The Annals of Statistics
2014, Vol. 42, No. 3, 1131–1144
DOI: 10.1214/13-AOS1179

A REMARK ON THE RATES OF CONVERGENCE FOR INTEGRATED VOLATILITY ESTIMATION IN THE PRESENCE OF JUMPS

BY JEAN JACOD[1] AND MARKUS REISS

UPMC (Université Paris-6) and Humboldt-Universität zu Berlin

The optimal rate of convergence of estimators of the integrated volatility, for a discontinuous Itô semimartingale sampled at regularly spaced times and over a fixed time interval, has been a long-standing problem, at least when the jumps are not summable. In this paper, we study this optimal rate, in the minimax sense and for appropriate "bounded" nonparametric classes of semimartingales. We show that, if the rth powers of the jumps are summable for some $r \in [0, 2)$, the minimax rate is equal to $\min(\sqrt{n}, (n \log n)^{(2-r)/2})$, where n is the number of observations.

1. Introduction. Let X be a one-dimensional Itô semimartingale, which in particular means that its "continuous martingale part" has the form

$$X_t^c = \int_0^t \sigma_s \, dW_s,$$

where W is a standard Brownian motion, and the process σ_t is optional and (locally) squared integrable.

One of the long-standing problems is the estimation of the so-called integrated volatility, say at time 1, that is of the variable $C_1 = \int_0^1 c_s \, ds$, where $c_t = \sigma_t^2$ is the (squared) volatility, on the basis of discrete observations of X. A huge number of papers have been devoted to this question already, in various situations: when the process is continuous (so X is the sum of X^c above, plus possibly a drift term), or when it has jumps; when the process X is "perfectly" observed, or contaminated by noise; when the sampling times are equi-spaced, or when they are irregularly spaced.

Below, we focus on the basic case, where the sampling is at regularly spaced times i/n for $i = 0, \dots, n$, and when $X_{i/n}$ is observed without noise. Even in this simple situation, the question of the "optimal" rate of convergence of estimators toward C_1, as $n \to \infty$, is unanswered so far, when there are jumps which are "too active."

Received September 2012; revised September 2013.

[1] Supported in part by a Humboldt Research Award.

MSC2010 subject classifications. Primary 62C20, 62G20, 62M09; secondary 60H99, 60J75.

Key words and phrases. Semimartingale, volatility, jumps, infinite activity, discrete sampling, high frequency.

More precisely, estimators are known, which converge to C_1 with the rate $\sqrt{n}$, in the continuous case (the realized volatility, or "approximate quadratic variation" at time 1, achieves this rate), and also when X has jumps with a degree of activity, or Blumenthal–Getoor index, less than 1. This rate is optimal (in a minimax sense), for the following reason: if $X = \sigma W$ where $c = \sigma^2$ is a constant, so $C_1 = c$, we have a purely parametric setting for which the local asymptotic normality (LAN) holds with rate $\sqrt{n}$, and the realized volatility is indeed the MLE in this case.

However, when the degree r of jump activity is larger than 1, the best rates found in the literature are of the form $n^{((2-r)/2)-\varepsilon}$ for $\varepsilon > 0$ arbitrarily small (see below for more details). The difficulty comes of course from the essentially nonparametric feature of the problem, since we do not want to specify the law of the process X, apart from the fact that it is an Itô semimartingale, plus possibly some boundedness assumptions on its characteristics. In a purely parametric problem, for example, when X is a Lévy process with a *known* Lévy measure and the only unknown parameters are the variance c of the Gaussian part, and possibly the drift, then again the rate $\sqrt{n}$ is available for estimating c (this rate is achieved by the MLE, under very general circumstances). There has been a considerable interest in providing also nonparametric estimators that converge at rate $\sqrt{n}$, but as we show here, this is in general impossible.

In this paper, a bound for the minimax rate is determined, when the degree of activity is r or smaller [the precise definition of r is given in Assumption (L-r) below, and is slightly different from the usual Blumenthal–Getoor index]. We will see that the best possible rate is $(n \log n)^{(2-r)/2}$ when $r > 1$ (and of course $\sqrt{n}$ when $r \leq 1$). It is interesting to notice that the truncated realized volatility, which achieves the rate $n^{((2-r)/2)-\varepsilon}$ for any prespecified $\varepsilon > 0$ is indeed "almost" rate-optimal.

The paper is organized as follows: in Section 2, we state the assumptions and review some known results. The results of this paper are presented in Section 3, and the proofs are given in the last section.

2. Some known results. We consider a one-dimensional Itô semimartingale X on a filtered space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, which is observed at regularly spaced times $\frac{i}{n}$ for $i = 0, 1, \ldots, n$, over the (fixed) finite interval $[0, 1]$. The characteristics (B, C, ν) where B is the drift, C the integrated volatility and ν the Lévy system of X (see, e.g., Chapter 1 of [4]), thus have the form

$$B_t = \int_0^t b_s \, ds, \qquad C_t = \int_0^t c_s \, ds, \qquad \nu(dt, dx) = dt \, F_t(dx). \tag{2.1}$$

Here, b_t and c_t are optional (or, predictable) processes, with $c_t \geq 0$, and $F_t = F_{\omega,t}(dx)$ is an optional random measure, also called the Lévy measure, which accounts for the jumps of the process.

When X is continuous, the canonical way for estimating C_1 is to use the realized volatility, or approximate quadratic variation at time 1:

$$[X,X]_1^n = \sum_{i=1}^n (\Delta_i^n X)^2 \qquad \text{where } \Delta_i^n X = X_{i/n} - X_{(i-1)/n}, \tag{2.2}$$

which converges in probability to C_1. When further $\int_0^1 b_s^2\,ds$ and $\int_0^1 c_s^2\,ds$ are a.s. finite, we have the stable convergence in law at rate $\sqrt{n}$

$$\sqrt{n}([X,X]_1^n - C_1) \stackrel{\mathcal{L}-s}{\longrightarrow} \mathcal{U} \qquad \text{where } \mathcal{U} = \sqrt{2}\int_0^1 c_s\,dW'_s, \tag{2.3}$$

and where W' is a standard Brownian motion, defined on an extension of $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge 0}, \mathbb{P})$, and which is independent of the σ-field $\mathcal{F}$: see, for example, Theorem 5.4.2 in [4].

When X has jumps, the variables $[X,X]_1^n$ no longer converge to C_1, but to the "full" quadratic variation $[X,X]_1 = C_1 + \sum_{s\le 1}(\Delta X_s)^2$, where $\Delta X_s = X_s - X_{s-}$ denotes the jump size at time s. However, there are two known methods to consistently estimate C_1:

(1) *Truncated realized volatility*. One chooses a sequence v_n of positive truncation levels, typically of the form $v_n \asymp 1/n^{\varpi}$ for some $\varpi \in (0, 1/2)$, and considers

$$\widehat{C}(v_n)_1 = \sum_{i=1}^n (\Delta_i^n X)^2 1_{\{|\Delta_i^n X| \le v_n\}}. \tag{2.4}$$

(2) *Multipower variations*. One chooses an integer $k \ge 2$, and considers

$$\widehat{C}(k,n)_1 = \frac{1}{m_{2/k}^k} \sum_{i=1}^{n-k+1} \prod_{j=0}^{k-1} |\Delta_{i+j}^n X|^{2/k}, \tag{2.5}$$

where $m_p = \mathbb{E}(|U|^p)$ is the pth absolute moment of a standard normal variable U (other versions are possible; one may, e.g., take any product of k increments, with powers adding up to 2).

The first method has been introduced by Mancini in [5], the second one by Barndorff-Nielsen and Shephard in [2]. Both provide estimators which converge in probability to C_1, upon rather weak assumptions on the jumps.

The question of the rate of convergence, though, is still open, and we quickly review the known results, in the case of truncated realized volatility. One needs the following assumption, where r is a number in $[0, 2]$:

ASSUMPTION (L-r). The variables $\sup_{t\le 1}|b_t|$, $\sup_{t\le 1} c_t$ and $\sup_{t\le 1}\int(|x|^r \wedge 1)F_t(dx)$ are almost surely finite.

The larger r is, the weaker Assumption (L-r) is. (L-2) is a very weak assumption for an Itô semimartingale, whereas (L-r) when $r < 2$ puts restrictions on the jump activity, and is slightly stronger than saying that the Blumenthal–Getoor index of X (or, jump activity index) is not bigger than r. In particular, (L-1) is slightly stronger than the property of the jumps to be summable on each finite interval, for example, the jump part to have trajectories of finite variation. Note that a stable process of index $\beta \in (0, 2)$ satisfies (L-r) for all $r > \beta$, but not for $r \leq \beta$.

When (L-r) holds for some $r < 1$, the estimators $\widehat{C}(v_n)_1$ enjoy exactly the same CLT as in (2.3) with $\widehat{C}(v_n)$ in place of $[X, X]_t$, with the same limit, provided we have

$$v_n \asymp 1/n^{\varpi} \qquad \text{with } \frac{1}{4-2r} < \varpi < \frac{1}{2}. \tag{2.6}$$

When (L-r) holds for some $r \geq 1$, the CLT with rate $\sqrt{n}$ no longer holds for $\widehat{C}(v_n)$, but we have when $v_n \asymp 1/n^{\varpi}$ with $\varpi \in (0, 1/r)$:

$$0 < \varpi < \tfrac{1}{2} \qquad \Longrightarrow \qquad n^{\varpi(2-r)}\big(\widehat{C}(v_n)_1 - C_1\big) \xrightarrow{\mathbb{P}} 0 \tag{2.7}$$

(convergence in probability). These results are shown in [3], and Mancini in [6] has proved that when the jumps of X are those of a stable process with index β [so (L-r) holds for all $r > \beta$, but not for $r = \beta$], and when $\beta \geq 1$, the estimator converges exactly at rate $n^{\varpi(2-\beta)}$, in the sense that the sequence $n^{\varpi(2-\beta)}(\widehat{C}(v_n)_1 - C_1)$ converges to a nontrivial limit (in probability, and not in law, in this case): this rate is less than $\sqrt{n}$, as it is in (2.7), and no proper CLT is available in this case.

Turning now to multipowers, we have analogous results, except that one needs stronger assumptions: basically, (L-r) plus the fact that the process c_t is also an Itô semimartingale, and never vanishes: the CLT for $\widehat{C}(k, n)_1$ holds when $r < 1$, with $\sqrt{2}$ replaced by a suitable (bigger) constant depending on k; see [1]. When $r = 1$, Vetter in [7] proves that there is a CLT at rate $\sqrt{n}$ with a nonvanishing bias term. When $r > 1$ nothing is formally known, but the presence of the bias term when $r = 1$ suggests that for $r > 1$ the rate is less than $\sqrt{n}$.

3. The results. We are in a nonparametric setting, in which the process X is not specified [apart from the fact that it satisfies (L-r) for some r], and even the space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ is not specified. The meaning of "optimality" or "rate-optimality" is not a priori clear; and, to begin with, even the quantity to estimate, namely C_1, depends of course on the space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ and on X.

A possible setting is as follows. We consider a family $\mathcal{S}$ of Itô semimartingales satisfying (L-r), each one being defined on its own filtered space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, and the quantity to estimate is the associated integrated volatility $C(X)_1$. Each X in $\mathcal{S}$ takes its values, as a process, in the Skorokhod space $\mathbb{D}^1$ of all càdlàg functions on $\mathbb{R}_+$, and the image by X of the observed σ-field

$\sigma(X_{i/n} : i = 0, \dots, n)$ is the σ-field $\mathcal{D}_n = \sigma(x(i/n) : i = 0, 1, \dots, n)$ of $\mathbb{D}^1$. For any $X \in \mathcal{S}$ we denote by $\mathbb{P}_X^n$ the restriction to $\mathcal{D}_n$ of the law of X.

An estimator at stage n is a $\mathcal{D}_n$-measurable function $X \mapsto \widehat{C}(X)_i^n$ on $\mathbb{D}^1$. We say that a sequence $\widehat{C}_1^n$ of such estimators achieves the uniform rate w_n (with $w_n \to \infty$) on $\mathcal{S}$, for estimating C_1, if the family $w_n(\widehat{C}(X)_1^n - C(X)_1)$ is tight, uniformly in n and in $X \in \mathcal{S}$, that is, $|\widehat{C}(X)_1^n - C(X)_1| = O_P(w_n^{-1})$ uniformly in $X \in \mathcal{S}$.

Of course, if $\mathcal{S}^r$ denotes the set of *all* Itô semimartingales satisfying (L-r), there cannot be any uniform rate because, to begin with, the variables $C(X)_1$ are not uniformly tight when X runs through $\mathcal{S}^r$: we need to restrict our attention to subfamilies of $\mathcal{S}^r$ which are "bounded" in some sense. In view of the formulation of (L-r), it is natural to consider, for any $A > 0$, the class

$$\begin{aligned} &\mathcal{S}_A^r = \text{the set of all Itô semimartingales with} \\ &|b_t| + c_t + \textstyle\int (|x|^r \wedge 1) F_t(dx) \le A \text{ for all } t. \end{aligned} \tag{3.1}$$

We also denote by $\mathcal{S}_A^{r,L}$ the subclass of all Lévy processes belonging to $\mathcal{S}_A^r$.

The main result of this paper is the following theorem.

THEOREM 3.1. *Let $r \in [0, 2)$ and $A > 0$. Any uniform rate w_n for estimating $C(X)_1$, within the class $\mathcal{S}_A^{r,L}$, hence also within the bigger class $\mathcal{S}_A^r$, satisfies* (*up to a multiplicative constant*, *of course*)

$$w_n \le \rho_n := \begin{cases} \sqrt{n}, & \text{if } r \le 1, \\ (n \log n)^{(2-r)/2}, & \text{if } r > 1. \end{cases} \tag{3.2}$$

The results recalled in the previous section show that the truncated estimators $\widehat{C}(v_n)_1$ (which are estimators in the sense specified above) achieve the rate ρ_n when $r < 1$, and at least $n^{\varpi(2-r)}$ when $r \ge 1$, for any X satisfying (L-r). We indeed have (slightly) more:

THEOREM 3.2. *Let $r \in [0, 2)$ and $A > 0$, and take $v_n \asymp 1/n^{\varpi}$. The truncated estimators $\widehat{C}(v_n)_1$ have the uniform rate w_n below, within $\mathcal{S}_A^r$, for estimating $C(X)_1$,*

$$w_n = \begin{cases} \sqrt{n}, & \text{if } r < 1 \text{ and } \dfrac{1}{4 - 2r} \le \varpi < \dfrac{1}{2}, \\ n^{\varpi(2-r)}, & \text{if } r \ge 1 \text{ and } 0 < \varpi < \dfrac{1}{2}. \end{cases} \tag{3.3}$$

When $r < 1$, the truncated estimators $\widehat{C}(v_n)_1$ achieve the uniform rate $\sqrt{n}$, and as seen in the previous section they even enjoy a CLT. When $r \ge 1$ we have the uniform rate $n^{\varpi(2-r)}$, although for any given X we indeed have a "faster" rate, as seen in (2.7); however, this faster rate is not uniform in $X \in \mathcal{S}_A^r$, as could be seen by taking a sequence of Lévy processes with characteristics $(0, 1, G_n)$, with $\int(|x|^r \wedge 1)G_n(dx) \le 1$ (so $X^n \in \mathcal{S}_1^r$ for all n), but such that $\sup_n \int_{\{|x|\le\varepsilon\}} |x|^r G_n(dx)$ does *not* tend to 0 as $\varepsilon \to 0$.

We then conclude that the truncated estimators are uniformly rate optimal when $r < 1$, and otherwise they approach the bound ρ_n, up to $n^{-\varepsilon}$ with $\varepsilon > 0$ arbitrarily small, upon choosing ϖ close enough to $\frac{1}{2}$.

Let us finally show that on the restricted class $\mathcal{S}_A^{r,L}$ of Lévy processes the rate ρ_n of (3.2) can be achieved exactly and thus constitutes the exact minimax optimal rate: this means that for any $r \in [0, 2)$ and any $A > 0$ one can find estimators for $C(X)_1$ enjoying the uniform rate ρ_n. When $r < 1$, we already know this (even for the much larger class $\mathcal{S}_A^r$) by the previous theorem, but for all $r \in [0, 2)$ we can construct estimators with the uniform rate ρ_n on $\mathcal{S}_A^{r,L}$ as follows. For any process X, we consider the empirical characteristic function of the increments, at each stage n (below $u \in \mathbb{R}$):

$$\widehat{\phi}_n(u) = \frac{1}{n}\sum_{j=1}^n e^{iu\Delta_j^n X}. \tag{3.4}$$

Then we set

$$\widehat{C}'(u)_1 = -\frac{2n}{u^2}\left(\log\left|\widehat{\phi}_n(u)\right|\right)1_{\{\widehat{\phi}_n(u)\neq 0\}}. \tag{3.5}$$

THEOREM 3.3. *For all $A > 0$ and $r \in [0, 2)$, the estimators $\widehat{C}'(u_n)_1$ with*

$$u_n = \begin{cases} \sqrt{n}, & \text{if } r \le 1, \\ \sqrt{(r-1)n\log n}/\sqrt{A}, & \text{if } r > 1 \end{cases} \tag{3.6}$$

attain the uniform rate ρ_n for estimating $C(X)_1$, within the class $\mathcal{S}_A^{r,L}$ of Lévy processes.

REMARK 3.4. When $r \le 1$ the estimators $\widehat{C}'(u_n)_1$ are likely to enjoy a Central Limit theorem with rate ρ_n, and with a bias when $r = 1$.

When $r > 1$, and upon examining the proof [see (4.15) and (4.17), e.g.], the estimation error $\widehat{C}'(u_n)_1 - C(X)_1$ is the sum of a random part, which is easily seen to enjoy a CLT with rate $n^{(2-r)/2}\log n$, and a nonrandom part equal to $\Gamma_n = \frac{2\rho_n}{u_n^2}\int(1-\cos(u_n x))F(dx)$, where F is the Lévy measure of the Lévy process X under consideration. It turns out that $|\rho_n\Gamma_n| \le \int(u_n^{-r} \wedge |x|^r)F(dx)$ tends to 0 by Lebesgue's theorem, so, for any given X we indeed have $\rho_n(\widehat{C}'(u_n)_1 - C(X)_1) \to 0$ in probability: this convergence is of course not uniform in $X \in S_A^{r,L}$, otherwise the conclusion of Theorem 3.1 would be violated. Now, depending on whether $\rho_n\Gamma_n(\log n)^{r/2}$ converges or diverges—and both occurrences are possible—we have a CLT with rate $\rho_n(\log n)^{r/2}$, or we have a slower effective rate (still at least ρ_n, of course) with the normalized error converging in probability to a nontrivial limit.

Note that the argumentation is in line with the standard nonparametric error decomposition in a bias and variance part. Our estimator uses that the characteristic exponent for high frequencies u_n separates the Brownian from the jump part according to the ratio u_n^2/u_n^r. We have reliable empirical access to this exponent only up to frequency u_n (otherwise the stochastic error explodes due to a Gaussian deconvolution setting). So far, we do not know whether this spectral approach yields the same optimal rate on the larger class $\mathcal{S}_A^r$.

4. Proofs.

4.1. *Proof of Theorem* 3.1. *The bound* $w_n \le \sqrt{n}$. For proving this bound, it is enough to show that it already holds on the subclass $\mathcal{S}_A^{\mathrm{BM}}$ of all Brownian motions with unit variance $c \le A$ (so $\mathcal{S}_A^{\mathrm{BM}} \subset \mathcal{S}_A^{r,L}$ for all $r \in [0, 2]$).

In this case, and as already mentioned in the Introduction, the increments follow the parametric model $N(0, c/n)^{\otimes n}$ with parameter c running through $[0, A]$, for which the LAN property holds with rate $\sqrt{n}$, and the result follows.

The bound $w_n \le (n \log n)^{(2-r)/2}$ *when* $r \in (0, 2)$. By scaling, if the result holds for one $A > 0$, it holds for all $A > 0$. Hence, in order to find a bound on the uniform rate w_n on $\mathcal{S}_A^{r,L}$, hence a fortiori on $\mathcal{S}_A^r$, it is enough to construct two sequences X^n and Y^n of Lévy processes belonging to $\mathcal{S}_K^{r,L}$ for $n \ge 2$ and some constant K, with the following two properties, where $a_n = (n \log n)^{-(2-r)/2}$:

$$\bullet \text{ we have } C(X^n)_1 = 1 + a_n \text{ and } C(Y^n)_1 = 1 \text{ identically,} \tag{4.1}$$

$$\bullet \text{ the total variation distance between } \mathbb{P}_{X^n}^n \text{ and } \mathbb{P}_{Y^n}^n \text{ tends to 0.} \tag{4.2}$$

Indeed, letting $\widehat{C}(X)_1$ be a sequence of estimators with uniform rate $w_n \to \infty$ on $\mathcal{S}_A^r$ (or, even, on $\mathcal{S}_A^{r,L}$), the two sequences $w_n(\widehat{C}(X^n)_1^n - (1 + a_n))$ and $w_n(\widehat{C}(Y^n)_t^n - 1)$ are tight under $\mathbb{P}_{X^n}^n$ and $\mathbb{P}_{Y^n}^n$, respectively, by (4.1). Then (4.2) implies that the sequence $w_n(\widehat{C}(Y^n)_1^n - (1 + a_n))$ is also tight under $\mathbb{P}_{Y^n}^n$. This is possible only if the sequence $w_n a_n$ is bounded. So $1/a_n$ is an upper bound for any uniform rate on $\mathcal{S}_K^{r,L}$ (up to a multiplicative constant, of course).

The proof of (4.1) and (4.2) is divided into several steps:

(1) We take Lévy processes X^n and Y^n with respective characteristics $(0, 1 + a_n, F_n)$ and $(0, 1, G_n)$, with Lévy measures F_n, G_n satisfying

$$\int (|x|^r \wedge 1) F_n(dx) \le K, \qquad \int (|x|^r \wedge 1) G_n(dx) \le K \tag{4.3}$$

for some constant K (below constants change from line to line, and may depend on r, and are all denoted as K).

By construction, we have (4.1) and $X^n, Y^n \in \mathcal{S}_K^{r,L}$ for a constant K [which may differ from the one in (4.3)], and we need to choose the above measures F_n and G_n in such a way that (4.2) is satisfied.

(2) We take $u_n = 2/a_n^{1/(2-r)} = 2\sqrt{n \log n}$ and the even functions $h_n \in C^2(\mathbb{R})$ defined for $u \geq 0$ by

$$h_n(u) = a_n\big(1_{\{u \leq u_n\}} + e^{-(u-u_n)^3} 1_{\{u > u_n\}}\big).$$

We use the following convention for the Fourier transform, namely $\mathcal{F}g(u) = \int e^{iux} g(x)\,dx$, so the inverse is $\mathcal{F}^{-1}h(x) = \frac{1}{2\pi}\int e^{-iux} h(u)\,du$. We also denote as $f^{(q)}$ the qth derivative of any q-differentiable function f.

Since $h_n^{(q)} \in \mathbb{L}^p$ for all $p \geq 1$ and $q = 0, 1, 2$, we can define $H_n = \mathcal{F}^{-1}h_n$, and we have $h_n^{(q)} = i^q \mathcal{F}^{-1} H_{n,q}$, where $H_{n,q}(x) = x^q H_n(x)$. By the Plancherel identity we deduce

$$\begin{aligned} &\|H_n\|_{\mathbb{L}^2} \leq K a_n u_n^{1/2} \leq K a_n^{(3-2r)/(4-2r)}, \qquad q = 1, 2 \\ &\quad \Rightarrow \quad \|H_{n,q}\|_{\mathbb{L}^2} \leq \big\|h_n^{(q)}\big\|_{\mathbb{L}^2} \leq K a_n. \end{aligned} \tag{4.4}$$

Then the Cauchy–Schwarz inequality applied to the functions $\frac{1}{\sqrt{1+x^2}}$ and $H_n(x)\sqrt{1+x^2}$ yields

$$\int \big|H_n(x)\big|\,dx \leq K\big(1 + a_n^{(3-2r)/(4-2r)}\big) < \infty \tag{4.5}$$

[note that $\|H_n\|_{\mathbb{L}^1}$ is bounded in n when $r \leq 3/2$, but not otherwise; we also have $H_n(0) > a_n u_n \to \infty$]. Therefore, the two measures

$$F_n(dx) = \frac{|H_n(x)|}{x^2}\,dx, \qquad G_n(dx) = F_n(dx) + \frac{H_n(x)}{x^2}\,dx$$

are nonnegative and integrate x^2, hence are Lévy measures.

This construction will satisfy (4.2) mainly because the definition of the two Lévy measures and the constant value of h_n for $|u| \leq u_n$ imply that the difference between the two characteristic exponents vanishes for $|u| \leq u_n$, as we shall prove next.

(3) Splitting the integration domain into the sets $\{|u| \leq u_n\}$ and $\{|u| > u_n\}$ in the integral $\int e^{-iux} h_n(u)\,du$, we get

$$\begin{aligned} \big|H_n(x)\big| &\leq K a_n\left(\frac{|\sin(u_n x)|}{|x|} + 1\right) \\ &\leq K a_n\left(u_n 1_{\{|x| \leq 1/u_n\}} + \frac{1}{|x|} 1_{\{1/u_n < |x| \leq 1\}} + 1_{\{|x| > 1\}}\right). \end{aligned}$$

In turn, the integral $\int \frac{|x|^r \wedge 1}{x^2} |H_n(x)|\,dx$ can be split into integrals on the sets $\{|x| \leq 1/u_n\}$, $\{1/u_n < |x| \leq 1\}$ and $\{|x| > 1\}$, and recalling $1 < r < 2$ we deduce from the above that

$$\int \frac{|x|^r \wedge 1}{x^2} \big|H_n(x)\big|\,dx \leq K a_n\big(u_n^{2-r} + 1\big) \leq K.$$

It follows that the measures F_n and G_n satisfy (4.3), and it remains to prove (4.2).

(4) We denote by ϕ_n and ψ_n the characteristic functions of $X^n_{1/n}$ and $Y^n_{1/n}$, and $\eta_n = \phi_n - \psi_n$. These functions are real (because H_n is an even function) and are given by

$$\phi_n(u) = \exp\Big(-\frac{1}{2n}\big(u^2 + a_n u^2 + 2\widetilde{\phi}_n(u)\big)\Big),$$
$$\psi_n(u) = \exp\Big(-\frac{1}{2n}\big(u^2 + 2\widetilde{\phi}_n(u) + 2\widetilde{\eta}_n(u)\big)\Big),$$

where

$$\widetilde{\phi}_n(u) = \int \big(1 - \cos(ux)\big)\frac{|H_n(x)|}{x^2}\,dx,$$
$$\widetilde{\eta}_n(u) = \int \big(1 - \cos(ux)\big)\frac{H_n(x)}{x^2}\,dx.$$

We proceed to studying $\widetilde{\phi}_n$ and $\widetilde{\eta}_n$. Equation (4.4) applied with $q = 1, 2$ implies that $\widetilde{\phi}_n$ and $\widetilde{\eta}_n$ are twice differentiable. First, we have $\widetilde{\phi}'_n(u) = \int \sin(ux)\frac{|H_n(x)|}{x}\,dx$, hence (4.5) yields

$$(4.6)\qquad \begin{aligned} 0 \le \widetilde{\phi}_n(u) &\le K\big(1 + a_n^{(3-2r)/(4-2r)}\big)u^2, \\ \big|\widetilde{\phi}'_n(u)\big| &\le K\big(1 + a_n^{(3-2r)/(4-2r)}\big)|u|. \end{aligned}$$

Second, $\widetilde{\eta}''_n(u) = \int \cos(ux) H_n(x)\,dx = h_n(u)$, whereas $\widetilde{\eta}(0) = \widetilde{\eta}'_n(0) = 0$, and this yields

$$(4.7)\qquad \begin{aligned} |u| \le u_n \quad &\Rightarrow \quad \widetilde{\eta}_n(u) = \frac{a_n u^2}{2}, \qquad \widetilde{\eta}'_n(u) = a_n u, \\ |u| \ge u_n \quad &\Rightarrow \quad \big|\widetilde{\eta}_n(u)\big| \le \frac{a_n u^2}{2}, \qquad \big|\widetilde{\eta}'_n(u)\big| \le a_n|u|. \end{aligned}$$

(5) Since X^n and Y^n have a nonvanishing Gaussian part, the variables $X^n_{1/n}$ and $Y^n_{1/n}$ have densities, denoted by f_n and g_n, and we set $k_n = f_n - g_n$. Since X^n and Y^n are Lévy processes, the variation distance between $\mathbb{P}^n_{X^n}$ and $\mathbb{P}^n_{Y^n}$ is not more than n times $\int |k_n(x)|\,dx$, and we are thus left to show that $n\int |k_n(x)|\,dx \to 0$.

To check this, we use the same argument as for (4.5): if $k_{n,1}(x) = x k_n(x)$, by the Cauchy–Schwarz inequality we have $\int |k_n(x)|\,dx \le K(\|k_n\|_{\mathbb{L}^2} + \|k_{n,1}\|_{\mathbb{L}^2})$, whereas $\eta_n = \mathcal{F}k_n$ and also, since η_n is twice differentiable, $\eta'_n = i\mathcal{F}k_{n,1}$. By Plancherel identity, it is thus enough to prove that

$$(4.8)\qquad n^2\int \big|\eta_n(u)\big|^2\,du \to 0, \qquad n^2\int \big|\eta'_n(u)\big|^2\,du \to 0.$$

We have $\widetilde{\phi}_n \ge 0$ and $\widetilde{\phi}_n + \widetilde{\eta}_n \ge 0$, which implies $\phi_n(u) \le e^{-u^2/2n}$ and $\psi_n(u) \le e^{-u^2/2n}$, whereas $2\overline{\eta}_n(u) = a_n u^2$ if $|u| \le u_n$ and $|2\overline{\eta}_n(u)| \le a_n u^2$ if $|u| > u_n$

by (4.7). Therefore,

$$
\begin{aligned}
|\eta_n(u)| &= \psi_n(u)\left|1-\frac{\phi_n(u)}{\psi_n(u)}\right| \\
&= \psi_n(u)\left|1-e^{-(a_n u^2-2\widetilde{\eta}_n(u))/(2n)}\right| \le \frac{a_n u^2}{2n} e^{-u^2/2n} 1_{\{|u|>u_n\}},
\end{aligned}
$$

and also, upon using (4.6),

$$
\begin{aligned}
|\eta_n'(u)| &= \frac{1}{n}\left|(u+ua_n+\widetilde{\phi}_n'(u))\phi_n(u)-(u+\widetilde{\phi}_n'(u)+\overline{\eta}_n'(u))\psi_n(u)\right| \\
&\le \frac{1}{n}\left(a_n|u|e^{-u^2/2n}+|\widetilde{\eta}_n'(u)|e^{-u^2/2n}+|u+\widetilde{\phi}_n'(u)||\eta_n(u)|\right)1_{\{|u|>u_n\}} \\
&\le K a_n \frac{|u|}{n} e^{-u^2/2n}\left(1+\left(1+a_n^{(3-2r)/(4-2r)}\right)\frac{u^2}{n}\right)1_{\{|u|>u_n\}}.
\end{aligned}
$$

Now, since $u_n = 2\sqrt{n\log n}$, we have $\int_{\{|u|>u_n\}}(\frac{u^2}{n})^q e^{-u^2/n}\,du \le K\frac{(\log n)^{q-1}}{n^{7/2}}$ for $q = 1,2,3$. Since further $a_n^{(3-2r)/(4-2r)}/\sqrt{n} \to 0$, we deduce

$$
\int |\eta_n(u)|^2\,du \le K\frac{\log n}{n^{7/2}}, \qquad \int |\eta_n'(u)|^2\,du \le K\frac{(\log n)^2}{n^{7-1/2}}.
$$

Then (4.8) follows, and the proof is complete.

4.2. *Proof of Theorem* 3.2. The proof requires several steps:

(1) Any $X \in \mathcal{S}_A^r$ can be written as follows, on some space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge 0}, \mathbb{P})$:

$$
\begin{aligned}
X_t = X_0 &+ \int_0^t b_s\,ds + \int_0^t \sqrt{c_s}\,dW_s \\
&+ \int_0^t\int_E \delta(s,z)1_{\{\|\delta(s,z)\|\le 1\}}(\mu-\nu)(ds,dz) \qquad (4.9)\\
&+ \int_0^t\int_E \delta(s,z)1_{\{\|\delta(s,z)\|>1\}}\mu(ds,dz).
\end{aligned}
$$

Here, b and c are as in (L-r), and W is a standard Brownian motion, and μ is a Poisson random measure on $\mathbb{R}_+\times\mathbb{R}$ with intensity measure $\nu(dt,dz) = dt\otimes dz$, and $\delta = \delta(\omega,t,z)$ is a predictable function on $\Omega\times\mathbb{R}_+\times\mathbb{R}$. The connection between δ and F_t is that $F_{\omega,t}$ is the image of Lebesgue measure by the map $z\mapsto \delta(\omega,t,z)$, restricted to $\mathbb{R}\setminus\{0\}$.

We use the decomposition $X = X'+Y+Z$, where

$$
X_t' = X_0 + \int_0^t b_s\,ds + \int_0^t \sqrt{c_s}\,dW_s
$$

and Y and Z are, respectively, the last two terms in (4.9).

With w_n given by (3.3), it is clearly enough to prove that, for some constant K only depending on A, r, ϖ (as will be all constants K below, changing from line to line), we have

$$\mathbb{E}\big(\big|\widehat{C}(v_n)_1 - C_1\big|\big) \le K/w_n. \tag{4.10}$$

(2) Here, we recall estimates on the increments of X' and Y, the later coming from Lemmas 2.1.5 and 2.1.6 of [4], and where $p > 0$ is arbitrary (the constants K_p below depend on p in addition to r, A). Namely, since $\int_{\{|x|\le 1\}} |x|^r F_t(dx) \le A$, we have uniformly in $s \in [(i-1)/n, i/n]$:

$$\begin{aligned} &\mathbb{E}\big(\big|X'_s - X'_{(i-1)/n}\big|^p\big) \le K_p n^{-p/2}, \\ &\mathbb{E}\big(\big|Y_s - Y_{(i-1)/n}\big|^p\big) \le K n^{-(p/r)\wedge 1}. \end{aligned} \tag{4.11}$$

We will also use the following estimates, which follow from the property $F_t(\{x : |x| > 1\}) \le A$ and from the fact that if $\Delta_i^n Z \ne 0$ there is at least one jump of Z within the interval $(\frac{i-1}{n}, \frac{i}{n}]$ (this estimate follows from Lemma 2.1.7 of [4] applied to the counting process $\sum_{s\le t} 1_{\{\Delta Z_s \ne 0\}}$):

$$\mathbb{P}\big(\Delta_i^n Z \ne 0\big) \le \frac{K}{n}. \tag{4.12}$$

(3) With the notation (2.2), Itô's formula yields $[X', X']_1^n - C_1 = U_n + V_n$, where

$$\begin{aligned} U_n &= \sum_{i=1}^n \mathbb{E}\big(\zeta_i^n | \mathcal{F}_{(i-1)/n}\big), \\ \zeta_i^n &= 2\int_{(i-1)/n}^{i/n} \big(X'_s - X'_{(i-1)/n}\big) b_s\, ds, \\ V_n &= \sum_{i=1}^n \xi_i^n, \\ \xi_i^n &= 2\int_{(i-1)/n}^{i/n} \big(X'_s - X'_{(i-1)/n}\big)\sqrt{c_s}\, dW_s + \zeta_i^n - \mathbb{E}\big(\zeta_i^n | \mathcal{F}_{(i-1)/n}\big). \end{aligned}$$

Equation (4.11) yields

$$\big|\mathbb{E}\big(\zeta_i^n | \mathcal{F}_{(i-1)/n}\big)\big| \le K/n^{3/2}, \qquad \mathbb{E}\big((\xi_i^n)^2\big) + \mathbb{E}\big((\zeta_i^n)^2\big) \le K/n^2,$$

whereas $\mathbb{E}(\xi_i^n | \mathcal{F}_{(i-1)/n}) = 0$. Thus we have $\mathbb{E}(|U_n|) \le K/\sqrt{n}$ and $\mathbb{E}(V_n^2) \le K/n$, implying

$$\mathbb{E}\big(\big|[X', X']_1^n - C_1\big|\big) \le K/\sqrt{n}. \tag{4.13}$$

Therefore, it remains to prove that

$$\mathbb{E}\big(\big|\widehat{C}(v_n)_1 - [X', X']_1^n\big|\big) \le K/w_n. \tag{4.14}$$

(4) Consider the case $r < 1$ first. By Lemma 13.2.6 of [4], applied with $k = 1$ and $F(x) = x^2$, hence $s' = 2$ and $m = s = p' = 1$ and $\theta = 0$ (with the notation of this lemma), we have

$$\mathbb{E}\left(\left|\widehat{C}(v_n)_1 - \sum_{i=1}^{n} (\Delta_i^n X')^2 1_{\{|\Delta_i^n X'| \le v_n\}}\right|\right) \le \frac{K}{n^{(2-r)\varpi}} \le \frac{K}{\sqrt{n}},$$

where the last inequality follows from $\varpi \ge \frac{1}{4-2r}$. On the other hand, (4.11) and Markov inequality yield $\mathbb{E}((\Delta_i^n X')^2 1_{\{|\Delta_i^n X'| > v_n\}}) \le K_p / n^{1+p(1-2\varpi)/2}$ for any $p > 0$, and upon taking $p = \frac{1}{1-2\varpi}$ we obtain

$$\mathbb{E}\left(\left|[X', X']_1^n - \sum_{i=1}^{n} (\Delta_i^n X')^2 1_{\{|\Delta_i^n X'| \le v_n\}}\right|\right) \le \frac{K}{\sqrt{n}}.$$

These two estimates readily give (4.14).

(5) Now we turn to the case $r \ge 1$. One has $\widehat{C}(v_n)_1 - [X', X']_1^n = \sum_{j=1}^3 U(j)_n$, where $U(j)_n = \sum_{i=1}^n \eta(j)_i^n$ and

$$\eta(1)_i^n = (\Delta_i^n X)^2 1_{\{|\Delta_i^n X| \le v_n\}} - (\Delta_i^n X')^2 - 2\Delta_i^n X' \Delta_i^n Y,$$
$$\eta(2)_i^n = 2\mathbb{E}(\Delta_i^n X' \Delta_i^n Y | \mathcal{F}_{(i-1)/n}), \qquad \eta(3)_i^n = 2\Delta_i^n X' \Delta_i^n Y - \eta(2)_i^n.$$

Itô's formula yields, with the notation $\gamma_s = \int_{\{|z| \le 1\}} z^2 F_s(dz)$, and taking advantage of the facts that Y and $\int_0^t \sqrt{c_s}\, dW_s$ are two orthogonal martingales and that $Y_t^2 - \int_0^t \gamma_s\, ds$ is a martingale:

$$\eta(2)_i^n = 2\mathbb{E}\left(\int_{(i-1)/n}^{i/n} (X'_s - X'_{(i-1)/n}) b_s\, ds \Big| \mathcal{F}_{(i-1)/n}\right)$$

$$\begin{aligned}
&\mathbb{E}((\Delta_i^n X' \Delta_i^n Y)^2 | \mathcal{F}_{(i-1)/n}) \\
&\quad = \mathbb{E}\left(\int_{(i-1)/n}^{i/n} (Y_s - Y_{(i-1)/n})^2 c_s\, ds \Big| \mathcal{F}_{(i-1)/n}\right) \\
&\qquad + 2\mathbb{E}\left(\int_{(i-1)/n}^{i/n} (X'_s - X'_{(i-1)/n})(Y_s - Y_{(i-1)/n})^2 b_s\, ds \Big| \mathcal{F}_{(i-1)/n}\right) \\
&\qquad + \mathbb{E}\left(\int_{(i-1)/n}^{i/n} (X'_s - X'_{(i-1)/n})^2 \gamma_s\, ds \Big| \mathcal{F}_{(i-1)/n}\right).
\end{aligned}$$

Then standard estimates and (4.11), plus Hölder's inequality, yield (the first bound is a.s.)

$$|\eta(2)_i^n| \le \frac{K}{n^{3/2}}, \qquad \mathbb{E}((\eta(3)_i^n)^2) \le \frac{K}{n^2}.$$

Since $\mathbb{E}(\eta(3)_i^n | \mathcal{F}_{(i-1)/n}) = 0$, these estimates yield $|U(2)_n| \le K/\sqrt{n}$ and $\mathbb{E}(U(3)_n^2) \le K/n$, hence it is enough to show that $\mathbb{E}(|U(1)_n|) \le K/w_n$.

(6) Recalling $r \geq 1$, the following inequality is easy to check, for $x, y, z \in \mathbb{R}$ and $v \in (0, 1/4]$:

$$\begin{aligned}&\big|(x+y+z)^2 1_{\{|x+y+z| \leq v\}} - x^2 - 2xy\big| \\ &\quad \leq 2v^2 1_{\{z \neq 0\}} + 6|xy| 1_{\{|x|>v/2\}} + 6x^2 1_{\{|x|>v/2\}} + 16 v^{2-r}|y|^r.\end{aligned}$$

It follows that $|\eta(1)_i^n| \leq K \sum_{j=1}^5 \xi(j)_i^n$, where

$$\xi(1)_i^n = v_n^2 1_{\{\Delta_i^n Z \neq 0\}}, \qquad \xi(2)_i^n = \big|\Delta_i^n X' \Delta_i^n Y\big| 1_{\{|\Delta_i^n X'|>v_n/2\}},$$
$$\xi(3)_i^n = \big(\Delta_i^n X'\big)^2 1_{\{|\Delta_i^n X'| \geq v_n/2\}}, \qquad \xi(4)_i^n = v_n^{2-r}\big|\Delta_i^n Y\big|^r.$$

Equation (4.12) yields $\mathbb{E}(\xi(1)_i^n) \leq K/n^{1+2\varpi}$, and (4.11) yields $\mathbb{E}(\xi(4)_i^n) \leq K/n^{1+(2-r)\varpi}$. Another application of (4.11), plus Hölder and Markov inequalities, give us $\mathbb{E}(\xi(j)_i^n) \leq K_p/n^{1+p(1-2\varpi)/2}$ for $j = 2, 3$. Upon taking p large enough, we obtain

$$\mathbb{E}\big(\xi(j)_i^n\big) \leq K/n w_n$$

for $j = 1, 2, 3, 4, 5$. We deduce $\mathbb{E}(|U(1)_n|) \leq K/w_n$, and the proof is complete.

4.3. *Proof of Theorem* 3.3. We let $X \in \mathcal{S}_A^{r,L}$, where $r \in [0, 2)$ and $A > 0$ are given. The characteristic triple of X is (b, c, F) and the characteristic function of $X_{1/n}$ is

$$\phi_n(u) = \exp\bigg(\frac{1}{n}\bigg(iub - \frac{cu^2}{2} + \int \big(e^{iux} - 1 - iux 1_{\{|x| \leq 1\}}\big) F(dx)\bigg)\bigg).$$

Then $|\phi_n(u_n)| = e^{(-1/(2n))(cu_n^2 + \gamma_n)}$, where $\gamma_n = 2\int(1 - \cos(u_n x)) F(dx)$. As soon as n is large enough we have $u_n \geq 1$, hence, since $1 - \cos y \leq 1 \wedge y^2 \leq |y|^r \wedge 1$,

$$\begin{aligned}0 \leq \gamma_n &\leq 2 \int \big(|u_n x|^r \wedge 1\big) F(dx) \leq 2u_n^r \int \big(|x|^r \wedge 1\big) F(dx) \\ &\leq 2u_n^2 \int \big(|x|^r \wedge 1\big) F(dx).\end{aligned}$$

Because $c + \int (|x|^r \wedge 1) F(dx) \leq A$ by hypothesis, and in view of the form of u_n in (3.6), by singling out the two cases $r \leq 1$ and $r > 1$ this implies that, with $\Gamma = e^A$,

$$\frac{1}{|\phi_n(u_n)|} = e^{u_n^2 (c+\gamma_n)/2n} \leq \Gamma n^{(r-1)^+/2}. \tag{4.15}$$

The estimation error $\widehat{C}'(u_n)_1 - c$ is the sum $G_n + H_n$ of the deterministic and stochastic errors:

$$\begin{aligned}G_n &= -\frac{2n}{u_n^2} \log\big|\phi_n(u_n)\big| - c = \frac{\gamma_n}{u_n^2}, \\ H_n &= \frac{2n}{u_n^2}\big(\log\big|\phi_n(u_n)\big| - \big(\log\big|\widehat{\phi}_n(u_n)\big|\big) 1_{\{\widehat{\phi}_n(u_n) \neq 0\}}\big).\end{aligned}$$

The previous estimates on γ_n readily yield

$$|G_n| \le \frac{2A}{u_n^{2-r}}. \tag{4.16}$$

Second, we study H_n. The variables $\exp(iu_n\Delta_j^n X)$ are i.i.d. as j varies, with modulus 1 and expectation $\phi_n(u_n)$, hence $V_n = \widehat{\phi}_n(u_n) - \phi_n(u_n)$ satisfies $\mathbb{E}(|V_n|^2) \le 1/n$. In view of (4.15), on the set $\{|V_n| \le 1/n^{r/4}\}$ we have $|V_n/\phi_n(u_n)| \le 1/2$ and $\widehat{\phi}_n(u_n) = V_n + \phi_n(u_n) \ne 0$ as soon as $n \ge n_0 = (2\Gamma)^{4/((2-r)\wedge r)}$, in which case we deduce, for some universal constant K:

$$|H_n| = \frac{2n}{u_n^2}\left|\log\left|1 + \frac{V_n}{\phi_n(u_n)}\right|\right| \le K\frac{n|V_n|}{u_n^2|\phi_n(u_n)|}.$$

Henceforth, if $n \ge n_0$,

$$\mathbb{E}\big(|H_n|1_{\{|V_n|\le 1/n^{r/4}\}}\big) \le \begin{cases} \dfrac{K\Gamma}{\sqrt{n}}, & \text{if } r \le 1,\\ \dfrac{KA\Gamma}{(r-1)n^{(2-r)/2}\log n}, & \text{if } r > 1.\end{cases} \tag{4.17}$$

Putting together (4.16) and (4.17), plus the fact that $\mathbb{P}(|V_n| > 1/n^{r/4}) \le 1/n^{(2-r)/2}$ (by Bienaymé–Tchebycheff inequality) tends to zero, and the equality $\widehat{C}'(u_n)_1 - c = G_n + H_n$, we deduce that $\rho_n(\widehat{C}'(u_n)_1 - c)$ [with the notation (3.2)] is tight, uniformly in $X \in \mathcal{S}_A^{r,L}$.

REFERENCES

[1] BARNDORFF-NIELSEN, O. E., GRAVERSEN, S. E., JACOD, J., PODOLSKIJ, M. and SHEPHARD, N. (2006). A central limit theorem for realised power and bipower variations of continuous semimartingales. In *From Stochastic Calculus to Mathematical Finance* (Yu. Kabanov, R. Liptser and J. Stoyanov, eds.) 33–68. Springer, Berlin. MR2233534

[2] BARNDORFF-NIELSEN, O. E. and SHEPHARD, N. (2004). Power and bipower variation with stochastic volatility and jumps. *J. Financ. Econom.* **2** 1–48.

[3] JACOD, J. (2008). Asymptotic properties of realized power variations and related functionals of semimartingales. *Stochastic Process. Appl.* **118** 517–559. MR2394762

[4] JACOD, J. and PROTTER, P. (2012). *Discretization of Processes*. Springer, Heidelberg. MR2859096

[5] MANCINI, C. (2001). Disentangling the jumps of the diffusion in a geometric jumping Brownian motion. *G. Ist. Ital. Attuari* **LXIV** 19–47.

[6] MANCINI, C. (2011). The speed of convergence of the threshold estimator of integrated variance. *Stochastic Process. Appl.* **121** 845–855. MR2770909

[7] VETTER, M. (2010). Limit theorems for bipower variation of semimartingales. *Stochastic Process. Appl.* **120** 22–38. MR2565850

INSTITUT DE MATHÉMATIQUES DE JUSSIEU
UPMC (UNIVERSITÉ PARIS-6)
4 PLACE JUSSIEU
75005-PARIS
FRANCE
E-MAIL: jean.jacod@upmc.fr

INSTITUT FÜR MATHEMATIK
HUMBOLDT-UNIVERSITÄT ZU BERLIN
UNTERDEN LINDEN, 6
10099-BERLIN
GERMANY
E-MAIL: mreiss@math.hu-berlin.de