*The Annals of Statistics*
2013, Vol. 41, No. 3, 1462–1484
DOI: 10.1214/13-AOS1115


# QUARTICITY AND OTHER FUNCTIONALS OF VOLATILITY: EFFICIENT ESTIMATION

BY JEAN JACOD AND MATHIEU ROSENBAUM

*Université Pierre et Marie Curie (Paris 6)*

We consider a multidimensional Itô semimartingale regularly sampled on $[0, t]$ at high frequency $1/\Delta_n$, with $\Delta_n$ going to zero. The goal of this paper is to provide an estimator for the integral over $[0, t]$ of a given function of the volatility matrix. To approximate the integral, we simply use a Riemann sum based on local estimators of the pointwise volatility. We show that although the accuracy of the pointwise estimation is at most $\Delta_n^{1/4}$, this procedure reaches the parametric rate $\Delta_n^{1/2}$, as it is usually the case in integrated functionals estimation. After a suitable bias correction, we obtain an unbiased central limit theorem for our estimator and show that it is asymptotically efficient within some classes of sub models.

**1. Introduction.** Let $X$ be a semimartingale, which is observed at discrete times $i\Delta_n$ for $i = 0, 1, \ldots$, over a finite time interval $[0, T]$, with a discretization mesh $\Delta_n$ which is small and eventually goes to 0 (high-frequency setting). One of the main problems encountered in practice is the estimation of the integrated (squared) volatility (in finance terms), or equivalently of the continuous part of the quadratic variation $[X, X]_t$.

By now, this is a well-understood problem, at least when $X$ is an Itô semimartingale. For example, in the continuous one-dimensional case, if $X$ takes the form

$$X_t = X_0 + \int_0^t b_s\, ds + \int_0^t \sigma_s\, dW_s$$

the approximate quadratic variation $\sum_{i=1}^{[t/\Delta_n]} (X_{i\Delta_n} - X_{(i-1)\Delta_n})^2$, which of course converges to $[X, X]_t = \int_0^t \sigma_s^2\, ds$, enjoys a central limit theorem (CLT): the difference between these two processes, normalized by $\frac{1}{\sqrt{\Delta_n}}$, converges stably in law to a limit which is conditionally on $X$ a continuous Gaussian martingale with quadratic variation (equivalently, with variance) twice the so-called "quarticity," that is, $2\int_0^t \sigma_s^4\, ds$.

Although later we consider a much more general framework, allowing $X$ to be multi-dimensional and with jumps, in the Introduction we pursue the discussion in

Received July 2012; revised March 2013.
*MSC2010 subject classifications.* 60F05, 60G44, 62F12.
*Key words and phrases.* Semimartingale, high frequency data, volatility estimation, central limit theorem, efficient estimation.

this special one-dimensional continuous case. In various statistical problems one needs to estimate not only the quarticity, but functionals of the form

$$V(g)_t = \int_0^t g(c_s)\,ds \qquad \text{where } c_s = \sigma_s^2$$

(for relatively general test functions $g$, and to derive associated CLTs, see [5]); notice that we plug in the "spot" squared volatility $c_t$ rather than $\sigma_t$, since in any case it is impossible to determine the sign of $\sigma_t$ on the basis of the observation of the path $t \mapsto X_t$. The case $g(x) = x$ corresponds to the usual integrated volatility, and $g(x) = x^2$ to the quarticity.

Toward this aim, two methods are currently at hand:

(1) The first one is available if $g(x) = \mathbb{E}(f(U(x)_1, \dots, U(x)_k))$ for all $x \geq 0$, where the $U(x)_j$'s are independent $\mathcal{N}(0, x)$ variables and $f$ is a continuous function on $\mathbb{R}^k$, of polynomial growth. Then we know that

$$U^n(f)_t = \Delta_n \sum_{i=1}^{[t/\Delta_n]-k+1} f\left(\frac{\Delta_i^n X}{\sqrt{\Delta_n}}, \dots, \frac{\Delta_{i+k-1}^n X}{\sqrt{\Delta_n}}\right) \qquad \text{where } \Delta_i^n X = X_{i\Delta_n} - X_{(i-1)\Delta_n}, \tag{1.1}$$

converges to $V(g)_t$ in probability, and if $f$ is $C^1$ the rate of convergence is $1/\sqrt{\Delta_n}$, and in the associated CLT the limiting conditional variance is $\int_0^t F(c_s)\,ds$ for a suitable function $F$.

(2) The second one consists in using estimators for the spot volatility and approximating the integral $V(g)_t$ by Riemann sums, in which the spot volatility is replaced by its estimator; that is, we set

$$V^n(g)_t = \Delta_n \sum_{i=1}^{[t/\Delta_n]-k_n+1} g(\widehat{c}_i^n) \qquad \text{where } \widehat{c}_i^n = \frac{1}{k_n \Delta_n} \sum_{j=0}^{k_n-1} (\Delta_{i+j}^n X)^2 \tag{1.2}$$

for an arbitrary sequence of integers such that $k_n \to \infty$ and $k_n \Delta_n \to 0$. Then one knows that $V^n(g)_t \xrightarrow{\mathbb{P}} V(g)_t$ (when $g$ is continuous and of polynomial growth). But so far nothing is known about the rate of convergence of these estimators when $k_n$ goes to infinity (the situation $k_n = k$ not depending on $n$ is studied in [11] where the rate $1/\sqrt{\Delta_n}$ is obtained for power functions).

The first method is quite powerful and gives optimal rates, but the special form of $g$ puts strong constraints on this function [e.g., it is $C^\infty$ on $(0, \infty)$, and much more]. To tell the truth, in the one-dimensional case, by far the most useful test functions $g$ are the powers $g_p(x) = x^p$ (recall that $x \geq 0$ here) for $p > 0$, which are associated as above with $f_p(x) = |x|^{2p}/m_{2p}$, where $m_q$ is the $q$th absolute moment of $\mathcal{N}(0, 1)$. Nevertheless, some functions $g$ of interest might not be, or not in an obvious way, of this form or, more generally, linear combinations of functions of this form. In the multivariate case, however, with $X$ being $d$-dimensional and thus $U$ above as well, one typically finds asymptotic variances which are complicated functions of the $d \times d$-dimensional spot volatility. This is, for instance, the case when studying multipower variations for integrated volatility estimation in the presence of jumps; see, for example, [5]. In this situation and more generally for an arbitrary (smooth) function $g$ on the set $\mathcal{M}_d^+$ of all $d \times d$ symmetric nonnegative matrices, it is rather a difficult task in practice to find an integer $k \geq 1$ and a function $f$ on $(\mathbb{R}^d)^k$ such that, for all $x \in \mathcal{M}_d^+$, we have $g(x) = \mathbb{E}(f(U(x)_1, \ldots, U(x)_k))$, where again the $U(x)_j$'s are ($d$-dimensional) i.i.d. $\mathcal{N}(0, x)$.

In addition, this first method does not provide efficient estimation in general. To see that, consider the toy example $X_t = \sigma W_t$, where $\sigma$ is a constant, $c = \sigma^2$, $\Delta_n = \frac{1}{n}$ and $T = 1$. We thus observe the increments $\Delta_i^n X$ for $i = 1, \ldots, n$, or equivalently the $n$ variables $Y_i = \Delta_i^n X/\sqrt{\Delta_n}$. These variables are i.i.d. $\mathcal{N}(0, c)$, so the asymptotically best estimators for $c$ (efficient in all possible senses, and also the MLE) are $\widehat{c}_n = \frac{1}{n}\sum_{i=1}^n (Y_i)^2 = \sum_{i=1}^n (\Delta_i^n X)^2$, with convergence rate $\sqrt{n}$ and asymptotic variance $2c^2$. If instead one wants to estimate $c^p$ for some $p \neq 1$ in $(0, \infty)$, one can use $\widehat{c}(p)_n = \frac{1}{nm_{2p}}\sum_{i=1}^n |Y_i|^{2p} = \frac{n^{p-1}}{m_{2p}}\sum_{i=1}^n |\Delta_i^n X|^{2p}$, and the ordinary central limit theorem tells us that the rate of convergence is again $\sqrt{n}$, and the asymptotic variance is $\frac{m_{4p}-m_{2p}^2}{m_{2p}^2}c^{2p}$: this is exactly what the first method above does. But this is not optimal, the asymptotically optimal estimators being $(\widehat{c}_n)^p$ (the MLE again), with rate $\sqrt{n}$ and asymptotic variance $2p^2c^{2p}$, smaller than the previous one when $p \neq 1$. Now, taking $(\widehat{c}_n)^p$ is exactly what the second method (1.2) does.

The aim of this paper is to develop the second method, and in particular to provide a central limit theorem, with the rate $1/\sqrt{\Delta_n}$ (as it is usually the case in a nonparametric setting for integrated functionals estimation; see, e.g., [2, 3]), and with an asymptotic variance always smaller than if one uses the first method. This can be viewed as an extension, in several directions, of the "block method" of Mykland and Zhang in [11]. About efficiency, and despite the title of the paper, we do not really examine the question in the general nonparametric or semi-parametric setting assumed below, since even for the simpler problem of estimating the integrated volatility, the concept of efficiency is not well established so far. Instead, we will term as "efficient" a procedure which is efficient in the usual sense for the sub-model consisting in the toy model $X_t = \sigma W_t$ above, and efficient in the sense of the Hajek convolution theorem, for the Markov-type model recently studied by Clément, Delattre and Gloter in [4] and of the form

$$dX_t = a(X_t)\,dt + f(t, X_t, Y_t)\,dW_t, \qquad dY_t = \overline{b}_t\,dt + \overline{\sigma}_t d\overline{W}_t, \tag{1.3}$$

where $a, f$ are unknown smooth enough functions and $\overline{b}, \overline{\sigma}$ arbitrary processes and where the two Brownian motions $W, \overline{W}$ are independent.

This will be done in the multivariate setting and when $X$ possibly has jumps (upon suitably truncating the increments in (1.2) if it is the case, in the spirit of [9, 10]), and under the additional assumptions that $c_t$ itself is an Itô semimartingale and that, when $X$ jumps, these jumps are summable, which are exactly the same assumptions under which the truncated versions of $U^n(f)$ in (1.1) converge with rate $1/\sqrt{\Delta_n}$.

The paper is organized as follows: Section 2 is devoted to presenting the assumptions. Results are given in Section 3, and all proofs are gathered in Section 4.

**2. Setting and assumptions.** The underlying process $X$ is $d$-dimensional, and observed at the times $i\Delta_n$ for $i = 0, 1, \dots,$ within a fixed interval of interest $[0, t]$. For any process $Y$ we use the notation $\Delta_i^n Y$ defined in (1.1) for the increment over the $i$th observation interval. We assume that the sequence $\Delta_n$ goes to 0. The precise assumptions on $X$ are as follows:

First, $X$ is an Itô semimartingale on a filtered space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$. It can be written in its Grigelionis form, using a $d$-dimensional Brownian motion $W$ and a Poisson random measure $\mu$ on $\mathbb{R}_+ \times E$, where $E$ is an auxiliary Polish space and with the (nonrandom) intensity measure $\nu(dt, dz) = dt \otimes \lambda(dz)$ for some $\sigma$-finite measure $\lambda$ on $E$,

$$\begin{aligned} X_t = X_0 &+ \int_0^t b_s\, ds + \int_0^t \sigma_s\, dW_s \\ &+ \int_0^t \int_E \delta(s, z) 1_{\{\|\delta(s,z)\|\leq 1\}} (\mu - \nu)(ds, dz) \\ &+ \int_0^t \int_E \delta(s, z) 1_{\{\|\delta(s,z)\|>1\}} \mu(ds, dz). \end{aligned} \tag{2.1}$$

This is a vector-type notation: the process $b_t$ is $\mathbb{R}^d$-valued optional, the process $\sigma_t$ is $\mathbb{R}^d \otimes \mathbb{R}^d$-valued optional, $\delta = \delta(\omega, t, z)$ is a predictable $\mathbb{R}^d$-valued function on $\Omega \times \mathbb{R}_+ \times E$ and $\|\cdot\|$ denotes the Euclidean norm on any finite-dimensional linear space. Besides the measurability requirements above, and for any $r \in [0, 2]$, we introduce the assumption:

ASSUMPTION (H-$r$). There are a sequence $(J_n)$ of nonnegative bounded $\lambda$-integrable functions on $E$ and a sequence $(\tau_n)$ of stopping times increasing to $\infty$, such that

$$\begin{aligned} t < \tau_n(\omega) \quad &\Rightarrow \quad \|b_t(\omega)\| \leq n, \qquad \|\sigma_t(\omega)\| \leq n, \\ t \leq \tau_n(\omega) \quad &\Rightarrow \quad \|\delta(\omega, t, z)\|^r \wedge 1 \leq J_n(z). \end{aligned} \tag{2.2}$$

The spot volatility process $c_t = \sigma_t \sigma_t^*$ ($^*$ denotes transpose) takes its values in the set $\mathcal{M}_d^+$ of all nonnegative symmetric $d \times d$ matrices. We will indeed suppose that $c_t$ is again an Itô semimartingale, and we consider the following assumption:

ASSUMPTION (A-$r$). The process $X$ satisfies Assumption (H-$r$), the associated volatility process $c$ satisfies (H-2) and the processes $b_t$ and, when $r \le 1$, $b'_t = b_t - \int \delta(t,z) 1_{\{\|\delta(t,z)\| \le 1\}} \lambda(dz)$ are càglàd or càdlàg.

The bigger $r$ is, the weaker Assumption (A-$r$) is, and when (A-0) holds the process $X$ has finitely many jumps on each finite interval. Since we suppose in the theorems of the next section that $r < 1$, the last condition in (2.2) implies that $b'_t$ is indeed well defined, and it is the "genuine" drift, in the sense that this is the drift after removing the sum $\sum_{s \le t} \Delta X_s$ of all jumps (which here are summable, and we even have $\sum_{s \le t} \|\Delta X_s\|^r < \infty$ a.s. here).

## 3. The results.

3.1. *A (seemingly) natural choice for the window $k_n$.* In order to define the estimators of the spot volatility, we need to fix a sequence $k_n$ of integers and a sequence $u_n$ of cut-off levels in $(0, \infty]$. The $\mathcal{M}_d^+$-valued variables $\widehat{c}_i^n$ are defined, componentwise, as

$$\widehat{c}_i^{n,lm} = \frac{1}{k_n \Delta_n} \sum_{j=0}^{k_n - 1} \Delta_{i+j}^n X^l \Delta_{i+j}^n X^m 1_{\{\|\Delta_{i+j}^n X\| \le u_n\}}, \tag{3.1}$$

and they implicitly depend on $\Delta_n, k_n, u_n$.

A natural idea is to choose the sequence $k_n$ satisfying, as $n \to \infty$,

$$k_n \sim \frac{\theta}{\sqrt{\Delta_n}}, \qquad \theta \in (0, \infty). \tag{3.2}$$

Indeed, one knows that $\widehat{c}_{[t/\Delta_n]}^n \xrightarrow{\mathbb{P}} c_t$ for any $t$, as soon as $k_n \to \infty$ and $k_n \Delta_n \to 0$, and there is an associated central limit theorem under Assumption (A-$r$) for some $r < 2$, with rate $\min(1/\sqrt{k_n}, 1/\sqrt{k_n \Delta_n})$, which reaches its biggest value $1/\Delta_n^{1/4}$ when $k_n \asymp 1/\sqrt{\Delta_n}$: this choice of $k_n$ ensures a balance between the involved "statistical error" which is of order $1/\sqrt{k_n}$, and the variation of $c_t$ over the interval $[t, t + k_n \Delta_n]$, which is of order $\sqrt{k_n \Delta_n}$ because $c_t$ is an Itô semimartingale (and even when it jumps); see [1, 5].

By Theorem 9.4.1 of [5], and again as soon as $k_n \to \infty$ and $k_n \Delta_n \to 0$, one also knows that

$$V(g)_t^n := \Delta_n \sum_{i=1}^{[t/\Delta_n] - k_n + 1} g(\widehat{c}_i^n) \quad \overset{\text{u.c.p.}}{\Longrightarrow} \quad V(g)_t := \int_0^t g(c_s)\, ds \tag{3.3}$$

(convergence in probability, uniform over each compact interval; by convention $\sum_{i=a}^b v_i = 0$ whenever $b < a$), as soon as the function $g$ on $\mathcal{M}_d^+$ is continuous with $|g(x)| \leq K(1+\|x\|^p)$ for some constants $K$, $p$, and under either one of the following three conditions:

(3.4)
- (A-0) holds, $X$ is continuous, $\frac{u_n}{\Delta_n^\varepsilon} \to \infty$ for some $\varepsilon < \frac{1}{2}$ (e.g., $u_n \equiv \infty$);
- (A-$r$) holds for some $r < 2$ and $p \leq 1$ and $u_n \asymp \Delta_n^\varpi$ for some $\varpi \in (0, \frac{1}{2})$;
- (A-$r$) holds for some $r < 2$ and $p > 1$ and $u_n \asymp \Delta_n^\varpi$ for some $\varpi \in [\frac{p-1}{2p-r}, \frac{1}{2})$.

Notice the upper limit in definition (3.3) of $V^n(g)_t$: this is to ensure that $V^n(g)_t$ is actually computable from the observations up to the time horizon $t$. Note also that when $X$ is continuous, the truncation in (3.1) is useless: one may use (3.1) with $u_n \equiv \infty$, which reduces to (1.2) in the one-dimensional case.

Now, we want to determine at which rate convergence (3.3) takes place. This amounts to proving an associated central limit theorem. Under the restriction $r < 1$ and an appropriate choice of the truncation levels, such a CLT is available for $V(g)^n$, with the rate $1/\sqrt{\Delta_n}$, but the limit exhibits a bias term.

Below, $g$ is a smooth function on $\mathcal{M}_d^+$, and the two first partial derivatives are denoted as $\partial_{jk} g$ and $\partial^2_{jk,lm} g$, since any $x \in \mathcal{M}_d^+$ has $d^2$ components $x^{jk}$. The family of all partial derivatives of order $j$ is simply denoted as $\partial^j g$.

THEOREM 3.1. *Assume Assumption (A-$r$) for some $r < 1$. Let $g$ be a $C^3$ function on $\mathcal{M}_d^+$ such that*

$$\|\partial^j g(x)\| \leq K\left(1+\|x\|^{p-j}\right), \qquad j = 0, 1, 2, 3,$$

*for some constants $K > 0$, $p \geq 3$. Either suppose that $X$ is continuous and $u_n/\Delta_n^\varepsilon \to \infty$ for some $\varepsilon < 1/2$ (e.g., $u_n \equiv \infty$, so there is no truncation at all), or suppose that*

$$u_n \asymp \Delta_n^\varpi, \qquad \frac{2p-1}{2(2p-r)} \leq \varpi < \frac{1}{2}.$$

*Then we have the finite-dimensional (in time) stable convergence in law*

$$\frac{1}{\sqrt{\Delta_n}}\left(V(g)_t^n - V(g)_t\right) \stackrel{\mathcal{L}_f-s}{\longrightarrow} A_t^1 + A_t^2 + A_t^3 + A_t^4 + Z_t,$$

*where $Z$ is a process defined on an extension $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, (\widetilde{\mathcal{F}}_t)_{t\geq 0}, \widetilde{\mathbb{P}})$ of $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, which conditionally on $\mathcal{F}$ is a continuous centered Gaussian martingale with variance*

$$\widetilde{\mathbb{E}}\left((Z_t)^2 \mid \mathcal{F}\right) = \sum_{j,k,l,m=1}^d \int_0^t \partial_{jk} g(c_s)\, \partial_{lm} g(c_s) \left(c_s^{jl} c_s^{km} + c_s^{jm} c_s^{kl}\right) ds,$$

*and where*

$$A_t^1 = -\frac{\theta}{2}\big(g(c_0) + g(c_t)\big),$$

$$A_t^2 = \frac{1}{2\theta} \sum_{j,k,l,m=1}^{d} \int_0^t \partial^2_{jk,lm} g(c_s)\big(c_s^{jl} c_s^{km} + c_s^{jm} c_s^{kl}\big)\, ds,$$

$$A_t^3 = -\frac{\theta}{12} \sum_{j,k,l,m=1}^{d} \int_0^t \partial^2_{jk,lm} g(c_s)\widetilde{c}_s^{jk,lm}\, ds,$$

*where $\widetilde{c}_s$ is the volatility process of $c_t$,*

$$A_t^4 = \theta \sum_{s \le t} G(c_{s-}, \Delta c_s)$$

*with $G(x, y) = \int_0^1 (g(x + wy) - (1 - w)g(x) - wg(x + y))\, dw$.*

Note that $|G(x, y)| \le K(1 + \|x\|)^p \|y\|^2$, so the sum defining $A_t^4$ is absolutely convergent, and vanishes when $c_t$ is continuous.

The bias has four parts:

(1) The first part $A^1$ is a border effect, easily eliminated by taking

$$\widetilde{V}(g)_t^n = V(g)_t^n + \frac{(k_n - 1)\Delta_n}{2}\big(g(\widehat{c}_1^n) + g(\widehat{c}^n_{[t/\Delta_n]-k_n+1})\big) \tag{3.5}$$

instead of $V(g)_t^n$: we then have $\frac{1}{\sqrt{\Delta_n}}(\widetilde{V}(g)_t^n - V(g)_t) \xrightarrow{\mathcal{L}_f-s} A_t^2 + A_t^3 + A_t^4 + Z_t$, and this convergence is even functional in time when $c_t$ is continuous.

(2) The second part $A^2$ is continuous in time and is present even for the toy model $X_t = \sqrt{c}W_t$ with $c$ a constant and $\Delta_n = \frac{1}{n}$ and $T = 1$. In this simple case it can be interpreted as follows: instead of taking the "optimal" $g(\widehat{c}_n)$ for estimating $g(c)$, with $\widehat{c}_n = \sum_{i=1}^n (\Delta_i^n X)^2$, one takes $\frac{1}{n}\sum_{i=1}^n g(\widehat{c}_i^n)$ with $\widehat{c}_i^n$ a "local" estimator of $c$. This adds a statistical error which results in a bias.

(3) The third and fourth parts $A^3$ and $A^4$ are, respectively, continuous and purely discontinuous, due to the continuous part and to the jumps of the volatility process $c_t$ itself. These two biases disappear if we take $\theta = 0$ in (3.2) (with still $k_n \to \infty$).

The only test function $g$ for which the biases $A^2, A^3, A^4$ disappear is the identity $g(x) = x$. This is because, in this case, and up to border terms, $\widetilde{V}(g)_t^n$ is nothing but the realized quadratic variation itself and the spot estimators $\widehat{c}_i^n$ actually merge together and disappear as such.

It is possible to consistently estimate $A_t^2, A_t^3, A_t^4$, and thus de-bias $\widetilde{V}(g)_t^n$ and obtain a CLT with a conditionally centered Gaussian limit. Consistent estimators for $A_t^2$ are easy to derive, since $A_t^2 = V(f)_t$ for the function $f(x) =$

$\sum_{j,k,l,m} \partial^2_{jk,lm} g(x)(x^{jl}x^{km} + x^{jm}x^{kl})$. Consistent estimators for $A^3_t$ and $A^4_t$, involving the volatility and the jumps of $c_t$, are more complicated to describe, especially the last one, and also likely to have poor performances. All the details about the way to remove the bias together with the proof of Theorem 3.1 can be found in [7].

3.2. *A suitable window $k_n$.* In front of the difficulties involved in de-biasing the estimators $V(g)^n_t$ above, we in fact choose a window size $k_n$ smaller than the one in (3.2). Namely, we choose $k_n$ such that, as $n \to \infty$,

$$k_n^3 \Delta_n \to \infty, \qquad k_n^2 \Delta_n \to 0. \tag{3.6}$$

Of course, the second condition enables us to make the first and last two bias terms in Theorem 3.1 vanish, which is technically very convenient. However, it amplifies the first bias term, which becomes the leading term in the difference $V(g)^n - V(g)$, and thus a prior de-biasing is necessary if we want a rate $1/\sqrt{\Delta_n}$. This leads us to consider the following estimator:

$$\begin{aligned} V'(g)^n_t = \Delta_n \sum_{i=1}^{[t/\Delta_n]-k_n+1} \Biggl( g(\widehat{c}^n_i) - \frac{1}{2k_n} \sum_{j,k,l,m=1}^{d} \partial^2_{jk,lm} g(\widehat{c}^n_i) \\ \times (\widehat{c}^{n,jl}_i \widehat{c}^{n,km}_i + \widehat{c}^{n,jm}_i \widehat{c}^{n,kl}_i)\Biggr). \end{aligned} \tag{3.7}$$

This estimator uses overlapping intervals, in the sense that we estimate $c_{(i-1)\Delta_n}$ on the basis of the time window $((i-1)\Delta_n, (i+k_n-1)\Delta_n]$, and then sum over all $i$'s. Another version is indeed possible, which does not use overlapping intervals and is as follows:

$$\begin{aligned} V''(g)^n_t = k_n \Delta_n \sum_{i=0}^{[t/k_n\Delta_n]-1} \Biggl( g(\widehat{c}^n_{ik_n+1}) \\ - \frac{1}{2k_n} \sum_{j,k,l,m=1}^{d} \partial^2_{jk,lm} g(\widehat{c}^n_{ik_n+1}) \\ \times (\widehat{c}^{n,jl}_{ik_n+1} \widehat{c}^{n,km}_{ik_n+1} + \widehat{c}^{n,jm}_{ik_n+1} \widehat{c}^{n,kl}_{ik_n+1})\Biggr). \end{aligned} \tag{3.8}$$

We can now give the final version of our associated central limit theorems.

THEOREM 3.2. *Assume Assumption (A-$r$) for some $r < 1$. Let $g$ be a $C^3$ function on $\mathcal{M}^+_d$ such that*

$$\|\partial^j g(x)\| \le K(1 + \|x\|^{p-j}), \qquad j = 0, 1, 2, 3, \tag{3.9}$$

*for some constants $K > 0$, $p \geq 3$. Either suppose that $X$ is continuous and $u_n/\Delta_n^{\varepsilon} \to \infty$ for some $\varepsilon < 1/2$ (e.g., $u_n \equiv \infty$, so there is no truncation at all), or suppose that*

$$u_n \asymp \Delta_n^{\varpi}, \qquad \frac{2p-1}{2(2p-r)} \leq \varpi < \frac{1}{2}. \tag{3.10}$$

*Then under (3.6) we have the two (functional in time) stable convergences in law*

$$\frac{1}{\sqrt{\Delta_n}}\big(V'(g)^n - V(g)\big) \stackrel{\mathcal{L}-s}{\Longrightarrow} Z, \qquad \frac{1}{\sqrt{\Delta_n}}\big(V''(g)^n - V(g)\big) \stackrel{\mathcal{L}-s}{\Longrightarrow} Z, \tag{3.11}$$

*where $Z$ is a process defined on an extension $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, (\widetilde{\mathcal{F}}_t)_{t\geq 0}, \widetilde{\mathbb{P}})$ of $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, which conditionally on $\mathcal{F}$ is a continuous centered Gaussian martingale with variance*

$$\widetilde{\mathbb{E}}\big((Z_t)^2 \mid \mathcal{F}\big) = \sum_{j,k,l,m=1}^{d} \int_0^t \partial_{jk} g(c_s)\, \partial_{lm} g(c_s) \big(c_s^{jl} c_s^{km} + c_s^{jm} c_s^{kl}\big)\, ds. \tag{3.12}$$

REMARK 3.3. When $X$ jumps, the requirement (3.10) is exactly the same as in Theorem 3.1, and it implies $r < 1$. This restriction is not a surprise, since one needs $r \leq 1$ in order to estimate the integrated volatility by the (truncated) realized volatility, with a rate of convergence $1/\sqrt{\Delta_n}$. Indeed, it is shown in [6] that if $r > 1$, the optimal rate in the minimax sense is $(\sqrt{\Delta_n}\log(1/\sqrt{\Delta_n}))^{-(2-r)/2}$. When $r = 1$ it is likely that the CLT still holds for an appropriate choice of the sequence $u_n$, and with another additional bias; see, for example, [12] for a slightly different context. Here we let this borderline case aside.

REMARK 3.4. The limiting process $Z$ is the same in both Theorems 3.1 and 3.2, but in the latter case the functional convergence always holds. It is also the same for (the normalized versions of) the processes $V'(g)^n$ and $V''(g)^n$, which is somewhat a surprise since in many instances using overlapping intervals instead of nonoverlapping intervals results in a strictly smaller asymptotic variance; this is for example the case for multipower variations, see Theorem 11.2.1 in [5]. However, in practice, it is probably advisable to use $V'(g)^n$ rather than $V''(g)^n$, because the former estimator is likely to be less sensitive to way-off values of the spot estimators $\widehat{c}_i^n$ than the latter one, due to the "smoothing" embedded in its definition.

REMARK 3.5. The $C^3$ property of $g$ is somewhat restrictive, as, for example, in the one-dimensional case it rules out the powers $g(x) = x^r$ with $r \in (0,3) \setminus \{1,2\}$. It could be proved that, in the one-dimensional case again, and if the processes $c_t$ and $c_{t-}$ do not vanish (equivalently, the process $1/c_t$ is locally bounded), the result still holds when $g$ is $C^3$ on $(0,\infty)$ and satisfies (3.9) with an arbitrary $p > 0$: here again, the fact that $1/c_t$ is locally bounded is also necessary for having a CLT for the functionals of (1.1) (say, with $k = 1$) when the test function $f$ is $C^1$ outside 0 only.

REMARK 3.6. One should compare this result with those of Mykland and Zhang in [11]: in that paper [in which only the continuous one-dimensional case and the test functions $g(x) = x^r$ are considered] the authors propose to take $k_n = k$ in (3.1). Of course (3.6) fails, but $V(g)^n$ in this case is actually of the form (1.1) and a CLT holds for $\frac{1}{\sqrt{\Delta_n}}(\alpha(g,k)V(g)^n - V(g))$ [without de-biasing term, but with an appropriate multiplicative factor $\alpha(g,k)$, which is explicitly known]: the asymptotic variance is bigger than in (3.12), but approaches this value when $k$ is large.

An advantage of Mykland–Zhang's approach is that when $g$ is positive, hence $V(g)_t$ as well, the estimators are also positive. In contrast, $V'(g)_t^n$ in (3.7) may be negative even when $g \geq 0$ everywhere. Thus if this positivity issue is important for a specific application, taking $k_n = k$ "large" and the estimator $\alpha(g,k)V(g)_t^n$ might be advisable, although it seems to work only when $g$ is a power function. Moreover, if $V'(g)_t^n$ is negative, it probably means that there is not enough data in order to obtain a relevant estimation.

It is simple to make this CLT "feasible," that is, usable in practice for determining a confidence interval for $V(g)_t$ at any time $t > 0$. Indeed, we can define the following function on $\mathcal{M}_d^+$:

$$\overline{h}(x) = \sum_{j,k,l,m=1}^{d} \partial_{jk} g(x)\, \partial_{lm} g(x) \big(x^{jl} x^{km} + x^{jm} x^{kl}\big), \tag{3.13}$$

which is continuous with $\overline{h}(x) \leq K(1 + \|x\|^{2p-2})$, and nonnegative (and positive at each $x$ such that $\partial g(x) \neq 0$). (3.10) implies the last condition in (3.4), and we have $V(\overline{h})^n \stackrel{\text{u.c.p.}}{\Longrightarrow} V(\overline{h})$, with $V(\overline{h})_t$ being the right-hand side of (3.12). Then we readily deduce:

COROLLARY 3.7. *Under the assumptions of the previous theorem, for any $t > 0$ we have the following stable convergence in law, where $Y$ is an $\mathcal{N}(0,1)$ variable*:

$$\frac{V'(g)_t^n - V(g)_t}{\sqrt{\Delta_n V(\overline{h})_t^n}} \xrightarrow{\mathcal{L}-s} Y \qquad \text{in restriction to the set } \big\{V(\overline{h})_t > 0\big\}, \tag{3.14}$$

*and the same holds with $V''(g)_t^n$ instead of $V'(g)_t^n$.*

3.3. *Optimality of the procedures.* We address now the question of the optimality of our procedures.

For simplicity, we restrict our attention to the one-dimensional case $d = 1$. We denote by $\mathcal{S}$ the class of all one-dimensional continuous semimartingales $X$ of the form (1.3), with $a, f$ being $C^3$ functions with bounded derivatives with further $f$ bounded away from 0, and $W, \overline{W}$ being two independent Brownian motions, and $\overline{b}_t, \overline{\sigma}_t$ being Lebesgue square-integrable processes, optional with respect to the filtration generated by $\overline{W}$, and with $(\overline{\sigma}_t)^2$ bounded away from 0. Such an $X$ satisfies (A-0), with $\sigma_t = f(t, X_t, Y_t)$.

Let $t > 0$. In the following, we say that a sequence of estimators $(T_t^n)_{n\geq 1}$ of $V(g)_t$ satisfy Property $\mathcal{P}$ over $\mathcal{S}$ if:

(1) the estimator $T_t^n$ is a function of $(X_{i\Delta_n} : 0 \leq i \leq [t/\Delta_n])$;

(2) for any $X \in \mathcal{S}$, the variables $\frac{1}{\sqrt{\Delta_n}}(T_t^n - V(g)_t)$ converge stably in law to a limit $Z'_t$ (depending of $g$ of course), defined on an extension of the space.

The following theorem gives three small steps toward optimality.

THEOREM 3.8. *Let $d = 1$ and $g$ be a $C^3$ function on $\mathbb{R}_+$ satisfying* (3.9) *and which is strictly increasing*, *or strictly decreasing*.

(a) *For the parametric model $X_t = \sigma W_t$, where $c_t = \sigma_t^2 = c$ is a constant* (*the toy example of the Introduction*), *for any $t > 0$, the estimators $V'(g)_t^n$ and $V''(g)_t^n$ are asymptotically efficient* (*in Le Cam's sense*) *for estimating the number $tg(c)$.*

(b) *Let $(T_t^n)_{n\geq 1}$ be a sequence of estimators satisfying $\mathcal{P}$ over the class of continuous processes $X$ for which* (A-0) *holds. Assume $Z'_t$ has a conditional variance of the form*

$$\widetilde{\mathbb{E}}\big((Z'_t)^2 \mid \mathcal{F}\big) = \int_0^t H(c_s)\,ds \tag{3.15}$$

*for some nonnegative Borel function $H$. Then necessarily $H \geq \overline{h}$, as given by* (3.13), *and in particular,*

$$\widetilde{\mathbb{E}}\big((Z'_t)^2 \mid \mathcal{F}\big) \geq \widetilde{\mathbb{E}}\big((Z_t)^2 \mid \mathcal{F}\big). \tag{3.16}$$

(c) *The estimators $V'(g)_t^n$ and $V''(g)_t^n$ are optimal over $\mathcal{S}$ in the following sense*: *for any sequence $(T_t^n)$ of estimators satisfying $\mathcal{P}$ over $\mathcal{S}$, the limiting variable $Z'_t$ can be realized as $Z_t + Z''_t$, where $Z_t$ is the limiting process in* (3.11), *and the variable $Z''_t$ is independent of $Z_t$ conditionally on $\mathcal{F}$.*

Part (b) of Theorem 3.8 shows in particular that the estimators $U^n(f)_t$ given in (1.1) for estimating $g(x) = \mathbb{E}(f(\sqrt{x}U))$ have always an asymptotic variance bigger than or equal to the variance (3.12).

Part (c) states that our estimators achieve the lower bounds of Hajek *convolution theorem* over the class $\mathcal{S}$. This convolution theorem for the subclass $\mathcal{S}$ is due to Clément, Delattre and Gloter; see [4]. It in particular implies that for given $t$, any rate optimal estimator over $\mathcal{S}$ has a limiting variance which is larger than those of $Z_t$ the limiting process in (3.11).

So far, however, a "general" theory of optimality in our nonparametric context seems still out of reach.

EXAMPLE 3.9 (Quarticity). Suppose $d = 1$, and take $g(x) = x^2$, so we want to estimate the quarticity $\int_0^t c_s^2\,ds$. In this case an "optimal" estimator for the quarticity is

$$\Delta_n\left(1-\frac{2}{k_n}\right)\sum_{i=1}^{[t/\Delta_n]-k_n+1}(\widehat{c}_i^n)^2.$$

The asymptotic variance is $8\int_0^t c_s^4\,ds$, to be compared with the asymptotic variance of the more usual estimators $\frac{1}{3\Delta_n}\sum_{i=1}^{[t/\Delta_n]}(\Delta_i^n X)^4$, which is $\frac{32}{3}\int_0^t c_s^4\,ds$.

REMARK 3.10. Although taking (3.6) eliminates the bias terms $A_t^1$, $A_t^3$ and $A_t^4$ showing in Theorem 3.1, it might be judicious to still eliminate the (asymptotically negligible) bias $A_t^1$ by adding to $V'(g)_t^n$ the same correction term $\frac{(k_n-1)\Delta_n}{2}(g(\widehat{c}_1^n)+g(\widehat{c}_{[t/\Delta_n]-k_n+1}^n)$ as in (3.5).

Due to their probable instability, it does not seem advisable, though, to eliminate the biases $A_t^3$ and $A_t^4$ by using (with the proper normalization) the method of [7].

**4. Proofs.** Under Assumption (A-$r$), not only do we have (2.1), but we can write $c_t$ in a similar fashion:

$$c_t = c_0 + \int_0^t \widetilde{b}_s\,ds + \int_0^t \widetilde{\sigma}_s\,dW'_s + \int_0^t\int_E \widetilde{\delta}(s,z)1_{\{\|\widetilde{\delta}(s,z)\|\le 1\}}(\mu-\nu)(ds,dz)$$
$$+\int_0^t\int_E \widetilde{\delta}(s,z)1_{\{\|\widetilde{\delta}(s,z)\|>1\}}\mu(ds,dz)$$

(here, $W'$ is a $d^2$-dimensional Brownian motion, possibly correlated with $W$). Then, according to the localization Lemma 4.4.9 of [5] [for the assumption (K) in that lemma], it is enough to show Theorem 3.2 under the following stronger assumption:

ASSUMPTION (SA-$r$). We have Assumption (A-$r$). Moreover we have, for a $\lambda$-integrable function $J$ on $E$ and a constant $A$,

$$\|b\|,\ \|\widetilde{b}\|,\ \|c\|,\ \|\widetilde{c}\|,\ J\le A,\qquad \|\delta(\omega,t,z)\|^r\le J(z),$$
$$\|\widetilde{\delta}(\omega,t,z)\|^2\le J(z). \tag{4.1}$$

In the sequel we suppose that $X$ satisfies Assumption (SA-$r$), and also that (3.6) holds: these assumptions are typically not recalled. Below, all constants are denoted by $K$, and they vary from line to line. They may implicitly depend on the process $X$ [usually through $A$ in (4.1)]. When they depend on an additional parameter $p$, we write $K_p$.

Recall the notation $b'_t$ in Assumption (A-$r$). We will usually replace the discontinuous process $X$ by the continuous process

$$X'_t = \int_0^t b'_s\,ds + \int_0^t \sigma_s\,dW_s, \tag{4.2}$$

connected with $X$ by $X_t = X_0 + X'_t + \sum_{s\le t}\Delta X_s$. Note that $b'$ is bounded, and without loss of generality we will use below its càdlàg version.

4.1. *Estimates*. (1) First, we recall well-known estimates for $X'$ and $c$. Under (4.1) and for $s,t\ge 0$ and $q\ge 0$, we have

$$\begin{aligned}
&\mathbb{E}\Big(\sup_{w\in[0,s]}\|X'_{t+w}-X'_t\|^q \mid \mathcal{F}_t\Big) \le K_q s^{q/2},\\
&\big\|\mathbb{E}\big(X'_{t+s}-X'_t \mid \mathcal{F}_s\big)\big\| \le Ks,\\
&\mathbb{E}\Big(\sup_{w\in[0,s]}\|c_{t+w}-c_t\|^q \mid \mathcal{F}_t\Big) \le K_q s^{1\wedge(q/2)},\\
&\big\|\mathbb{E}(c_{t+s}-c_t \mid \mathcal{F}_s)\big\| \le Ks.
\end{aligned}\tag{4.3}$$

We need slightly more refined estimates for $X'$, and before giving them we introduce some simplifying notation,

$$\begin{aligned}
&c_i^n = c_{(i-1)\Delta_n}, \qquad \mathcal{F}_i^n = \mathcal{F}_{(i-1)\Delta_n},\\
&\eta_{t,s} = \sup\big(\big\|b'_{t+u}-b'_t\big\|^2 : u\in[0,s]\big),\\
&\eta_{i,j}^n = \sqrt{\mathbb{E}\big(\eta_{(i-1)\Delta_n, j\Delta_n} \mid \mathcal{F}_i^n\big)}, \qquad \eta_i^n = \eta_{i,k_n}^n.
\end{aligned}\tag{4.4}$$

LEMMA 4.1. *We have*

$$\begin{aligned}
&\big|\mathbb{E}\big(\Delta_i^n X'^j\,\Delta_i^n X'^m \mid \mathcal{F}_i^n\big) - c_i^{n,jm}\Delta_n\big|\\
&\qquad \le K\Delta_n^{3/2}\big(\sqrt{\Delta_n}+\eta_{i,1}^n\big),\\
&\big|\mathbb{E}\big(\Delta_i^n X'^j\,\Delta_i^n X'^k\,\Delta_i^n X'^l\,\Delta_i^n X'^m \mid \mathcal{F}_i^n\big) - \big(c_i^{n,jk}c_i^{n,lm} + c_i^{n,jl}c_i^{n,km} + c_i^{n,jm}c_i^{n,kl}\big)\Delta_n^2\big|\\
&\qquad \le K\Delta_n^{5/2}.
\end{aligned}$$

PROOF. For simplicity we prove the result when $i=1$, so $\Delta_1^n X' = X'_{\Delta_n}$, but upon shifting time the proof for $i>1$ is the same.

First we have $X'_t = M_t + tb'_0 + \int_0^t (b'_s - b'_0)\,ds$, where $M$ is a martingale with $M_0 = 0$. Taking the $\mathcal{F}_0$-conditional expectation thus yields

$$\big\|\mathbb{E}\big(X'_t \mid \mathcal{F}_0\big) - tb'_0\big\| \le t\eta_{0,t}. \tag{4.5}$$

Next, Itô's formula yields that $X_t'^j X_t'^m$ is the sum of a martingale vanishing at 0, plus

$$b_0'^j \int_0^t X_s'^m\,ds + b_0'^m \int_0^t X_s'^j\,ds + \int_0^t X_s'^m (b_s'^j - b_0'^j)\,ds + \int_0^t X_s'^j (b_s'^m - b_0'^m)\,ds + c_0^{jm} t + \int_0^t (c_s^{jm} - c_0^{jm})\,ds.$$

Upon taking the conditional expectation, and using the Cauchy–Schwarz inequality and the first and the last parts of (4.3), plus (4.5), we readily deduce

$$\big|\mathbb{E}(X_t'^j X_t'^m \mid \mathcal{F}_0) - t c_0^{jm}\big| \le K t^{3/2}(\sqrt{t} + \eta_{0,t}). \tag{4.6}$$

With $t = \Delta_n$, this gives the first claim. Finally, for any indices $j_1, \dots, j_4$ Itô's formula yields a martingale $M$ vanishing at 0 such that

$$\begin{aligned}\prod_{l=1}^4 \Delta_1^n X'^{j_l} = M_{\Delta_n} &+ \sum_{l=1}^p \int_0^{\Delta_n} b_s'^{j_l} \prod_{1\le m\le p, m\ne l} X_s'^{j_m}\,ds \\ &+ \frac{1}{2} \sum_{1\le l,l'\le d, l\ne l'} c_0^{j_l j_{l'}} \int_0^{\Delta_n} \prod_{1\le m\le 4, m\ne l,l'} X_s'^{j_m}\,ds \\ &+ \frac{1}{2} \sum_{1\le l,l'\le d, l\ne l'} \int_0^{\Delta_n} (c_s^{j_l j_{l'}} - c_0^{j_l j_{l'}}) \prod_{1\le m\le 4, m\ne l,l'} X_s'^{j_m}\,ds.\end{aligned} \tag{4.7}$$

Again, we take the $\mathcal{F}_0$-conditional expectation and we deal with the second, the third and the last term in the right-hand side above by Fubini's theorem and the Cauchy–Schwarz inequality. For the fourth term we use (4.6), and a simple calculation yields the second claim. □

LEMMA 4.2. *For all $t > 0$ we have $\Delta_n \mathbb{E}(\sum_{i=1}^{[t/\Delta_n]} \eta_i^n) \to 0$, and for all $j, k$ such that $j + k \le k_n$ we have $\mathbb{E}(\eta_{i+j,k}^n \mid \mathcal{F}_i^n) \le \eta_i^n$.*

PROOF. The second claim follows from the definitions of $\eta_i^n$ and $\eta_{i,j}^n$ and the Cauchy–Schwarz inequality. For the first claim, we observe that $\mathbb{E}((\eta_i^n)^2)$ is smaller than a constant always, and than $\frac{1}{\Delta_n}\int_{(i-2)\Delta_n}^{(i-1)\Delta_n} \mathbb{E}((\eta_{s,2k_n+1})^2)\,ds$ when $i \ge 2$. Hence by the Cauchy–Schwarz inequality,

$$\begin{aligned}\Delta_n \mathbb{E}\left(\sum_{i=1}^{[t/\Delta_n]} \eta_i^n\right) &\le \left(t\mathbb{E}\left(\Delta_n \sum_{i=1}^{[t/\Delta_n]} (\eta_i^n)^2\right)\right)^{1/2} \\ &\le \left(K t \Delta_n + \mathbb{E}\left(t \int_0^t (\eta_{s,2k_n+1})^2\,ds\right)\right)^{1/2}.\end{aligned}$$

We have $\eta_{s,2k_n+1} \le K$, and the càdlàg property of $b'$ yields that $\eta_{s,2k_n+1}(\omega) \to 0$ for all $\omega$, and all $s$ except for countably many strictly positive values (depending on $\omega$). Then, the first claim follows by the dominated convergence theorem. $\square$

(2) It is much easier (although unfeasible in practice) to replace $\widehat{c}_i^n$ in (3.3) by the estimators based on the process $X'$, as given by (4.2). Namely, we will replace $\widehat{c}_i^n$ by the following:

$$\widehat{c}_i^{\prime n} = \frac{1}{k_n \Delta_n} \sum_{j=0}^{k_n-1} \Delta_{i+j}^n X' \Delta_{i+j}^n X'^*.$$

The comparison between $\widehat{c}_i^n$ and $\widehat{c}_i^{\prime n}$ is based on the following consequence of Lemma 13.2.6 of [5], applied with $F(x) = xx^*$, so $k = 1$ and $p' = s' = 2$ and $s = 1$ and $\varepsilon = 0$ (because $r < 1$) with the notation of that lemma. Namely, we have for all $q \ge 1$ and for some sequence $a_n$ going to 0,

$$\begin{aligned} &\mathbb{E}\bigl(\bigl\|(\Delta_i^n X \Delta_i^n X^*)1_{\{\|\Delta_i^n X\| \le u_n\}} - (\Delta_i^n X' \Delta_i^n X'^*)1_{\{\|\Delta_i^n X'\| \le u_n\}}\bigr\|^q\bigr) \\ &\quad \le K_q a_n \Delta_n^{(2q-r)\varpi+1}. \end{aligned}$$

Since $\mathbb{E}(\|\Delta_i^n X'\|^{2q}) \le K_q \Delta_n^q$ for any $q > 0$ by classical estimates, implying by Markov's inequality that $\mathbb{E}(\|\Delta_i^n X'\|^{2q} 1_{\{\|\Delta_i^n X'\| > u_n\}}) \le K \Delta_n^{q+q'(1-2\varpi)}$ for any $q' > 0$, by taking $q' > \frac{1}{1-2\varpi}$, we then easily deduce

$$\mathbb{E}\bigl(\|\widehat{c}_i^n - \widehat{c}_i^{\prime n}\|^q\bigr) \le K_q a_n \Delta_n^{(2q-r)\varpi+1-q}. \tag{4.8}$$

(3) Let us introduce the following $\mathbb{R}^d \otimes \mathbb{R}^d$-valued variables:

$$\begin{aligned} \alpha_i^n &= \Delta_i^n X' \Delta_i^n X'^* - c_i^n \Delta_n, \\ \beta_i^n &= \widehat{c}_i^{\prime n} - c_i^n = \frac{1}{k_n \Delta_n} \sum_{j=0}^{k_n-1} \bigl(\alpha_{i+j}^n + (c_{i+j}^n - c_i^n)\Delta_n\bigr). \end{aligned} \tag{4.9}$$

From (4.3) we get that for all $q \ge 0$,

$$\mathbb{E}\bigl(\|\alpha_i^n\|^q \mid \mathcal{F}_i^n\bigr) \le K_q \Delta_n^q, \qquad \bigl\|\mathbb{E}(\alpha_i^n \mid \mathcal{F}_i^n)\bigr\| \le K \Delta_n^{3/2}. \tag{4.10}$$

This and the Burkholder–Gundy and Hölder inequalities give us, for $q \ge 2$, that $\mathbb{E}(\|\sum_{j=0}^{k_n-1} \alpha_{i+j}^n\|^q \mid \mathcal{F}_i^n) \le K_q \Delta_n^q k_n^{q/2}$. This and (4.3) and again Hölder's inequality yield

$$q \ge 2 \quad \Rightarrow \quad \mathbb{E}\bigl(\|\beta_i^n\|^q \mid \mathcal{F}_i^n\bigr) \le K_q\bigl(k_n^{-q/2} + k_n \Delta_n\bigr). \tag{4.11}$$

Lemma 4.1 allows us for better estimates for $\alpha_i^n$, namely

$$\begin{aligned} &\bigl\|\mathbb{E}(\alpha_i^n \mid \mathcal{F}_i^n)\bigr\| \le K \Delta_n^{3/2}\bigl(\sqrt{\Delta_n} + \eta_{i,1}^n\bigr), \\ &\bigl|\mathbb{E}(\alpha_i^{n,jk} \alpha_i^{n,lm} \mid \mathcal{F}_i^n) - (c_i^{n,jl} c_i^{n,km} + c_i^{n,jm} c_i^{n,kl})\Delta_n^2\bigr| \le K \Delta_n^{5/2}. \end{aligned} \tag{4.12}$$

LEMMA 4.3. *We have*

$$\|\mathbb{E}(\beta_i^n \mid \mathcal{F}_i^n)\| \le K\sqrt{\Delta_n}\big(k_n\sqrt{\Delta_n}+\eta_i^n\big),$$
$$\Big|\mathbb{E}(\beta_i^{n,jk}\beta_i^{n,lm} \mid \mathcal{F}_i^n) - \frac{1}{k_n}\big(c_i^{n,jl}c_i^{n,km}+c_i^{n,jm}c_i^{n,kl}\big)\Big|$$
$$\le K\sqrt{\Delta_n}\big(k_n^{-1/2}+k_n\sqrt{\Delta_n}+\eta_i^n\big).$$

PROOF. The first claim follows from (4.3), (4.12) and the last part of Lemma 4.2. For the second one, we set $\xi_i^n = c_i^{n,jl}c_i^{n,km}+c_i^{n,jm}c_i^{n,kl}$ and $\zeta_{i,j}^n=\alpha_{i+j}^n+(c_{i+j}^n-c_i^n)\Delta_n$ and write $\beta_i^{n,jk}\beta_i^{n,lm}$ as

$$\begin{aligned}&\frac{1}{k_n^2\Delta_n^2}\sum_{u=0}^{k_n-1}\zeta_{i,u}^{n,jk}\zeta_{i,u}^{n,lm}+\frac{1}{k_n^2\Delta_n^2}\sum_{u=0}^{k_n-2}\sum_{v=u+1}^{k_n-1}\zeta_{i,u}^{n,jk}\zeta_{i,v}^{n,lm}\\ &\quad+\frac{1}{k_n^2\Delta_n^2}\sum_{u=0}^{k_n-2}\sum_{v=u+1}^{k_n-1}\zeta_{i,u}^{n,lm}\zeta_{i,v}^{n,jk}.\end{aligned}\tag{4.13}$$

First, we have

$$|\zeta_{i,u}^{n,jk}\zeta_{i,u}^{n,lm}-\alpha_{i+u}^{n,jk}\alpha_{i+u}^{n,lm}|\le 2\Delta_n\|c_{i+u}^n-c_i^n\|\|\alpha_{i+u}^n\|+\Delta_n^2\|c_{i+u}^n-c_i^n\|^2,$$

whose $\mathcal{F}_i^n$-conditional expectation is less than $K\Delta_n^{5/2}k_n^{1/2}$ by (4.3) and (4.10). The boundedness of $c_t$ and (4.3) yield $|\mathbb{E}(\xi_{i+u}^n \mid \mathcal{F}_i^n)-\xi_i^n|\le Kk_n\Delta_n$. Then (4.12) gives us that the $\mathcal{F}_i^n$-conditional expectation of the first term in (4.13), minus $\frac{1}{k_n}\xi_i^n$, is less than $K\sqrt{\Delta_n}/\sqrt{k_n}$.

Second, (4.3) and (4.12), plus the first claim of Lemma 4.1, yield, when $0\le u<v<k_n$,

$$\begin{aligned}\big|\mathbb{E}(\zeta_{i,v}^{n,jk}\mid\mathcal{F}_{i+u+1}^n)-(c_{i+u+1}^{n,jk}-c_i^{n,jk})\Delta_n\big|&\le K\Delta_n^{3/2}\big(k_n\sqrt{\Delta_n}+\eta_{i+v,1}^n\big),\\ \big|\mathbb{E}\big(\alpha_{i+u}^{n,lm}(c_{i+u+1}^{n,jk}-c_{i+u}^{n,jk})\mid\mathcal{F}_{i+u}^n\big)\big|&\le K\Delta_n^{3/2}\big(\sqrt{\Delta_n}+\eta_{i+u,1}^n\big),\\ \big|\mathbb{E}\big(\alpha_{i+u}^{n,lm}(c_{i+u}^{n,jk}-c_i^{n,jk})\mid\mathcal{F}_{i+u}^n\big)\big|&\le K\Delta_n^{3/2}\big(\sqrt{\Delta_n}+\eta_{i+u,1}^n\big),\\ \big|\mathbb{E}\big((c_{i+u}^{n,lm}-c_i^{n,lm})(c_{i+u+1}^{n,jk}-c_i^{n,jk})\mid\mathcal{F}_i^n\big)\big|&\le Kk_n\Delta_n.\end{aligned}$$

Since $\zeta_{i+u}^n$ is $\mathcal{F}_{i+u+1}^n$-measurable, and using (4.10) and the second part of Lemma 4.2, the $\mathcal{F}_i^n$-conditional expectation of the last term of (4.13) is smaller than $K\sqrt{\Delta_n}(k_n\sqrt{\Delta_n}+\eta_i^n)$. The same is obviously true for the second term, and we readily deduce the second claim of the lemma. □

4.2. *Proof of Theorem* 3.2. Using the key property $\widehat{c}_i'^n = c_i^n+\beta_i^n$ and the definition (4.9) of $\beta_i^n$, a simple calculation shows the decomposition $\frac{1}{\sqrt{\Delta_n}}(V'(g)_t^n -$

$V(g)_t) = \sum_{j=1}^{5} V_t^{n,j}$, as soon as $t > k_n\Delta_n$, and where

$$V_t^{n,1} = \sqrt{\Delta_n} \sum_{i=1}^{[t/\Delta_n]-k_n+1} \Biggl( g(\widehat{c}_i^n) - g(\widehat{c}_i'^n) - \frac{1}{2k_n} \sum_{j,k,l,m=1}^{d} \bigl(\partial_{jk,lm}^2 g(\widehat{c}_i^n)(\widehat{c}_i^{n,jl}\widehat{c}_i^{n,km} + \widehat{c}_i^{n,jm}\widehat{c}_i^{n,kl}) - \partial_{jk,lm}^2 g(\widehat{c}_i'^n)(\widehat{c}_i'^{n,jl}\widehat{c}_i'^{n,km} + \widehat{c}_i'^{n,jm}\widehat{c}_i'^{n,kl})\bigr)\Biggr),$$

$$V_t^{n,2} = \frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^{[t/\Delta_n]-k_n+1} \int_{(i-1)\Delta_n}^{i\Delta_n} \bigl(g(c_i^n) - g(c_s)\bigr)\,ds - \frac{1}{\sqrt{\Delta_n}} \int_{\Delta_n([t/\Delta_n]-k_n+1)}^{t} g(c_s)\,ds,$$

$$V_t^{n,3} = \sqrt{\Delta_n} \sum_{i=1}^{[t/\Delta_n]-k_n+1} \sum_{l,m=1}^{d} \partial_{lm} g(c_i^n) \frac{1}{k_n} \sum_{u=0}^{k_n-1} (c_{i+u}^{n,lm} - c_i^{n,lm}),$$

$$V_t^{n,4} = \sqrt{\Delta_n} \sum_{i=1}^{[t/\Delta_n]-k_n+1} \Biggl( g(c_i^n + \beta_i^n) - g(c_i^n) - \sum_{l,m=1}^{d} \partial_{lm} g(c_i^n)\beta_i^{n,lm} - \frac{1}{2k_n} \sum_{j,k,l,m=1}^{d} \partial_{jk,lm}^2 g(c_i^n + \beta_i^n) \times \bigl((c_i^{n,jl} + \beta_i^{n,jl})(c_i^{n,km} + \beta_i^{n,km}) + (c_i^{n,jm} + \beta_i^{n,jm})(c_i^{n,kl} + \beta_i^{n,kl})\bigr)\Biggr),$$

$$V_t^{n,5} = \frac{1}{k_n\sqrt{\Delta_n}} \sum_{i=1}^{[t/\Delta_n]-k_n+1} \sum_{l,m=1}^{d} \partial_{lm} g(c_i^n) \sum_{u=0}^{k_n-1} \alpha_{i+u}^{n,lm}.$$

The leading term is $V^{n,5}$, and the first claim in (3.11), about $V'(g)^n$, is a consequence of the following two lemmas:

LEMMA 4.4. *For $v = 1, 2, 3, 4$ we have* $V^{n,v} \stackrel{u.c.p.}{\Longrightarrow} 0$.

LEMMA 4.5. *With $Z$ as in Theorem 3.2, we have the functional stable convergence in law*

$$V^{n,5} \stackrel{\mathcal{L}-s}{\Longrightarrow} Z. \tag{4.14}$$

PROOF OF LEMMA 4.4. *The case* $v = 1$: We define functions $h_n$ on $\mathcal{M}_d^+$ by

$$h_n(x) = g(x) - \frac{1}{2k_n} \sum_{j,k,l,m=1}^{d} \partial^2_{jk,lm} g(x)\big(x^{jl}x^{km} + x^{jm}x^{kl}\big).$$

From (3.9) we obtain $|h_n(x) - h_n(y)| \le K(1 + \|y\|)^{p-1})\|x - y\| + K\|x - y\|^p$ (uniformly in $n$). So if $\eta_i^n$ is the $i$th summand in the definition of $V_t^{n,1}$, we get

$$\big|\eta_i^n\big| \le K\big(1 + \|\widehat{c}_i^n\|^{p-1} + \|\widehat{c}_i'^n\|^{p-1}\big)\|\widehat{c}_i^n - \widehat{c}_i'^n\| + K\|\widehat{c}_-^n\widehat{c}_i'^n\|^p.$$

Recalling the last part of (4.10), and by (4.8), Hölder's inequality and the fact that $(2p - r)\varpi + 1 - p < \frac{1}{q}((2q - r)\varpi + 1 - q)$ when $q > 1$ is small enough, because $\varpi < \frac{1}{2}$, we deduce $\mathbb{E}(|g(\widehat{c}_i^n) - g(\widehat{c}_i'^n)|) \le K a_n \Delta_n^{(2p-r)\varpi+1-p}$ and thus

$$\mathbb{E}\Big(\sup_{s\le t}\big|V_s^{n,1}\big|\Big) \le K t a_n \Delta_n^{(2p-r)\varpi+1/2-p}.$$

In view of (3.10), we deduce the result for $v = 1$.

*The case* $v = 2$: Since $g(c_s)$ is bounded, it is obvious that the absolute value of the last term in $V_t^{n,2}$ is smaller than $Kk_n\sqrt{\Delta_n}$, which goes to 0 by (3.6). Since $g$ is $C^2$, the convergence of the first term in $V_t^{n,2}$ to 0 in probability, locally uniformly in $t$, is well known; see, for example, the proof of (5.3.24) in [5], in which one replaces $\rho_{c_s}(f)$ by $g(c_s)$. Thus the result holds for $v = 2$.

*The case* $v = 3$: Letting $\zeta_i^n = \sum_{l,m=1}^{d} \partial_{lm} g(c_i^n) \frac{1}{k_n} \sum_{u=0}^{k_n-1} (c_{i+u}^{n,lm} - c_i^{n,lm})$ be the $i$th summand in the definition of $V_t^{n,3}$, and $N(n, j, t)$ be the integer part of $([t/\Delta_n] - k_n - j + 1)/k_n$, we have

$$V_t^{n,3} = \sqrt{\Delta_n} \sum_{j=1}^{k_n} H(j)_t^n \qquad \text{where } H(j)_t^n = \sum_{i=0}^{N(n,j,t)} \zeta_{j+k_n i}^n.$$

From (4.3) and the Cauchy–Schwarz inequality, we get

$$\big|\mathbb{E}\big(\zeta_i^n \mid \mathcal{F}_i^n\big)\big| \le Kk_n\Delta_n, \qquad \mathbb{E}\big(\big|\zeta_i^n\big|^2 \mid \mathcal{F}_i^n\big) \le Kk_n\Delta_n.$$

Then Doob's inequality, the $\mathcal{F}_{j+k_n(i+1)}^n$-measurability of $\zeta_{j+k_n i}^n$, and $N(n, j, t) \le t/k_n\Delta_n$ imply

$$\begin{aligned}\mathbb{E}\Big(\sup_{s\le t}\big|H(j)_s^n\big|\Big) \le{}& \sum_{i=0}^{N(n,j,t)} \mathbb{E}\big(\big|\mathbb{E}\big(\zeta_{j+k_n i}^n \mid \mathcal{F}_{j+k_n i}^n\big)\big|\big) \\ &+ \left(4\sum_{i=0}^{N(n,j,t)} \mathbb{E}\big(\big(\zeta(j)_{j+k_n i}^n\big)^2\big)\right)^{1/2} \\ \le{}& K(t + \sqrt{t}).\end{aligned}$$

Since $|V_t^{n,3}| \le \sqrt{\Delta_n}\sum_{j=1}^{k_n}|H(j)_t^n|$ and $k_n\sqrt{\Delta_n} \to 0$, we deduce the result for $v = 3$.

*The case* $v = 4$: The $i$th summand in the definition of $V_t^{n,4}$ is $v_i^n + w_i^n$, where

$$v_i^n = \frac{1}{2}\sum_{j,k,l,m=1}^{d} \partial^2_{jk,lm} g(c_i^n)\Big(\beta_i^{n,jk}\beta_i^{n,lm} - \frac{1}{k_n}\big(c_i^{n,jl}c_i^{n,km} + c_i^{n,jm}c_i^{n,kl}\big)\Big),$$

$$|w_i^n| \le K\big(1 + \|\beta_i^n\|^{p-3}\big)\|\beta_i^n\|^3 + \frac{K}{k_n}\big(1 + \|\beta_i^n\|^{p-1}\big)\|\beta_i^n\|$$

[use (3.9) and $\|c_t\| \le K$ repeatedly], and we thus have $V_t^{n,4} = G_t^n + \sum_{j=1}^{k_n} H(j)_t^n$, with $N(n, j, t)$ as in the previous step and

$$G_t^n = \sqrt{\Delta_n}\sum_{i=1}^{[t/\Delta_n]-k_n+1}\big(w_i^n + \mathbb{E}(v_i^n \mid \mathcal{F}_i^n)\big),$$

$$H(j)_t^n = \sum_{i=0}^{N(n,j,t)} \zeta(j)_i^n, \qquad \zeta(j)_i^n = \sqrt{\Delta_n}\big(v_{j+k_n i}^n - \mathbb{E}(v_{j+k_n i}^n \mid \mathcal{F}_{j+k_n i}^n)\big).$$

In view of Lemma 4.3 and (4.11), plus Hölder's inequality, we have

$$\big|\mathbb{E}(v_i^n \mid \mathcal{F}_i^n)\big| \le K\sqrt{\Delta_n}\big(k_n\sqrt{\Delta_n} + \eta_i^n\big),$$

$$\mathbb{E}\big(|w_i^n|\big) \le K\left(\frac{1}{k_n^{3/2}} + k_n\Delta_n + \frac{\sqrt{\Delta_n}}{\sqrt{k_n}}\right),$$

and thus (3.6) and Lemma 4.2 yield

$$\mathbb{E}\Big(\sup_{s\le t}|G_s^n|\Big) \le \mathbb{E}\left(\sum_{i=1}^{[t/\Delta_n]}\sqrt{\Delta_n}\big(|w_i^n| + |\mathbb{E}(v_i^n \mid \mathcal{F}_i^n)|\big)\right) \to 0.$$

Moreover (4.11) and $k_n^{-2} \le K k_n\Delta_n$ yield $\mathbb{E}(|\zeta(j)_i^n|^2) \le K\Delta_n^2 k_n$, whereas $\zeta(j)_i^n$ is a martingale increment for the filtration $(\mathcal{F}_{j+k_n i}^n)_{i\ge 0}$, hence Doob's inequality and $N(n, j, t) \le t/k_n\Delta_n$ imply

$$\mathbb{E}\Big(\sup_{s\le t}|H(j)_s^n|\Big) \le \left(\sum_{i=0}^{N(n,j,t)}\mathbb{E}\big((\zeta(j)_i^n)^2\big)\right)^{1/2} \le Kt\Delta_n.$$

Since $|V_t^{n,4}| \le |G_t^n| + \sum_{j=1}^{k_n}|H(j)_t^n|$, we deduce the result for $v = 4$. □

PROOF OF LEMMA 4.5. We can rewrite $V^{n,5}$ as

$$V_t^{n,5} = \frac{1}{\sqrt{\Delta_n}}\sum_{i=1}^{[t/\Delta_n]}\sum_{l,m=1}^{d} w_i^{n,lm}\alpha_i^{n,lm},$$

where

$$w_i^{n,lm} = \frac{1}{k_n} \sum_{j=(i-[t/\Delta_n]+k_n-1)^+}^{(i-1)\wedge(k_n-1)} \partial_{lm} g(c_{i-j}^n).$$

Observe that $w_i^n$ and $\alpha_i^n$ are measurable with respect to $\mathcal{F}_i^n$ and $\mathcal{F}_{i+1}^n$, respectively, so by Theorem IX.7.28 of [8] (with $G = 0$ and $Z = 0$ in the notation of that theorem) it suffices to prove the following four convergences in probability, for all $t > 0$ and all component indices:

$$\frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^{[t/\Delta_n]-k_n+1} w_i^{n,lm} \mathbb{E}(\alpha_i^{n,lm} \mid \mathcal{F}_i^n) \xrightarrow{\mathbb{P}} 0, \tag{4.15}$$

$$\begin{aligned} &\frac{1}{\Delta_n} \sum_{i=1}^{[t/\Delta_n]-k_n+1} w_i^{n,jk} w_i^{n,lm} \mathbb{E}(\alpha_i^{n,jk} \alpha_i^{n,lm} \mid \mathcal{F}_i^n) \\ &\quad \xrightarrow{\mathbb{P}} \int_0^t \partial_{jk} g(c_s)\, \partial_{lm} g(c_s) (c_s^{jl} c_s^{km} + c_s^{jm} c_s^{kl})\, ds, \end{aligned} \tag{4.16}$$

$$\frac{1}{\Delta_n^2} \sum_{i=1}^{[t/\Delta_n]-k_n+1} \|w_i^n\|^4 \mathbb{E}(\|\alpha_i^n\|^4 \mid \mathcal{F}_i^n) \xrightarrow{\mathbb{P}} 0, \tag{4.17}$$

$$\frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^{[t/\Delta_n]-k_n+1} w_i^{n,lm} \mathbb{E}(\alpha_i^{n,lm} \Delta_i^n N \mid \mathcal{F}_i^n) \xrightarrow{\mathbb{P}} 0, \tag{4.18}$$

where $N = W^k$ for some $k$, or is an arbitrary bounded martingale, orthogonal to $W$.

Lemma 4.2, (4.10), (4.12) and the property $\|w_i^n\| \le K$ readily imply (4.15) and (4.17). In view of the form of $\alpha_i^n$, a usual argument (see, e.g., [5]) shows that in fact $\mathbb{E}(\alpha_i^{n,lm} \Delta_i^n N \mid \mathcal{F}_i^n) = 0$ for all $N$ as above, and hence (4.18) holds.

For (4.16), by (4.12) it suffices to prove that

$$\begin{aligned} &\Delta_n \sum_{i=1}^{[t/\Delta_n]-k_n+1} w_i^{n,jk} w_i^{n,lm} (c_i^{n,jl} c_i^{n,km} + c_i^{n,jm} c_i^{n,kl}) \\ &\quad \xrightarrow{\mathbb{P}} \int_0^t \partial_{jk} g(c_s)\, \partial_{lm} g(c_s) (c_s^{jl} c_s^{km} + c_s^{jm} c_s^{kl})\, ds. \end{aligned}$$

In view of the definition of $w_i^n$, for each $t$ we have $w_{i(n,t)}^{n,jk} \to \partial_{jk} g(c_t)$ and $c_{i(n,t)}^{n,jk} \to c_t^{jk}$ almost surely if $|i(n,t)\Delta_n - t| \le k_n \Delta_n$, and the above convergence follows by the dominated convergence theorem, thus ending the proof of (4.14). $\square$

PROOF OF THE SECOND CLAIM IN (3.11). The proof is basically the same as for the first claim. We have the decomposition $\frac{1}{\sqrt{\Delta_n}}(V''(g)_t^n - V(g)_t) =$

$\sum_{j=1}^{5} \overline{V}_t^{n,j}$, where

$$\overline{V}_t^{n,1} = k_n\sqrt{\Delta_n} \sum_{i=0}^{[t/k_n\Delta_n]-1} \big(g(\widehat{c}_{k_ni+1}^n) - g(\widehat{c}'^n_{k_ni+1})\big),$$

$$\overline{V}_t^{n,2} = \frac{1}{\sqrt{\Delta_n}} \sum_{i=0}^{[t/k_n\Delta_n]-1} \int_{k_ni\Delta_n}^{k_n(i+1)\Delta_n} \big(g(c_{k_ni\Delta_n}) - g(c_s)\big)\, ds$$
$$- \frac{1}{\sqrt{\Delta_n}} \int_{k_n\Delta_n([t/k_n\Delta_n])}^{t} g(c_s)\, ds,$$

$$\overline{V}_t^{n,3} = k_n\sqrt{\Delta_n} \sum_{i=0}^{[t/k_n\Delta_n]-1} \sum_{l,m=1}^{d} \partial_{lm} g(c_{k_ni+1}^n) \frac{1}{k_n} \sum_{u=0}^{k_n-1} (c_{k_ni+1+u}^{n,lm} - c_{k_ni+1}^{n,lm}),$$

$$\begin{aligned}\overline{V}_t^{n,4} = {}& k_n\sqrt{\Delta_n}\\ &\times \sum_{i=0}^{[t/k_n\Delta_n]-1} \Bigg( g(c_{k_ni+1}^n + \beta_{k_ni+1}^n) - g(c_{k_ni+1}^n)\\ &\qquad - \sum_{l,m=1}^{d} \partial_{lm} g(c_{k_ni+1}^n)\beta_{k_ni+1}^{n,lm}\\ &\qquad - \frac{1}{2k_n} \sum_{j,k,l,m=1}^{d} \partial_{jk,lm}^2 g(c_{k_ni+1}^n + \beta_{k_ni+1}^n)\\ &\qquad\qquad \times \big((c_{k_ni+1}^{n,jl} + \beta_{k_ni+1}^{n,jl})(c_{k_ni+1}^{n,km} + \beta_{k_ni+1}^{n,km})\\ &\qquad\qquad + (c_{k_ni+1}^{n,jm} + \beta_{k_ni+1}^{n,jm})(c_{k_ni+1}^{n,kl} + \beta_{k_ni+1}^{n,kl})\big)\Bigg),\end{aligned}$$

$$\overline{V}_t^{n,5} = \frac{1}{\sqrt{\Delta_n}} \sum_{i=0}^{[t/k_n\Delta_n]-1} \sum_{l,m=1}^{d} \partial_{lm} g(c_{k_ni+1}^n) \sum_{u=0}^{k_n-1} \alpha_{k_ni+u+1}^{n,lm}.$$

The proofs of Lemmas 4.4 and 4.5 carry over to $\overline{V}^{n,v}$ instead of $V^{n,v}$, for $v = 1,2,3,4,5$, almost word for word, except for the following points:

(1) For Lemma 4.4, cases $v = 3,4$, there is no need to consider the $k_n$ processes $H(j)^n$; a single process $H^n$ is enough, and the proof is simpler.

(2) For Lemma 4.4, case $v = 2$, the proof of the u.c.p. convergence to 0 of the first term in the definition of $\overline{V}^{n,2}$ should be reworked as follows: the $i$th summand $\zeta_i^n$ in this term is $\mathcal{F}_{k_n(i+1)}^n$-measurable, and by (3.9) and (4.3) it satisfies

$$\big|\mathbb{E}(\zeta_i^n \mid \mathcal{F}_{k_ni}^n)\big| \le K(k_n\Delta_n)^2, \qquad \mathbb{E}(|\zeta_i^n|^2 \mid \mathcal{F}_{k_ni}^n) \le K(k_n\Delta_n)^3.$$

Then the claim follows from the usual martingale argument and $k_n\sqrt{\Delta_n} \to 0$.

(3) For Lemma 4.5, we have

$$\overline{V}_t^{n,5} = \frac{1}{\sqrt{\Delta_n}} \sum_{i=1}^{k_n[t/k_n\Delta_n]} \sum_{l,m=1}^{d} \partial_{lm} g\big(c^n_{1+k_n[(j-1)/k_n]}\big)\alpha_i^{n,lm},$$

and the rest of the proof is similar. $\square$

4.3. *Proof of Theorem* 3.8(a) *and* (b). (a) is almost obvious: indeed, $\overline{V}(g)_t^n$ converges with the rate $\frac{1}{\sqrt{\Delta_n}}$ and is asymptotically normal with asymptotic variance $2tg'(c)^2c^2$ ($g'$ is the derivative of $g$). However, since $g$ is one-to-one, the model index by the new parameter $tg(c)$ is regular, and the MLE is $tg(\widehat{c}_n)$, where $\widehat{c}_n = \sum_{i=1}^{[t/\Delta_n]}(\Delta_i^n X)^2$, and clearly $tg(\widehat{c}_n)$ has the same asymptotic properties as $\overline{V}(g)_t^n$: this proves the result.

(b) is also obvious: the properties of $T_t^n$ hold for all continuous processes $X$ satisfying (A-0). Then, using the toy model of (a), the optimality proved above implies that $tH(c) \geq 2tg'(c)^2c^2$ for any constant $c > 0$, that is, $H \geq \overline{h}$.

Finally, (c) is exactly Theorem 3 of [4] applied to the present setting.

**Acknowledgments.** We thank two anonymous referees for very constructive comments, which led to substantial improvements. We also thank Jia Li for pointing out some problems in an earlier draft.

## REFERENCES

[1] ALVAREZ, A., PANLOUP, F., PONTIER, M. and SAVY, N. (2012). Estimation of the instantaneous volatility. *Stat. Inference Stoch. Process.* **15** 27–59. MR2892587

[2] BICKEL, P. J. and RITOV, Y. (2003). Nonparametric estimators which can be "plugged-in". *Ann. Statist.* **31** 1033–1053. MR2001641

[3] BIRGÉ, L. and MASSART, P. (1995). Estimation of integral functionals of a density. *Ann. Statist.* **23** 11–29. MR1331653

[4] CLÉMENT, E., DELATTRE, S. and GLOTER, A. (2013). An infinite dimensional convolution theorem with applications to the efficient estimation of the integrated volatility. *Stochastic Process. Appl.* **123** 2500–2521. MR3054534

[5] JACOD, J. and PROTTER, P. (2012). *Discretization of Processes. Stochastic Modelling and Applied Probability* **67**. Springer, Heidelberg. MR2859096

[6] JACOD, J. and REISS, M. (2013). A remark on the rates of convergence for integrated volatility estimation in the presence of jumps. Preprint. Available at arXiv:1209.4173.

[7] JACOD, J. and ROSENBAUM, M. (2012). Estimation of volatility functionals: The case of a $\sqrt{n}$ window. Technical report, Laboratoire de Probabilités et Modèles Aléatoires, Univ. Pierre et Marie Curie. Available at arXiv:1212.1997.

[8] JACOD, J. and SHIRYAEV, A. N. (2003). *Limit Theorems for Stochastic Processes*, 2nd ed. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **288**. Springer, Berlin. MR1943877

[9] MANCINI, C. (2009). Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scand. J. Stat.* **36** 270–296. MR2528985

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

[11] MYKLAND, P. A. and ZHANG, L. (2009). Inference for continuous semimartingales observed at high frequency. *Econometrica* **77** 1403–1445. MR2561071

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

INSTITUT DE MATHÉMATIQUES DE JUSSIEU
CNRS—UMR 7586
UNIVERSITÉ PIERRE ET MARIE CURIE
4 PLACE JUSSIEU
75 005 PARIS
FRANCE
E-MAIL: jean.jacod@upmc.fr

LABORATOIRE DE PROBABILITÉS
ET MODÈLES ALÉATOIRES
CNRS—UMR 7599
UNIVERSITÉ PIERRE ET MARIE CURIE
4 PLACE JUSSIEU
75 005 PARIS
FRANCE
E-MAIL: mathieu.rosenbaum@upmc.fr