-\frac{\alpha}{2}} = \varepsilon^{2-\frac{\alpha}{2}}\big(o(1) - c\delta\big)$$

where c denotes a generic constant, or

$$K_{n1} > h\ell_{2,h}\varepsilon^{2-\alpha} + \ell_{1,h}^2 h^2 \Big[(c+\varepsilon^{1-\alpha})^2 I_{\{\alpha\neq 1\}} + \big(\ln^2 \frac{1}{\varepsilon}\big) I_{\{\alpha=1\}}\Big] + c\delta\varepsilon^{2-\frac{\alpha}{2}} = O(\varepsilon^{2-\frac{\alpha}{2}}),$$

However $K_{n1} \geq 0$ while for sufficiently small h the right hand term of the first inequality above is strictly negative, therefore $|H_{n1}| > \delta$ iff $K_{n1} > c\varepsilon^{2-\frac{\alpha}{2}}$, i.e. either

$$-c\varepsilon^{1-\frac{\alpha}{4}} \sim h(c+\varepsilon^{1-\alpha}) I_{\{\alpha\neq 1\}} + I_{\{\alpha=1\}} h \ln\frac{1}{\varepsilon} - c\varepsilon^{1-\frac{\alpha}{4}} > \int_0^{t_1}\int_{|x|\leq\varepsilon} x\tilde{\mu}(dx,dt)$$

or, for sufficiently small h, $\int_0^{t_1}\int_{|x|\leq\varepsilon} x\tilde{\mu}(dx,dt) > c\varepsilon^{1-\frac{\alpha}{4}}$, and so $|H_{n1}| > \delta$ iff

$$\left|\int_0^{t_1}\int_{|x|\leq\varepsilon} x\tilde{\mu}(dx,dt)\right| > c\varepsilon^{1-\frac{\alpha}{4}}.$$

This entails that for sufficiently small h,

$$P\{|H_{n1}| > \delta\} = P\left\{\Big|\int_0^{t_1}\int_{|x|\leq\varepsilon} x\tilde{\mu}(dx,dt)\Big| > c\varepsilon^{1-\frac{\alpha}{4}}\right\} \leq c\frac{E[|\int_0^{t_1}\int_{|x|\leq\varepsilon} x\tilde{\mu}(dx,dt)|^2]}{\varepsilon^{2-\frac{\alpha}{2}}} = h^{1-\frac{\alpha u}{2}} \to 0.$$

The first two factors of the r.h.s in (35) are dominated by

$$cn\frac{E^{\frac{1}{p}}\left[\left(K_{n1} - h\ell_{2,h}\varepsilon^{2-\alpha} - h^2\ell_{1,h}^2\Big[(c+\varepsilon^{1-\alpha})^2 I_{\{\alpha\neq 1\}} + (\ln^2\frac{1}{\varepsilon}) I_{\{\alpha=1\}}\Big]\right)^{2p}\right]}{\varepsilon^{4-\alpha}}$$

$$\leq cn\frac{E^{\frac{1}{p}}\left[K_{n1}^{2p}\right] + (h\varepsilon^{2-\alpha})^2 + h^4(1-\varepsilon^{1-\alpha})^4 + h^4\ln^4\frac{1}{\varepsilon}}{\varepsilon^{4-\alpha}},$$

The last three terms give no contribution to (35) since

$$n\frac{(h\varepsilon^{2-\alpha})^2 + h^4(1-\varepsilon^{1-\alpha})^4 + h^4\ln^4\frac{1}{\varepsilon}}{\varepsilon^{4-\alpha}} h^{(1-\frac{\alpha u}{2})\frac{1}{q}} \to 0.$$

On the other hand, by choosing e.g. $p = 5/4$ we have

$$E\left[K_{n1}^{2p}\right] = O(h\varepsilon^{5-\alpha}),$$

so we are left to deal with $n\frac{(h\varepsilon^{5-\alpha})^{\frac{1}{p}}}{\varepsilon^{4-\alpha}} h^{(1-\frac{\alpha u}{2})\frac{1}{q}} = \varepsilon^{\frac{\alpha}{2q}}$, so that the inequality in (33) is proved. □

Lemma 6.2. *As $h \to 0$: if $r_h \to 0$, $n = T/h$ and $\sup_{i=1..n}|a_{hi}| = O(r_h)$ then*

$$\sum_i |a_{hi}| I_{\{(\Delta_i X)^2 \leq r_h\}} - \sum_i |a_{hi}| I_{\{(\Delta_i M)^2 \leq 4r_h, \Delta_i N = 0\}} \stackrel{P}{\to} 0.$$

Proof. On $\{(\Delta_i X)^2 \leq r_h\}$ we have $|\Delta_i L| - |\Delta_i X_0| \leq |\Delta_i X| \leq \sqrt{r_h}$ and thus, by (7), for small h, $|\Delta_i L| \leq 2\sqrt{r_h}$, so that a.s.

$$\lim_{h\to 0}\sum_i |a_{hi}| I_{\{(\Delta_i X)^2 \leq r_h\}} \leq \lim_{h\to 0}\sum_i |a_{hi}| I_{\{(\Delta_i L)^2 \leq 4r_h\}}.$$

However

$$\sum_i |a_{hi}| I_{\{(\Delta_i L)^2 \leq 4r_h, \Delta_i N \neq 0\}} \leq \sup_i |a_{hi}| N_T \stackrel{a.s.}{\to} 0, \tag{36}$$

as $h \to 0$, and thus a.s.

$$\lim_{h\to 0}\sum_i |a_{hi}|I_{\{(\Delta_i X)^2\le r_h\}} \le \lim_{h\to 0}\sum_i |a_{hi}|I_{\{(\Delta_i L)^2\le 4r_h,\Delta_i N=0\}} = \lim_{h\to 0}\sum_i |a_{hi}|I_{\{(\Delta_i M)^2\le 4r_h,\Delta_i N=0\}}.$$

Now we show that on the other hand the positive quantity

$$\lim_{h\to 0}\sum_i |a_{hi}|(I_{\{(\Delta_i L)^2\le 4r_h,\Delta_i N=0\}} - I_{\{(\Delta_i X)^2\le r_h\}}) = 0 \quad a.s.$$

In fact $\{(\Delta_i L)^2 \le 4r_h, \Delta_i N = 0\} - \{(\Delta_i X)^2 \le r_h\} = \{(\Delta_i L)^2 \le 4r_h, \Delta_i N = 0, (\Delta_i X)^2 > r_h\} \subset$

$$\{|\Delta_i L| \le 2\sqrt{r_h}, \Delta_i N = 0, |\Delta_i X_0| + |\Delta_i M| > \sqrt{r_h}\} \subset \{|\Delta_i X_0| > \sqrt{r_h}/2\} \cup \{|\Delta_i M| \le 2\sqrt{r_h}, |\Delta_i M| > \sqrt{r_h}/2\}.$$

Since, by (18), a.s. for sufficiently small h $\sum_i |a_{hi}|I_{\{|\Delta_i X_0|>\sqrt{r_h}/2\}} = 0$, we a.s. have

$$\lim_{h\to 0}\sum_i |a_{hi}|(I_{\{(\Delta_i L)^2\le 4r_h,\Delta_i N=0\}} - I_{\{(\Delta_i X)^2\le r_h\}}) \le \lim_{h\to 0}\sum_i |a_{hi}|I_{\{|\Delta_i M|\le 2\sqrt{r_h},|\Delta_i M|>\sqrt{r_h}/2\}} \ ,$$

however, by Remark 2.2, as $h \to 0$

$$E[\sum_i |a_{hi}|I_{\{|\Delta_i M|\le 2\sqrt{r_h},|\Delta_i M|>\sqrt{r_h}/2\}}] \le O(r_h)nP\{|\Delta_i M| \le 2\sqrt{r_h}, |\Delta_i M| > \sqrt{r_h}/2\} \le$$

$$O(r_h)nP\{|\Delta_i M|I_{\{|\Delta_i M|\le 2\sqrt{r_h}\}} > \sqrt{r_h}/2\} \le O(r_h)n\frac{E[(\Delta_i M)^2 I_{\{|\Delta_i M|\le 2\sqrt{r_h}\}}]}{r_h} = O(r_h)n\frac{h\eta^2(2r_h^{\frac14})}{r_h} \to 0,$$

$\square$

Lemma 6.3. *Under the assumptions of theorem 2.5, for all* $\alpha \in [0, 2[$

$$\begin{aligned} \textstyle\sum_{i=1}^n (\Delta_i M)^2 I_{\{(\Delta_i M)^2\le r_h/16\}} - o_P(h^{1-\alpha/2}) &\le \textstyle\sum_{i=1}^n (\Delta_i M)^2 I_{\{(\Delta_i X)^2\le r_h,(\Delta_i M)^2\le 4r_h\}} \\ &\le \textstyle\sum_{i=1}^n (\Delta_i M)^2 I_{\{(\Delta_i M)^2\le 9r_h/4\}} + o_P(h^{1-\alpha/2}) \quad a.s. \end{aligned} \tag{37}$$

Proof. Let us first deal with $\sum_{i=1}^n (\Delta_i M)^2 I_{\{(\Delta_i X)^2>r_h,(\Delta_i M)^2\le 4r_h\}}$.

As in (24), on $\{(\Delta_i X)^2 > r_h\}$ we have either $|\Delta_i J| > \sqrt{r_h}/4$ or $|\Delta_i M| > \sqrt{r_h}/4$, so

$$\sum_{i=1}^n (\Delta_i M)^2 I_{\{(\Delta_i X)^2>r_h,(\Delta_i M)^2\le 4r_h\}} \le$$

$$\sum_{i=1}^n (\Delta_i M)^2 I_{\{(\Delta_i X)^2>r_h,\Delta_i J\ne 0,(\Delta_i M)^2\le 4r_h\}} + \sum_{i=1}^n (\Delta_i M)^2 I_{\{(\Delta_i X)^2>r_h,(\Delta_i M)^2>\frac{r_h}{16},(\Delta_i M)^2\le 4r_h\}}.$$

However

$$E\Big[\frac{\sum_{i=1}^n (\Delta_i M)^2 I_{\{(\Delta_i M)^2\le 4r_h,\Delta_i N\ne 0\}}}{h^{1-\alpha/2}}\Big] = O\Big(\frac{h\eta^2(r_h^{\frac14})N_T}{h^{1-\alpha/2}}\Big) \to 0,$$

$$\begin{aligned} \text{so} \qquad \sum_{i=1}^n (\Delta_i M)^2 I_{\{(\Delta_i X)^2>r_h,(\Delta_i M)^2\le 4r_h\}} &\le o_P(h^{1-\alpha/2}) + \sum_{i=1}^n (\Delta_i M)^2 I_{\{(\Delta_i M)^2\le 4r_h,(\Delta_i M)^2>r_h/16\}} \\ &= o_P(h^{1-\alpha/2}) + \sum_{i=1}^n (\Delta_i M)^2 I_{\{(\Delta_i M)^2\le 4r_h\}} - \sum_{i=1}^n (\Delta_i M)^2 I_{\{(\Delta_i M)^2\le 4r_h,(\Delta_i M)^2\le r_h/16\}} \\ &= o_P(h^{1-\alpha/2}) + \sum_{i=1}^n (\Delta_i M)^2 I_{\{(\Delta_i M)^2\le 4r_h\}} - \sum_{i=1}^n (\Delta_i M)^2 I_{\{(\Delta_i M)^2\le r_h/16\}}. \end{aligned} \tag{38}$$

Consider now $\sum_{i=1}^n (\Delta_i M)^2 I_{\{(\Delta_i M)^2 \le 4r_h, (\Delta_i M)^2 > 9r_h/4\}}$: on $\{2\sqrt{r_h} \ge |\Delta_i M| > \frac{3}{2}\sqrt{r_h}\}$ either $\Delta_i N \neq 0$, in which case

$$\frac{\sum_{i=1}^n (\Delta_i M)^2 I_{\{(\Delta_i M)^2 \le 4r_h, \Delta_i N \neq 0\}}}{h^{1-\alpha/2}} \xrightarrow{P} 0$$

as before, or $\Delta_i N = 0$, in which case $|\Delta_i X| > |\Delta_i M| - |\Delta_i X_0| > \frac{3}{2}\sqrt{r_h} - \frac{1}{2}\sqrt{r_h} = \sqrt{r_h}$ so

$$\sum_{i=1}^n (\Delta_i M)^2 I_{\{(\Delta_i X)^2 > r_h, (\Delta_i M)^2 \le 4r_h\}} + o_P(h^{1-\alpha/2}) \ge \sum_{i=1}^n (\Delta_i M)^2 I_{\{(\Delta_i M)^2 \le 4r_h, (\Delta_i M)^2 > 9r_h/4\}}$$

therefore

$$\sum_{i=1}^n (\Delta_i M)^2 I_{\{(\Delta_i X)^2 > r_h, (\Delta_i M)^2 \le 4r_h\}} \ge -o_P(h^{1-\alpha/2}) + \sum_{i=1}^n (\Delta_i M)^2 I_{\{(\Delta_i M)^2 \le 4r_h\}} - \sum_{i=1}^n (\Delta_i M)^2 I_{\{(\Delta_i M)^2 \le 9r_h/4\}}. \tag{39}$$

Combining now (38) and (39), we obtain (37) since

$$\sum_{i=1}^n (\Delta_i M)^2 I_{\{(\Delta_i X)^2 \le r_h, (\Delta_i M)^2 \le 4r_h\}} = \sum_{i=1}^n (\Delta_i M)^2 I_{\{(\Delta_i M)^2 \le 4r_h\}} - \sum_{i=1}^n (\Delta_i M)^2 I_{\{(\Delta_i X)^2 > r_h, (\Delta_i M)^2 \le 4r_h\}} \quad \square$$

Proof of theorem 2.5. Note that under $\beta > \frac{1}{2-\alpha/2}$ the assumptions of proposition 2.3 are satisfied. Since $X = X_1 + M$, we decompose

$$\frac{\hat{IV}_h - IV}{\sqrt{2h\hat{IQ}_h}} = \frac{\sum_{i=1}^n (\Delta_i X)^2 I_{\{(\Delta_i X)^2 \le r_h\}} - IV}{\sqrt{2h}\sqrt{\frac{\sum_i (\Delta_i X)^4 I_{\{(\Delta_i X)^2 \le r_h\}}}{3h}}} = \frac{\sum_{i=1}^n (\Delta_i X_1)^2 I_{\{(\Delta_i X_1)^2 \le 4r_h\}} - IV}{\sqrt{\frac{2}{3}\sum_i (\Delta_i X)^4 I_{\{(\Delta_i X)^2 \le r_h\}}}} + \tag{40}$$

$$+\frac{\sqrt{2hIQ}}{\sqrt{\frac{2}{3}\sum_i (\Delta_i X)^4 I_{\{(\Delta_i X)^2 \le r_h\}}}}\Bigg[\frac{\sum_{i=1}^n (\Delta_i X_1)^2 \left(I_{\{(\Delta_i X)^2 \le r_h\}} - I_{\{(\Delta_i X_1)^2 \le 4r_h\}}\right)}{\sqrt{2hIQ}} \tag{41}$$

$$+2\frac{\sum_{i=1}^n \Delta_i X_1 \Delta_i M I_{\{(\Delta_i X)^2 \le r_h\}}}{\sqrt{2hIQ}} + \frac{\sum_{i=1}^n (\Delta_i M)^2 I_{\{(\Delta_i X)^2 \le r_h\}}}{\sqrt{2hIQ}}\Bigg] := \sum_{j=1}^4 I_j(h).$$

The proof [20, Thm 2] shows that $I_1(h)$ converges stably in law to a standard Gaussian random variable. To show that the remaining terms either tend to zero or to infinity, we can assume w.l.g. that both a and σ are bounded a.s. If $(\Delta_i X)^2 \le r_h$ and $(\Delta_i X_1)^2 > 4r_h$ then $|\Delta_i M| > \sqrt{r_h}$ and $\Delta_i N \neq 0$, exactly as for $I_2(h)$ in Proposition 2.3. It follows that

$$P\Big\{\frac{\sum_{i=1}^n (\Delta_i X_1)^2 I_{\{(\Delta_i X)^2 \le r_h, (\Delta_i X_1)^2 > 4r_h\}}}{\sqrt{2hIQ}} \neq 0\Big\} \le nP\{\Delta_i N \neq 0, |\Delta_i M| > \sqrt{r_h}\} \to 0$$

by (19). The main factor of the remaining part of $I_2(h)$ is

$$\frac{\sum_{i=1}^n (\Delta_i X_1)^2 I_{\{(\Delta_i X)^2 > r_h, (\Delta_i X_1)^2 \le 4r_h\}}}{\sqrt{2hIQ}}.$$

We recall that on $\{|\Delta_i X_1| \le 2\sqrt{r_h}\}$ we have $\Delta_i N = 0$, thus $(\Delta_i X_1)^2 = (\Delta_i X_0)^2$. Moreover

$$\frac{\sum_{i=1}^n \left(\int_{t_{i-1}}^{t_i} a_u du\right)^2 I_{\{(\Delta_i X)^2 > r_h, (\Delta_i X_1)^2 \le 4r_h\}}}{\sqrt{2hIQ}} = O_P(\sqrt{h}) \to 0,$$

and, by (20)

$$\frac{1}{\sqrt{2hIQ}}\sum_{i=1}^n \int_{t_{i-1}}^{t_i} a_u du \int_{t_{i-1}}^{t_i} \sigma_u dW_u I_{\{(\Delta_i X)^2 > r_h, (\Delta_i X_1)^2 \le 4r_h\}} \le c\sqrt{h}\sqrt{h\ln\frac{1}{h}}\sum_{i=1}^n I_{\{(\Delta_i X)^2 > r_h\}}$$

$$= O\Big(h^{1-\alpha\beta/2}\sqrt{\ln\frac{1}{h}}\Big) \to 0.$$

Therefore in probability

$$\lim_{h\to 0} I_2(h) = \lim_{h\to 0} -\frac{\sum_{i=1}^n \big(\int_{t_{i-1}}^{t_i} \sigma_u dW_u\big)^2 I_{\{(\Delta_i X)^2>r_h,(\Delta_i X_1)^2\le 4r_h\}}}{\sqrt{2hIQ}}.$$

Now we show that term $I_3(h)/2$ in (41) tends to zero in probability. First recall that $\Delta_i X_1 = \Delta_i X_0 + \Delta_i J$, and within the sum $\sum_{i=1}^n \Delta_i J \Delta_i M I_{\{(\Delta_i X)^2\le r_h\}}/\sqrt{h}$ term i contributes only when $\Delta_i N \neq 0$, in which case we also have $(\Delta_i X_1)^2 > 4r_h$ and thus $|\Delta_i M| > \sqrt{r_h}$ as in (26). That implies

$$P\left\{\frac{\sum_{i=1}^n \Delta_i J \Delta_i M I_{\{(\Delta_i X)^2\le r_h\}}}{\sqrt{2hIQ}} \neq 0\right\} \le nP\{\Delta_i N \neq 0, |\Delta_i M| > \sqrt{r_h}\} \to 0.$$

As for $\frac{\sum_{i=1}^n \Delta_i X_0 \Delta_i M I_{\{(\Delta_i X)^2\le r_h\}}}{\sqrt{h}}$, as in the proof of lemma 6.2, we have

$$\frac{\sum_{i=1}^n \Delta_i X_0 \Delta_i M I_{\{(\Delta_i X)^2\le r_h\}}}{\sqrt{h}} = \frac{\sum_{i=1}^n \Delta_i X_0 \Delta_i M I_{\{(\Delta_i X)^2\le r_h,(\Delta_i L)^2\le 4r_h\}}}{\sqrt{h}}, \tag{42}$$

however since both $P\left\{\frac{1}{\sqrt{h}}\sum_{i=1}^n \Delta_i X_0 \Delta_i M I_{\{(\Delta_i X)^2\le r_h,(\Delta_i L)^2\le 4r_h,\Delta_i N\neq 0\}} \neq 0\right\}$ and $P\left\{\frac{1}{\sqrt{h}}\sum_{i=1}^n \Delta_i X_0 \cdot\ \Delta_i M I_{\{(\Delta_i X)^2\le r_h,(\right.$
are dominated by $nP\{\Delta_i N \neq 0, (\Delta_i M)^2 > cr_h\} \to 0$ we have

$$\lim_h \frac{1}{\sqrt{h}}\sum_{i=1}^n \Delta_i X_0 \Delta_i M I_{\{(\Delta_i X)^2\le r_h,(\Delta_i L)^2\le 4r_h\}} = \lim_h \frac{1}{\sqrt{h}}\sum_{i=1}^n \Delta_i X_0 \Delta_i M I_{\{(\Delta_i X)^2\le r_h,(\Delta_i L)^2\le 4r_h,\Delta_i N=0\}}$$

$$= \lim_h \frac{1}{\sqrt{h}}\sum_{i=1}^n \Delta_i X_0 \Delta_i M I_{\{(\Delta_i X)^2\le r_h,(\Delta_i M)^2\le 4r_h,\Delta_i N=0\}} = \lim_h \frac{1}{\sqrt{h}}\sum_{i=1}^n \Delta_i X_0 \Delta_i M I_{\{(\Delta_i X)^2\le r_h,(\Delta_i M)^2\le 4r_h\}}.$$

Moreover by the Cauchy-Schwartz inequality, we have

$$\frac{\sum_{i=1}^n \int_{t_{i-1}}^{t_i} a_u du \Delta_i M I_{\{(\Delta_i X)^2\le r_h,(\Delta_i M)^2\le 4r_h\}}}{\sqrt{h}} \le \frac{\sqrt{\sum_{i=1}^n (\int_{t_{i-1}}^{t_i} a_u du)^2}}{\sqrt{h}} \sqrt{\sum_{i=1}^n (\Delta_i M)^2 I_{\{(\Delta_i M)^2\le 4r_h\}}}$$

$$\le c\sqrt{\sum_{i=1}^n (\Delta_i M)^2 I_{\{(\Delta_i M)^2\le 4r_h\}}}, \tag{43}$$

which tends to zero in probability, since by remark 2.2 as $h \to 0$

$$E\left[\sum_{i=1}^n (\Delta_i M)^2 I_{\{(\Delta_i M)^2\le 4r_h\}}\right] = \int_0^T \int_{|x|\le 2r_h^{1/4}} x^2 \nu(dx) = T\eta^2(r_h^{1/4}) \to 0. \tag{44}$$

On the other hand

$$\frac{1}{\sqrt{h}}\sum_{i=1}^n \big(\int_{t_{i-1}}^{t_i} \sigma_u dW_u\big)\Delta_i M I_{\{(\Delta_i X)^2\le r_h,(\Delta_i M)^2\le 4r_h\}} = \frac{1}{\sqrt{h}}\sum_{i=1}^n \big(\int_{t_{i-1}}^{t_i} \sigma_u dW_u\big)\Delta_i M^{(h)} I_{\{(\Delta_i X)^2\le r_h,(\Delta_i M)^2\le 4r_h\}}$$

$$-\frac{1}{\sqrt{h}}\sum_{i=1}^n \big(\int_{t_{i-1}}^{t_i} \sigma_u dW_u\big) hd(2\sqrt[4]{r_h}) I_{\{(\Delta_i X)^2\le r_h,(\Delta_i M)^2\le 4r_h\}}, \tag{45}$$

where, using that $\int_{t_{i-1}}^{t_i} \sigma_u dW_u$ and $\Delta_i M^{(h)}$ are martingale increments with zero quadratic covariation, the $L^1(\Omega)$ norm of the first right-hand term is bounded by $\sqrt{E\Big[\frac{\sum_{i=1}^n \big(\int_{t_{i-1}}^{t_i} \sigma_u dW_u\big)^2 (\Delta_i M^{(h)})^2}{h}\Big]}$ which is dealt similarly as in (31) and tends to zero. Moreover

$$E\Big[\frac{1}{\sqrt{h}}\sum_{i=1}^n \big(\int_{t_{i-1}}^{t_i} \sigma_u dW_u\big) hd(2\sqrt[4]{r_h}) I_{\{(\Delta_i X)^2\le r_h,(\Delta_i M)^2\le 4r_h\}}\Big] =$$

$$c\sqrt{h}\Big[I_{\alpha\neq 1}(c+r_h^{\frac{1-\alpha}{4}})+I_{\alpha=1}\ln\frac{1}{r_h^{1/4}}\Big]E\Big[\sum_{i=1}^n\big(\int_{t_{i-1}}^{t_i}\sigma_u dW_u\big)I_{\{(\Delta_i X)^2\leq r_h,(\Delta_i M)^2\leq 4r_h\}}\Big]\leq$$

$$c\sqrt{h}\Big[I_{\alpha\neq 1}(c+r_h^{\frac{1-\alpha}{4}})+I_{\alpha=1}\ln\frac{1}{r_h^{1/4}}\Big]\sqrt{E\Big[\sum_{i=1}^n\big(\int_{t_{i-1}}^{t_i}\sigma_u dW_u\big)^2\Big]}\to 0.$$

Using that $\frac{\sqrt{2h\ IQ}}{\sqrt{2/3\sum_i(\Delta_i X)^4 I_{\{(\Delta_i X)^2\leq r_h\}}}}$ tends to 1 in probability, doing for $I_4(h)$ as in (42) and putting together the simplified version of $I_2(h)$ we obtain that $(\hat{IV}_h-IV)/\sqrt{2h\hat{IQ}_h}$ is the sum of a term which converges in distribution to a $N(0,1)$ r.v. plus a negligible term and a remainder

$$-\frac{\sum_{i=1}^n\big(\int_{t_{i-1}}^{t_i}\sigma_u dW_u\big)^2 I_{\{(\Delta_i X)^2>r_h,(\Delta_i X_1)^2\leq 4r_h\}}}{\sqrt{2hIQ}}+\frac{\sum_{i=1}^n(\Delta_i M)^2 I_{\{(\Delta_i X)^2\leq r_h,(\Delta_i M)^2\leq 4r_h\}}}{\sqrt{2hIQ}}. \tag{46}$$

a) if $\alpha<1$, the first term of (46) is negligible with respect to $\frac{r_h^{1-\alpha/2}}{\sqrt{2hIQ}}$, in fact

$$\frac{\sum_{i=1}^n\big(\int_{t_{i-1}}^{t_i}\sigma_u dW_u\big)^2 I_{\{(\Delta_i X)^2>r_h,(\Delta_i X_1)^2\leq 4r_h\}}}{r_h^{1-\alpha/2}}\leq\frac{\sum_{i=1}^n h\ln\frac{1}{h}I_{\{(\Delta_i X)^2>r_h\}}}{r_h^{1-\alpha/2}}$$

where

$$E\left[\frac{\sum_{i=1}^n h\ln\frac{1}{h}I_{\{(\Delta_i X)^2>r_h\}}}{r_h^{1-\alpha/2}}\right]\leq h^{1-\beta}\ln\frac{1}{h}\to 0.$$

Therefore (46) can be written as

$$\frac{r_h^{1-\alpha/2}}{\sqrt{2hIQ}}\Big[o_P(1)+\frac{\sum_{i=1}^n(\Delta_i M)^2 I_{\{(\Delta_i X)^2\leq r_h,(\Delta_i M)^2\leq 4r_h\}}}{r_h^{1-\alpha/2}}\Big]. \tag{47}$$

Using (37), lemma 2.1 i) and theorem 2.4 we reach

$$\frac{\sum_{i=1}^n(\Delta_i M)^2 I_{\{(\Delta_i X)^2\leq r_h,(\Delta_i M)^2\leq 4r_h\}}}{r_h^{1-\alpha/2}}\leq\frac{\sum_{i=1}^n(\Delta_i M)^2 I_{\{(\Delta_i M)^2\leq 9r_h/4\}}+o_P(h^{1-\alpha/2})}{r_h^{1-\alpha/2}}\sim$$

$$\frac{\sum_{i=1}^n(\Delta_i M)^2 I_{\{(\Delta_i M)^2\leq 9r_h/4\}}}{r_h^{1-\alpha/2}}\leq\frac{\sum_i\Big(\int_{t_{i-1}}^{t_i}\int_{|x|\leq 3\sqrt{r_h}/2}x\tilde\mu(dx,dt)-h\int_{3\sqrt{r_h}/2<|x|\leq 1}x\nu(dx)\Big)^2}{r_h^{1-\alpha/2}}$$

$$=R_h+Tc+Tc\Big(\frac{h}{r_h}\Big)^{\frac{\alpha}{2}}h^{1-\frac{\alpha}{2}}\stackrel{P}{\to}Tc$$

where term R_h has variance $\sim cr_h^{\alpha/2}\to 0$ so converges to zero in probability. Since $\frac{r_h^{1-\alpha/2}}{\sqrt{h}}\to 0$, we reach

$$\frac{\hat{IV}_h-IV}{\sqrt{2h\hat{IQ}_h}}\stackrel{st}{\to}N(0,1).$$

b) If $\alpha>1$ define $R_t:=\sum_{s\leq t}I_{\{|\Delta M_s|>\sqrt{h}\}}$ then by (37) last term (times $\sqrt{2IQ}$) in (46) dominates

$$\frac{\sum_{i=1}^n(\Delta_i M)^2 I_{\{(\Delta_i M)^2\leq r_h/16\}}-o_P(h^{1-\alpha/2})}{\sqrt{h}}=$$

$$\frac{1}{\sqrt{h}}\left[\sum_i(\Delta_i M)^2 I_{\{\Delta_i R=0\}}+\sum_i(\Delta_i M)^2\Big[I_{\{(\Delta_i M)^2\leq r_h/16\}}-I_{\{\Delta_i R=0\}}\Big]\right]-o_P(h^{\frac{1}{2}-\frac{\alpha}{2}})\geq$$

$$-o_P(h^{\frac{1}{2}-\frac{\alpha}{2}})+\frac{\sum_i\Big(\int_{t_{i-1}}^{t_i}\int_{|x|\leq\sqrt{h}}x\tilde\mu(dx,dt)-h\int_{\sqrt{h}<|x|\leq 1}x\nu(dx)\Big)^2}{\sqrt{h}}-\frac{\sum_i(\Delta_i M)^2 I_{\{(\Delta_i M)^2>r_h/16,\Delta_i R=0\}}}{\sqrt{h}}. \tag{48}$$

First

$$\sum_i (\Delta_i M)^2 I_{\{(\Delta_i M)^2>\frac{r_h}{16},\Delta_i R=0\}} = \sum_i \left[\Delta_i[M] + 2\int_{t_{i-1}}^{t_i}(M_{s_-}-M_{t_{i-1}})dM_s\right] I_{\{(\Delta_i M)^2>\frac{r_h}{16},\Delta_i R=0\}}:$$

As in Lemma 2.1 the sum of the right terms within brackets is of order $u_n=(n/\log n)^{1/\alpha}$, so that

$$\frac{\sum_i |\int_{t_{i-1}}^{t_i}(M_{s_-}-M_{t_{i-1}})dM_s|}{\sqrt{h}} = \frac{u_n\sum_i |\int_{t_{i-1}}^{t_i}(M_{s_-}-M_{t_{i-1}})dM_s|}{u_n\sqrt{h}} \stackrel{P}{\to} 0,$$

since $u_n\sqrt{h}=\left(\frac{n^{(1-\frac{\alpha}{2})}}{\log n}\right)^{\frac{1}{\alpha}}\to+\infty$. Theorem 2.4 applied with $u=1/2$ yields that with $\varepsilon=h^{\frac{1}{2}}$

$$\sum_i\left(\int_{t_{i-1}}^{t_i}\int_{|x|\le\sqrt{h}} x\tilde{\mu}(dx,dt) - h\int_{\sqrt{h}<|x|\le1}x\nu(dx)\right)^2 = \varepsilon^{2-\frac{\alpha}{2}}Y_h + Tc\varepsilon^{2-\alpha}+Tch\varepsilon^{2-2\alpha}.$$

where $var(Y_h)\to 1$. Therefore in (48) we remain with

$$h^{\frac{1}{2}-\frac{\alpha}{2}}\left[-o_P(1)+h^{\frac{\alpha}{4}}Y+Tc+Tch^{1-\frac{\alpha}{2}}-\frac{\sum_i\Delta_i[M]I_{\{(\Delta_i M)^2>r_h/16,\Delta_i R=0\}}}{h^{1-\frac{\alpha}{2}}}\right]\stackrel{a.s.}{\to}+\infty,$$

where the divergence is due to the facts that $h^{\frac{1}{2}-\frac{\alpha}{2}}\to+\infty$ while $\frac{\sum_i\Delta_i[M]I_{\{(\Delta_i M)^2>r_h/16,\Delta_i R=0\}}}{h^{1-\frac{\alpha}{2}}}$ tends to zero in probability since its expected value is dominated by

$$\frac{n}{h^{1-\frac{\alpha}{2}}}E^{\frac{1}{2}}\left[\left(\Delta_i[M]I_{\{\Delta_i R=0\}}\right)^2\right]P^{\frac{1}{2}}\{(\Delta_i M)^2>r_h/16,\Delta_i R=0\}$$

$$\le\frac{n}{h^{1-\frac{\alpha}{2}}}\left(h\int_{|x|\le\sqrt{h}}x^4\nu(dx)\right)^{\frac{1}{2}}h^{(2-\frac{\alpha}{2}-\beta)\frac{1}{2}}=h^{\frac{1-\beta}{2}}\to 0,$$

having used that

$$P\{(\Delta_i M)^2>r_h,\Delta_i R=0\}=P\{(\Delta_i M)^2 I_{\{\Delta_i R=0\}}>r_h\}\le\frac{E[(\Delta_i M)^2 I_{\{\Delta_i R=0\}}]}{r_h} \tag{49}$$

$$=\frac{h\int_{|x|\le\sqrt{h}}x^2\nu(dx)}{r_h}=h^{2-\frac{\alpha}{2}-\beta}.$$

On the other hand the first term in (46) is negligible with respect to $h^{\frac{1}{2}-\frac{\alpha}{2}}$ (the speed of divergence of $\left(\sum_{i=1}^n(\Delta_i M)^2 I_{\{(\Delta_i M)^2\le r_h/16\}}-o_P(h^{1-\frac{\alpha}{2}})\right)/\sqrt{h}$) because

$$\frac{\sum_{i=1}^n\left(\int_{t_{i-1}}^{t_i}\sigma_u dW_u\right)^2 I_{\{(\Delta_i X)^2>r_h,(\Delta_i X_1)^2\le 4r_h\}}}{\sqrt{h}h^{\frac{1}{2}-\frac{\alpha}{2}}}\le\frac{h\log\frac{1}{h}h^{-\frac{\alpha\beta}{2}}}{h^{1-\frac{\alpha}{2}}}=h^{\frac{\alpha}{2}(1-\beta)}\log\frac{1}{h}\to 0,$$

therefore (46) explodes to $+\infty$. Finally if $\alpha=1$ in (46) the first term is negligible, as

$$\frac{\sum_{i=1}^n\left(\int_{t_{i-1}}^{t_i}\sigma_u dW_u\right)^2 I_{\{(\Delta_i X)^2>r_h,(\Delta_i X_1)^2\le 4r_h\}}}{\sqrt{h}}=O_p(h^{\frac{1-\beta}{2}}\log\frac{1}{h})\to 0.$$

For the second term we take a $\delta>0$ such that $2/3<\beta+\delta<1$, we choose $\varepsilon=h^{\frac{\beta+\delta}{2}}$ and we make the same steps as to reach (48) for $\alpha>1$, but we consider $\tilde{R}_t=\sum_{s\le t}I_{\{|\Delta M_s|>\varepsilon\}}$ in place of R_t. Using also theorem 2.4 we obtain that the second term in (46) dominates

$$\frac{Y_h\varepsilon^{\frac{3}{2}}}{\sqrt{2hIQ}}+\frac{\varepsilon}{\sqrt{2hIQ}}-\frac{\sum\Delta_i[M]I_{\{(\Delta_i M)^2>\frac{r_h}{16},\Delta_i\tilde{R}=0\}}}{\sqrt{2hIQ}}-\frac{2\sum_{i=1}^n\int_{t_{i-1}}^{t_i}(M_{s_-}-M_{t_{i-1}})dM_s I_{\{(\Delta_i M)^2>\frac{r_h}{16},\Delta_i\tilde{R}=0\}}}{\sqrt{2hIQ}},$$

where the variance of Y_h tends to 1, so $Y_h \varepsilon^{\frac{3}{2}}/\sqrt{h}$ tends to zero in probability. The second term tends to $+\infty$ at rate $\varepsilon/\sqrt{h}$. The third term is negligible with respect to $\varepsilon/\sqrt{h}$: applying (49) with $\tilde{R}$ in place of R and the Cauchy-Schwarz inequality we get

$$E\Big[\frac{1}{\varepsilon}\sum\int_{t_{i-1}}^{t_i}\int_{|x|\le 1} x^2\mu(dx,dt) I_{\{(\Delta_i M)^2>\frac{r_h}{16},\Delta_i\tilde{R}=0\}}\Big] = O(h^{\frac{\delta}{2}}) \to 0;$$

finally the last term is also negligible since the speed of convergence to zero of the numerator is $u_n = n/\log^2 n$ (as in the proof of lemma 2.1) and $u_n\sqrt{h}\to +\infty$. So even for $\alpha = 1$ the normalized bias $(\hat{IV}_h - IV)/\sqrt{2h\hat{IQ}_h}$ diverges to $+\infty$. □

Proof of proposition 3.1. As in lemma 6.2 with $\sqrt{r_h}$ in place of r_h as bound for the $\max_{i=1..n}|a_{ni}|$, using that $\alpha < 1$ and applying lemma 2.1 i), we reach that $\hat{H}_h$ has the same limit in probability as

$$X_T - \sum_{i=1}^{n}(\Delta_i X_0 + \Delta_i M) I_{\{\Delta_i N=0,(\Delta_i M)^2\le r_h\}},$$

when $h\to 0$. Moreover, since a.s. $N_T < \infty$ and $\sum_{i=1}^n \Delta_i X_0 I_{\{(\Delta_i M)^2>r_h\}}) = O_P(h^{(1-\alpha\beta)/2}\sqrt{\log(1/h)}) \to 0$, taking $\tilde{R}_t = \sum_{s\le t} I_{\{|\Delta M_s|>\sqrt{r_h}\}}$, the above term has limit in probability equal to

$$X_T - \lim_h \sum_{i=1}^{n}(\Delta_i X_0 + \Delta_i M I_{\{(\Delta_i M)^2\le r_h\}}) = X_T - X_{0T} - \lim_h \Big[\sum_{i=1}^{n}\int_{t_{i-1}}^{t_i}\int_{|x|\le\sqrt{r_h}} x\tilde{\mu}(dx,dt)$$

$$-T\int_{\sqrt{r_h}<|x|\le 1} x\nu(dx)\Big] - \lim_h \sum_i \Delta_i M(I_{\{(\Delta_i M)^2\le r_h\}} - I_{\{\Delta_i\tilde{R}=0\}}):$$

using that $P\{\Delta_i\tilde{R}\ge 1\} = O(h^{1-\alpha\beta/2})$ as after (25) we reach $\sum_i \Delta_i M I_{\{(\Delta_i M)^2\le r_h,\Delta_i\tilde{R}\ge 1\}} = O_P(h^{(1-\alpha)\beta/2}) \to 0$; using Holder inequality with exponents $p = q = 2$ we have $\sum_i \Delta_i M I_{\{(\Delta_i M)^2>r_h,\Delta_i\tilde{R}=0\}} = O_P(\ r_h^{(1-\alpha)\beta/2}) \to 0$; finally $\int_0^T\int_{|x|\le\sqrt{r_h}} x\tilde{\mu}(dx,dt) \xrightarrow{L^2} 0$ and $\int_{\sqrt{r_h}<|x|\le 1} x\nu(dx) \to m$, so that $\hat{H}_{h,T} \xrightarrow{P} J_T + mT$. □

References

[1] Aït-Sahalia, Y. (2004). Disentangling volatility from jumps, *Journal of Financial Economics*, 74, 487-528.

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

[3] Aït-Sahalia, Y. and Jacod, J. (2007). Estimating the Degree of Activity of Jumps in High Frequency Data, forthcoming in the *Annals of Statistics*, 37, 2202–2244.

[4] Aït-Sahalia, Y. and Jacod, J. (2008). Testing whether jumps have finite or infinite activity. Working paper

[5] Alili, L. and Kyprianou, A. (2005). Some remarks on first passage of Lévy processes, the American put and pasting principles, *Annals of Applied Probability*, 15, pp. 2062–2080.

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

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

[8] Carr, P., Geman, H., Madan, D., Yor, M. (2002). The Fine Structure of Asset Returns: An Empirical Investigation, *Journal of Business*, **75**, No. 2, 305-332.

[9] Carr, P. and Wu, L.R. (2003). What type of process underlies options? A simple robust test, *Journal of Finance*, **LVIII**, No. 6, 2581–2610, December 2003.

[10] Cont, R. and Tankov, P. (2004). *Financial Modelling wih Jump Processes*, CRC Press.

[11] Ikeda, N., Watanabe, S. (1981). *Stochastic differential equations and diffusion processes*, Kodansha/ North Holland.

[12] Jacod, J. (2004): The Euler scheme for Lévy driven stochastic differential equations: limit theorems. *The Annals of Probability 32*(3A), 1830–1872.

[13] Jacod, J. (2008). Asymptotic properties of realized power variations and associated functions of semimartingales, *Stochastic processes and their applications*, *118*, 517–559.

[14] Jacod, J. (2007). Statistics and high-frequency data. Lecture Notes, SEMSTAT 2007, La Manga (Spain), May 2007.

[15] Jacod, J., Protter, P. (1998). Asymptotic error distributions for the Euler method for stochastic differential equations, *The Annals of Probability* **26**, 267-307.

[16] Kou, S. (2002). A jump-diffusion model for option pricing, *Management Science*, 48, pp. 1086–1101.

[17] Lee, S., and Mykland, P.A. (2008). Jumps in Financial Markets: A New Nonparametric Test and Jump Dynamics, *The Review of Financial Studies*, Vol. 21, Issue 6, pp. 2535-2563

[18] Madan, D.B. (2001). Purely discontinuous asset price processes, in: J. Cvitanic, E. Jouini and M. Musiela (Eds.): *Option Pricing, Interest Rates and Risk Management*, Cambridge University Press, pp 105–153.

[19] Mancini, C. (2004). Estimation of the parameters of jump of a general Poisson-diffusion model, *Scandinavian Actuarial Journal*, 1:42-52.

[20] Mancini, C., (2009). Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scandinavian Journal of Statistics*, vol.36, 270-296

[21] Mancini, C. and Renó, R., (2008). Threshold estimation of Markov models with jumps and interest rate modeling, *Journal of Econometrics*, forthcoming, available on ssrn.com

[22] Merton, R. (1976). Option pricing when underlying stock returns are discontinuous, *J. Financial Economics*, 3, pp. 125–144.

[23] Metivier, M. (1982). *Semimartingales: a course on stochastic processes.* De Gruyter.

[24] Protter, P.E. (2005). *Stochastic integration and differential equations*, Springer.

[25] Todorov, V., Tauchen, G. (2010). Activity Signature Functions for High-Frequency Data Analysis, *Journal of Econometrics*,154, pp.125-138

[26] Woerner, J. (2006b). *Analyzing the fine structure of continuous time stochastic processes*, working paper, University of Göttingen.