Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2011, Vol. 47, No. 4, 1197–1218
DOI: 10.1214/11-AIHP432

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES
www.imstat.org/aihp

Irregular sampling and central limit theorems for power variations: The continuous case[1]

Takaki Hayashi[a], Jean Jacod[b] and Nakahiro Yoshida[c]

[a]*Graduate School of Business Administration, Keio University, 4-1-1 Hiyoshi, Yokohama 223-8526, Japan. E-mail: takaki@kbs.keio.ac.jp*
[b]*Institut de mathématiques de Jussieu, Université P. et M. Curie (Paris-6), 4 place Jussieu, 75252 Paris, France (CNRS UMR 7586). E-mail: jean.jacod@upmc.fr*
[c]*Graduate School of Mathematical Sciences, University of Tokyo, 3-8-1 Komaba, Tokyo 153, Japan. E-mail: nakahiro@ms.u-tokyo.ac.jp*

Received 6 March 2009; revised 2 May 2011; accepted 9 May 2011

Abstract. In the context of high frequency data, one often has to deal with observations occurring at irregularly spaced times, at transaction times for example in finance. Here we examine how the estimation of the squared or other powers of the volatility is affected by irregularly spaced data. The emphasis is on the kind of assumptions on the sampling scheme which allow to provide consistent estimators, together with an associated central limit theorem, and especially when the sampling scheme depends on the observed process itself.

Résumé. Dans le contexte de données à haute fréquences, il est fréquent de recueillir les informations le long d'une grille irrégilière, par exemple aux instants de transaction pour les données financières. Dans cet article, nous étudions comment l'estimation de l'intégrale du carré, ou d'autres puissances, de la volatilité est affectée par l'irrégularité des données. L'accent est mis sur le type d'hypothèses qu'il est nécessaire de faire sur la répartition des observations, en particulier lorsque celles-ci dépendent du processus observé lui-même, de façon à obtenir un théorème limite central pour nos estimateurs.

MSC: Primary 60G44; 62M09; secondary 60G42; 62G20

Keywords: Quadratic variation; Discrete observations; Power variations; High frequency data; Stable convergence

1. Introduction

The approximation of the quadratic variation by the "realized" quadratic variation plays a very important role in practice, especially for financial data, since such data are necessarily reported at discrete times. Other quantities of interest, based on discrete observations as well, have been considered in many recent papers: these quantities are sums of functions of the successive increments of the process, usually some powers or absolute powers of those increments. They are used to estimate some characteristics of the jumps of the observed process, or the volatility of the continuous part, or for various testing problems about jumps for example.

However, if the behavior of the realized quadratic variation and other similar functionals is well known when the observations come in regularly, this is no longer the case when the observation times are irregularly spaced and, even worse, when they are random. Relatively few papers are so far available in that case: see [1,14] and [15] for deterministic observation times, and [4] for some special random times like hitting times, and [16] for a situation which

[1]Supported in part by JST Basic Research Programs PRESTO, Grants-in-Aid for Scientific Research of Japan Society for the Promotion of Science, No. 19340021 and No. 19530186, and by the Global COE Program "The research and training center for new developments in mathematics" of the Graduate School of Mathematical Sciences, University of Tokyo.

is relatively close to the present paper, and [6–9] and [2] when the process is multidimensional and the observation times exhibit some sort of relatively restricted randomness, or are random but independent of the observed process. In the five last papers, the situation is quite complex because the various components are observed at different times. All those papers are concerned with a continuous underlying process. One may also quote related works dealing with estimation of various parameters with random sampling, like [5] for diffusions and [3] for Markov processes.

Here, our aim is relatively modest: we consider only a continuous underlying process, which is 1-dimensional, quite a restrictive setting indeed which in particular lets aside the problem of non-synchronous observations for multivariate processes. The observation times are possibly random, and we try to find conditions on those times, allowing for limit theorems when the number of observations increases (the "high frequency" setting), which are as general as possible, and in particular accommodate some dependency between the sampling times and the process itself: this is the main novelty of this paper. We also consider the problem of the estimation of integrated powers of the volatility other than 2, for example the estimation of the quarticity, a quantity which turns out to be more and more frequently used in practice. In a sense, we are in the line of [1] and [16], and we try to weaken the assumptions on the observation scheme and the underlying observed process X as much as we can, together with providing an estimation method for other powers than the integrated volatility.

Of course, the assumptions about the observation times, as discussed below, are arbitrary to some extent. What we have in mind is to account for high frequency financial data, and especially those recording transaction times and prices, which are notoriously irregularly spaced. Note that in this paper we only consider a semi-realistic situation, where the process is supposed to be observed exactly at each observation time: we do not consider the problem of microstructure noise. For real applications to finance, this problem should be taken into account.

Let us be more specific. The process of interest is a 1-dimensional continuous semimartingale X on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, which is of Itô type, that is of the form

$$X_t = X_0 + \int_0^t b_s \,\mathrm{d}s + \int_0^t \sigma_s \,\mathrm{d}W_s, \tag{1.1}$$

where W is a Wiener process and b and σ are progressively measurable processes. In this paper our aim is to give estimators for the following "integrated" quantities:

$$V(p)_t = m_p \int_0^t |\sigma_s|^p \,\mathrm{d}s \tag{1.2}$$

(here m_p is the pth absolute moment of the standard normal law $\mathcal{N}(0,1)$, and it appears here for later convenience), when $p \geq 2$, together with "feasible" estimators for the variance or conditional variance of these estimators, so as to be able to construct confidence intervals for example. The restriction $p \geq 2$ is not essential, similar results hold when $1 < p < 2$, and also when $0 < p \leq 1$ under some additional assumptions, but $p \geq 2$ covers the most interesting situations and is simpler technically speaking.

At stage n the process X is observed along a strictly increasing sequence of – possibly random – finite times $T(n,i), i \geq 0$, starting at $T(n,0) = 0$, and we use the notation

$$\left.\begin{array}{ll} \Delta(n,i) = T(n,i) - T(n,i-1), & I(n,i) = \big(T(n,i-1), T(n,i)\big], \\ N_t^n = \inf\big(i\colon\ T(n,i) > t\big) - 1, & \pi_t^n = \sup_{i=1,\dots,N_t^n+1} \Delta(n,i) \end{array}\right\} \tag{1.3}$$

(π_t^n is the "mesh" up to time t, by convention $\inf(\varnothing) = \infty$ and $\sup(\varnothing) = 0$, and N_t^n is the number of observation times within $(0,t]$). Also, for any process Y we write

$$\Delta_i^n Y = Y_{T(n,i)} - Y_{T(n,i-1)}. \tag{1.4}$$

Among (many) other technical requirements, we will always assume the following minimal ones:

$$\left.\begin{array}{llll} n \geq 1 & \Rightarrow & T(n,i) \to \infty & \mathbb{P}\text{-a.s., as } i \to \infty, \\ t \geq 0 & \Rightarrow & \pi_t^n \xrightarrow{\mathbb{P}} 0 & \text{as } n \to \infty. \end{array}\right\} \tag{1.5}$$

With the convention $\sum_{i=1}^{0} = 0$, we take as our estimator for $V(p)_t$ of (1.2), at stage n, the variable

$$V^n(p)_t = \sum_{i=1}^{N_t^n} \Delta(n,i)^{1-p/2} \left|\Delta_i^n X\right|^p. \tag{1.6}$$

The upper limit of the sum is such that $V^n(p)_t$ involves the observations occurring up to time t only. As is well known, under regular sampling (that is, $T(n,i) = i\Delta_n$ for some time lag Δ_n going to 0), then $V^n(p) \stackrel{\text{u.c.p.}}{\longrightarrow} V(p)$ (this denotes "convergence in probability, locally uniform in time"), as soon as σ is càdlàg and b satisfies a suitable integrability condition and even under no condition at all when $p = 2$. Furthermore under some more regularity conditions, we have an associated CLT (Central Limit Theorem), and standardized versions are also available: see e.g. [12] and [11].

Our aim is to prove the same results for irregular sampling, and in particular to prove standardized versions of the CLT for the estimators. We already know, of course, that $V^n(2)$ converges to $V(2)$ as soon as the $T(n,i)$'s are stopping times subject to (1.5), but for the other values of p and for the CLT, we need appropriate additional assumptions on the sampling scheme. Note that for a regular scheme the factor $\Delta(n,i)^{1-p/2} = \Delta_n^{1-p/2}$ goes out of the sum, and this is of course how the "standard" result is stated. The idea to place this factor inside the sum is due to [1].

In Section 2 we state the precise assumptions on the underlying process X and the sampling schemes. Section 3 contains the basic results, and a discussion of their applicability. The proofs are gathered in Sections 4, 5 and 6.

2. Setting and assumptions

2.1. *Assumptions on X*

The assumptions on X will vary according to the power p in which we are interested. We always suppose that X has the form (1.1), and we use two different assumptions, with (B) stronger than (A) below:

***Assumption* (A).** *We have* (1.1), *and the process b is locally bounded*, *and the process σ is càdlàg* (= *right continuous with left limits*).

***Assumption* (B).** *We have* (1.1), *and the process σ is also a* (*possibly discontinuous*) *Itô semimartingale*, *which can be written as*

$$\left.\begin{array}{l}\sigma_t = \sigma_0 + \displaystyle\int_0^t \widetilde{b}_s\,\mathrm{d}s + \int_0^t \widetilde{\sigma}_s\,\mathrm{d}W_s + M_t + \sum_{s\le t} \Delta\sigma_s\, 1_{\{|\Delta\sigma_s|>1\}}, \quad \textit{where}\\ \bullet\ \textit{M is a local martingale with } |\Delta M_t| \le 1, \textit{ orthogonal to } W, \textit{ and its}\\ \quad \textit{predictable quadratic covariation process is } \langle M,M\rangle_t = \int_0^t a'_s\,\mathrm{d}s.\\ \bullet\ \textit{The predictable compensator of } \sum_{s\le t} 1_{\{|\Delta\sigma_s|>1\}} \textit{ is } \int_0^t a_s\,\mathrm{d}s.\end{array}\right\} \tag{2.1}$$

Moreover, the processes $\widetilde{b}$, a and a' are locally bounded, *and the processes $\widetilde{\sigma}$ and b are left continuous with right limits.*

In (B), M may have jumps, and also a non-vanishing continuous martingale part, which must then be a stochastic integral with respect to another Brownian motion independent of W.

The above assumptions are exactly those under which the theorems given below hold, for regular sampling schemes, except for the case of $V^n(2)$ which requires a bit less.

2.2. *The sampling scheme*

The sampling scheme, that is the collection $((T(n,i))_{i\ge 0}\colon n \ge 1)$, is subject to a number of conditions. There is first a structural assumption:

***Assumption* (C).** *There is a sub-filtration $(\mathcal{F}_t^0)_{t\ge 0}$ of $(\mathcal{F}_t)_{t\ge 0}$ with the following properties*:

- *W and b and σ are adapted to $(\mathcal{F}_t^0)$;*
- *any $(\mathcal{F}_t^0)$ martingale is also an $(\mathcal{F}_t)$-martingale;*
- *each variable $T(n,i)$ is an $(\mathcal{F}_t)$-stopping time which, conditionally on $\mathcal{F}_{T(n,i-1)}$, is independent of the σ-field $\mathcal{F}^0 = \bigvee_{t>0} \mathcal{F}_t^0$.*

This assumption is satisfied when the $T(n,i)$'s are non-random (*deterministic schemes*), and when the $T(n,i)$'s are independent of the processes (W,X,b,σ) (*independent schemes*), but it includes many other cases as well. However it excludes some *a priori* interesting situations: when $\mathcal{F}_t^0 = \mathcal{F}_t$, then (C) amounts to saying that those stopping times are "strongly predictable" in the sense that $T(n,i)$ is $\mathcal{F}_{T(n,i-1)}$-measurable, and this is quite restrictive; for example it excludes the case where the $T(n,i)$'s are the successive hitting times of a spatial grid by X, a case considered in [4] under some restrictive assumptions on X.

Apart from (C) and (1.5), the sampling scheme should also be not too wildly scattered, and its meshes π_t^n should converge to 0 at some deterministic rate. This rate is expressed through a sequence $r_n \to \infty$ of positive – non random – numbers. The assumptions below all involve this sequence r_n, in an implicit way.

Before stating the assumptions, and for any $q \geq 0$, we introduce the processes

$$A(q)_t^n = r_n^{q-1} \sum_{i=1}^{N_t^n} \Delta(n,i)^q. \tag{2.2}$$

The normalization r_n^{q-1} is motivated by the regular schemes $T(n,i) = i\Delta_n$, for which $A(q)_t^n = \Delta_n[t/\Delta_n]$ (with the choice $r_n = 1/\Delta_n$) converges towards t for any $q \geq 0$. Note also that $A(0)_t^n = N_t^n/r_n$, and since $t - \pi_t^n \leq A(1)_t^n \leq t$ we deduce that

$$A(1)_t^n \xrightarrow{\mathbb{P}} t \tag{2.3}$$

as soon as (1.5) holds. Then, with $q \geq 0$, we set:

***Assumption* (D(q)).** *We have (C) and (1.5), and there is a (necessarily nonnegative) $(\mathcal{F}_t^0)$-optional process $a(q)$, such that for all t we have*

$$A(q)_t^n \xrightarrow{\mathbb{P}} \int_0^t a(q)_s \,\mathrm{d}s. \tag{2.4}$$

Note that (D(q)) for some sequence r_n implies (D(q)) for any other sequence r'_n such that $r'_n/r_n \to \alpha \in [0,\infty)$, and the new limit in (2.4) and when $q > 1$ is then $\alpha^{q-1}a(q)$, and in particular vanishes when $r'_n/r_n \to 0$: the forthcoming theorems which explicitly involve r_n are true but "empty" when the limit in (2.4) vanishes identically. Regular sampling schemes with lag Δ_n satisfy (D(q)) for all $q \geq 0$, with $r_n = 1/\Delta_n$ and $a(q)_t = 1$.

Some important connections between these assumptions are stated in the next lemma, to be proved in Section 4.

Lemma 2.1. *Let $0 \leq q < p < q'$. Then for all $0 \leq s < t$ we have*

$$A(p)_t^n - A(p)_s^n \leq \bigl(A(q)_t^n - A(q)_s^n\bigr)^{(q'-p)/(q'-q)} \bigl(A(q')_t^n - A(q')_s^n\bigr)^{(p-q)/(q'-q)}. \tag{2.5}$$

Moreover, if (D(q)) holds for some $q \neq 1$ and if p is strictly between 1 and q, from any subsequence one may extract a further subsequence which satisfies (D(p)), and we have versions of $a(q)$ and $a(p)$ satisfying $a(p)_t \leq a(q)_t^{(p-1)/(q-1)}$.

Note that (2.4) for all t implies that $r_n^{q-1} \sum_{i=1}^{N_t^n} H_{T(n,i)} \Delta(n,i)^q \xrightarrow{\text{u.c.p.}} \int_0^t H_s a(q)_s \,\mathrm{d}s$ as soon as H is càdlàg. In some cases we need this convergence to hold at a rate faster than $1/\sqrt{r_n}$, and we express this in the following assumption. It is indeed a very strong assumption, usually not satisfied by random schemes, and in particular not satisfied by the independent schemes for which the $\Delta(n,i)$'s are i.i.d. when i varies, but not deterministic.

***Assumption* (D′(q)).** *We have* (D(q)), *and further for all* $t \geq 0$ *and all càdlàg* $(\mathcal{F}^0_t)$*-adapted processes* H *we have*

$$\sqrt{r_n}\left(r_n^{q-1}\sum_{i=1}^{N^n_t} H_{T(n,i)}\Delta(n,i)^q - \int_0^t H_s a(q)_s \,\mathrm{d}s\right) \stackrel{\text{u.c.p.}}{\longrightarrow} 0. \tag{2.6}$$

For a deterministic scheme, (D(q)) may or may not be satisfied, but there is no simple criterion to ensure that it holds. For a random scheme, it may be useful to describe conditions on the laws or on the conditional laws of the lags $\Delta(n,i)$, which ensure (D(q)). For each n and each $q \geq 0$ we choose an $(\mathcal{F}_t)$-optional $(0,\infty]$-valued process $G(q)^n$ such that

$$G(q)^n_{T(n,i-1)} = r_n^q \mathbb{E}\big(\Delta(n,i)^q | \mathcal{F}_{T(n,i-1)}\big). \tag{2.7}$$

This specifies $G(q)^n_t$ only at the times $t = T(n,i)$, so there are many such processes $G(q)^n$. A simple choice consists in taking $G(q)^n_t$ to be equal to the right side of (2.7) when $T(n,i-1) \leq t < T(n,i)$ (a piecewise constant process). But other choices are possible, and perhaps more appropriate in view of the forthcoming assumption. We can obviously take $G(0)^n_t = 1$, and by Hölder's inequality, we can and will choose processes $G(q)^n$ which satisfy

$$0 \leq p \leq q \quad \Rightarrow \quad G(p)^n \leq \big(G(q)^n\big)^{p/q}. \tag{2.8}$$

Then we set, with $q \geq 1$:

***Assumption* (E(q)).** *We have* (C) *and* (1.5), *and for each* $p \in [0,q]$ *there is a càdlàg process* $G(p)$, *adapted to* $(\mathcal{F}^0_t)$, *and further* $G(1)$ *and* $G(1)_-$ *do not vanish, such that for an appropriate choice of* $G(p)^n$ *we have*

$$G(p)^n \stackrel{\text{u.c.p.}}{\longrightarrow} G(p). \tag{2.9}$$

Lemma 2.2. *Assume* (E(q)) *for some* $q > 1$. *Then* (D(p)) *holds for all* $p \in [0,q)$, *with*

$$a(p)_t = \frac{G(p)_t}{G(1)_t}, \tag{2.10}$$

and in particular

$$\frac{1}{r_n} N^n_t \stackrel{\mathbb{P}}{\longrightarrow} \int_0^t \frac{1}{G(1)_s}\,\mathrm{d}s. \tag{2.11}$$

For some results below we will also need a rate of convergence in (2.9), which is expressed in the following:

***Assumption* (E′(q)).** *We have* (E(q)), *and for each* $p \in [0,q]$ *we have* $\sqrt{r_n}(G(p)^n - G(p)) \stackrel{\text{u.c.p.}}{\longrightarrow} 0$.

Finally, we introduce a kind of sampling schemes which are somehow restrictive but should accommodate many practical applications, and are called *mixed renewal schemes*. These schemes are constructed as follows: we consider a filtration $(\mathcal{F}^0_t)$ as in (C), and a double sequence $(\varepsilon(n,i)\colon i,n \geq 1)$ of i.i.d. positive variables on $(\Omega,\mathcal{F},\mathbb{P})$, independent of $\mathcal{F}^0$, with moments

$$m'_q = \mathbb{E}\big(\varepsilon(n,i)^q\big). \tag{2.12}$$

We may have $m'_q = \infty$ for $q > 1$, but we assume that $m'_1 < \infty$. We consider a sequence v^n of positive $(\mathcal{F}_t)$-adapted processes, and we define $T(n,i)$ by induction on i as follows:

$$T(n,0) = 0, \qquad T(n,i+1) = T(n,i) + \frac{1}{r_n} v^n_{T(n,i)} \varepsilon(n,i+1). \tag{2.13}$$

Then we take for $(\mathcal{F}_t)$ the smallest filtration containing $(\mathcal{F}_t^0)$ and such that each $T(n,i)$ is a stopping time. In this situation, a natural choice for the processes $G(q)^n$ of (2.7) is $G(q)^n_t = m'_q(v^n_t)^q$. Again, the next lemma is shown in Section 4.

Lemma 2.3. *Let $(T(n,i))$ be a mixed renewal scheme.*

(a) *We have* (C), *and also $T(n,i)\to\infty$ a.s. as $i\to\infty$ if $1/v^n$ is a locally bounded process.*

(b) *Suppose further that $v^n \xrightarrow{\text{u.c.p.}} v$ where v is an $(\mathcal{F}_t^0)$-adapted càdlàg process v such that both v and v_- do not vanish, and that $m'_q<\infty$ for some $q\ge 1$* (this is always true for $q=1$). *Then* (E(q)) *holds with $G(p)_t = m'_p(v_t)^p$ for all $p\in[0,q]$, and* (D(p)) *holds for all $p\in(0,q]$ with $a(p)_t = \frac{m'_p}{m'_1}(v_t)^{p-1}$.*

(c) *Under the assumptions of* (b), *and if $\sqrt{r_n}\,(v^n - v)\xrightarrow{\text{u.c.p.}} 0$, we have* (E$'$($q$)).

Remark 2.4. *In many practical situations, see e.g.* [10] *for concrete examples, one is led to consider simultaneously the scheme $(T(n,i))$, and another scheme $(T'(n,i))$ which is a "sub-scheme" of the first one; typically one considers only the even observations, that is we take $T'(n,i)=T(n,2i)$. Apart from* (1.5), *none of the previous assumptions is preserved by such a transformation, and in particular* (C) *is not satisfied in general by the new scheme, unless the original scheme is deterministic.*

In fact another assumption, which replaces (C), *has been introduced in* [8], *namely that $T(n,i)$ is a stopping time for the filtration $(\mathcal{F}_{(t-u_n)^+})_{t\ge 0}$, where $u_n = 1/r_n^{\xi}$ and $\xi\in(0,1)$. Such a property is obviously shared (upon a modification of ξ) by the sub-scheme $T'(n,i)$ above. In this setting, the particular case $p=2$ and $q=0$ of the forthcoming results has been proved (under stronger assumptions on X, though). We cannot fit this other assumption within our framework in general, but when* (D(q)) *holds for some $q>1$ then the same "localization procedure" as in Lemma* 4.1 *below implies that under this assumption we can modify the scheme so that* (C) *holds, and without modifying the asymptotics below, provided $\xi<1-1/q$.*

3. The results

3.1. *Limiting results*

As we will see, the behavior of $V^n(p)$ is not enough for our purposes, and we need to establish the convergence in probability for more general processes. For $p>0$ and $q\ge 0$ we set

$$V^n(p,q)_t = \sum_{i=1}^{N_t^n} \Delta(n,i)^{q+1-p/2}\left|\Delta_i^n X\right|^p, \tag{3.1}$$

so in particular $V^n(p)=V^n(p,0)$. In the regular sampling case $\Delta(n,i)=\Delta_n$ these processes all convey the same information, since $V^n(p,q)=\Delta_n^q V^n(p)$, but this is no longer the case in the irregular sampling case.

Our first result is a law of large numbers, which goes as follows:

Theorem 3.1. *Let $p\ge 1$ and $q\ge 0$. Assume* (A) *and* (D($q+1$)). *Then*

$$r_n^q V^n(p,q)_t \xrightarrow{\text{u.c.p.}} V(p,q)_t := m_p\int_0^t |\sigma_s|^p a(q+1)_s\,\mathrm{d}s. \tag{3.2}$$

In particular, under (A) *and* (C) *and* (1.5), *we have*

$$V^n(p)_t \xrightarrow{\text{u.c.p.}} V(p)_t = m_p\int_0^t |\sigma_s|^p\,\mathrm{d}s. \tag{3.3}$$

For applications, we need an associated central limit theorem. For its statement, we need to recall the notion of $\mathcal{F}^0$-stable convergence in law for a sequence of random variables (or processes) Y_n defined on $(\Omega,\mathcal{F},\mathbb{P})$, see [13] for more details. We say that Y_n *converge $\mathcal{F}^0$-stably in law* to Y, where Y is a variable defined on an extension $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ of $(\Omega, \mathcal{F}, \mathbb{P})$, if we have

$$\mathbb{E}\big(Zh(Y_n)\big) \to \widetilde{\mathbb{E}}\big(Zh(Y)\big): \quad Z \text{ bounded } \mathcal{F}^0\text{-measurable, } h \text{ continuous bounded.} \tag{3.4}$$

There are in fact two versions for the CLT. The first is associated with (3.3), and is thus the most useful in practice, and it also holds for (3.2) when $q > 0$ under a strong additional assumption:

Theorem 3.2. *Let $p \geq 2$, and assume (A) when $p = 2$ and (B) when $p > 2$. Let $q \geq 0$ and assume one of the following two sets of hypotheses*:

(i) $q = 0$ *and* (D(2));
(ii) $q > 0$ *and* (D($q + 1$)) *and* (D($2q + 2$)) *and* (D′($q + 1$)).

Then the processes $\sqrt{r_n}(r_n^q V^n(p, q) - V(p, q))$ converge $\mathcal{F}^0$-stably in law to

$$\overline{V}(p, q)_t = \sqrt{m_{2p} - m_p^2} \int_0^t |\sigma_s|^p \sqrt{a(2q + 2)_s}\, \mathrm{d}W'_s, \tag{3.5}$$

where W' is a standard Brownian motion, defined on an extension of $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ and independent of $\mathcal{F}$. Moreover, conditionally on $\mathcal{F}$, the process $\overline{V}(p, q)$ is a continuous centered Gaussian martingale with (*conditional*) *variance at time t*:

$$\big(m_{2p} - m_p^2\big) \int_0^t |\sigma_s|^{2p} a(2q + 2)_s\, \mathrm{d}s. \tag{3.6}$$

As mentioned before, (D′($q + 1$)) is a very strong assumption, which for example is never satisfied by mixed renewal schemes unless the variables $\varepsilon(n, i)$ are constant. So when $q > 0$ the previous CLT hardly applies.

In practice we usually want to estimate $V(p)_t$, so the above seems enough because we can use the set (i) of assumptions for this. However, as we will see in Section 3.2 below, we need also an estimate for the conditional variance (3.6) when $q = 0$, that is for $\frac{m_{2p}-m_p^2}{m_{2p}} V(2p, 1)$: for this, we may use $(1 - \frac{m_p^2}{m_{2p}}) r_n V^n(2p, 1)_t$ by virtue of (3.2). This is why we have introduced the processes $V^n(p, q)$ for $q > 0$. But now, for asserting the quality of the latter estimator, we need a CLT for the processes $V^n(p, q)$ under reasonable assumptions, weaker than (ii) above. This is achieved in the following result:

Theorem 3.3. *Let $p \geq 2$ and $q \geq 0$. Assume (B) and that the sampling scheme satisfies (E′($2q + 2$)) and (D($2q + 2$)), and that the $(\mathcal{F}_t^0)$-adapted processes $G(p)$ for $p \in [1, q + 1]$ are Itô semimartingales with the same properties as the process σ in Assumption (B). Then the processes $\sqrt{r_n}(r_n^q V^n(p, q) - V(p, q))$ converge $\mathcal{F}^0$-stably in law to*

$$\overline{V}(p, q)_t = \int_0^t |\sigma_s|^p \sqrt{a(2q + 2)_s w(p, q)_s}\, \mathrm{d}W'_s, \tag{3.7}$$

where W' is as in Theorem 3.2 and

$$w(p, q)_t = m_{2p} - m_p^2 \frac{2G(q + 2)_t G(q + 1)_t G(1)_t - (G(q + 1)_t)^2 G(2)_t}{G(2q + 2)_t (G(1)_t)^2}. \tag{3.8}$$

One may check that $2G(q + 2)G(q + 1)G(1) - (G(q + 1))^2 G(2) \leq G(2q + 2)(G(1))^2$ always, because $q \mapsto G(q)_t$ has the same structure as the moments of a random variable. Thus $w(p, q)_t \geq m_{2p} - m_p^2$, and this inequality is strict in general, unless $q = 0$ of course. When $q = 0$ this theorem is a special case of Theorem 3.2, with stronger assumptions on the sampling scheme.

Remark 3.4. *By Lemma 2.2, (E′($2q + 2$)) implies (D(p)) for all $p < 2q + 2$, but not necessarily for $p = 2q + 2$, and this is why we add the assumption (D($2q + 2$)).*

Remark 3.5. *When $q > 0$ we apparently have two different limits for the same sequence of processes, in the two previous theorems. However if both* (E($2q+2+\varepsilon$)) *and* (D′($q+1$)) *hold for some $q > 0$, one may check that $G(p)^n_t - (G(1)^n_t)^p \stackrel{\text{u.c.p.}}{\longrightarrow} 0$ for all $p \le 2q+2$, and this implies $G(p)_t = (G(1)_t)^p$, which in turn yields $w(p,q)_t = m_{2p} - m_p^2$. When the scheme is a mixed renewal scheme, this also implies that the variables $\varepsilon(n,i)$ are in fact all equal to a constant, and the scheme is strongly predictable.*

Remark 3.6. *We have considered $p > 1$ and $p \ge 2$ respectively in the theorems above: this considerably simplifies the proofs, but the same results are also available when $p \in (0,1]$ for Theorem* 3.1 *and $p \in (1,2)$ for Theorems* 3.2 *and* 3.3. *The last two theorems also hold for $p \in (0,1]$ when the processes σ and σ_- do not vanish. Note also that when $p = 2$, Theorem* 3.10 *holds under* (A) *instead of* (B), *but here again the proof is complicated.*

Remark 3.7. *Both CLTs above have a multidimensional extension, in the sense that we can consider a finite family $(p_j, q_j)_{1\le j\le d}$ of indices with $p_j \ge 2$ and $q_j \ge 0$. We then have the $\mathcal{F}^0$-stable convergence in law of the d-dimensional processes with components $\sqrt{r_n}(V^n(p_j,q_j) - V(p_j,q_j))$. In the setting of Theorem* 3.2 *for example, the limit is, conditionally on $\mathcal{F}$, a continuous centered d-dimensional Gaussian martingale on an extension of the space, with covariance at time t between the jth and kth components given by*

$$(m_{p_j+p_k} - m_{p_j} m_{p_k}) \int_0^t |\sigma_s|^{p_j+p_k} a(q_j+q_k+2)_s \, \mathrm{d}s,$$

to be compared with (3.6).

Remark 3.8. *There is another easy extension to more general test functions. Namely, instead of* (3.1) *we can consider the process*

$$V^n(f,q)_t = \sum_{i=1}^{N^n_t} \Delta(n,i)^{q+1} f\bigl(\Delta^n_i X/\sqrt{\Delta(n,i)}\bigr), \tag{3.9}$$

where f is a function on $\mathbb{R}$ with at most polynomial growth and some smoothness properties: it should be C^1 for the law of large numbers, and C^2 for the CLT. Let ρ_a denote the normal law $\mathcal{N}(0,a^2)$ and $\rho_a(f)$ the integral of f with respect to ρ_a. The law of large numbers reads as Theorem 3.1, *with the limit $\int_0^t \rho_{\sigma_s}(f) a(q+1)_s \, \mathrm{d}s$, and the CLTs are as Theorems* 3.2 *or* 3.3: *for example, for the former theorem, for the limiting conditional variance one should replace* (3.6) *by*

$$\int_0^t \bigl(\rho_{\sigma_s}(f^2) - \rho_{\sigma_s}(f)^2\bigr) a(2q+2)_s \, \mathrm{d}s.$$

Remark 3.9. *We can even extend all the results to a d-dimensional underlying process X of the form* (1.1), *with now W a d-dimensional Brownian motion. Taking test functions f on $\mathbb{R}^d$, we can consider the processes $V^n(f,q)$ of* (3.9), *and derive the same results as in the previous remark (now, if a is a $d\times d$ matrix, ρ_a denotes the d-dimensional law $\mathcal{N}(0, aa^\star)$). The proofs are exactly the same, with slightly more cumbersome notation.*

However, this trivial extension requires the observation times $T(n,i)$ to be the same for all components of X: this is a fundamental restriction, since when the observations are irregularly spaced they also are typically non-synchronous: this is why we have not treated explicitly this case here.

3.2. *Estimation of the integrated volatility and other powers*

In practice, one wants to estimates the variable $V(p)_t$ at some time $t > 0$, mostly when $p = 2$, but the case $p = 4$ is also of interest. An estimator is $V^n(p)_t$, and Theorem 3.2 gives a central limit theorem for this estimator: it says that the normalized estimation error $\sqrt{r_n}(V^n(p)_t - V(p)_t)$ is centered normal, conditionally on $\mathcal{F}^0$. Even for regular sampling schemes, for which $a(2)_s = 1$, this result is not directly usable in practice for deriving confidence intervals for example, because $\int_0^t |\sigma_s|^{2p}\,\mathrm{d}s$ is a priori unknown. But here things are worse because the process $a(2)$ is also unknown, and in fact we also ignore r_n. However, we can "standardize" by considering the variable

$$T_t^n = \sqrt{\frac{m_{2p}}{(m_{2p}-m_p^2)V^n(2p,1)_t}}\big(V^n(p)_t - V(p)_t\big). \tag{3.10}$$

Then the following holds:

Theorem 3.10. *Let $p \geq 2$. Assume (D(2)). Assume also (A) when $p=2$ and (B) otherwise. Then for any $t>0$ such that $\int_0^t |\sigma_s|^{2p} a(2)_s\,\mathrm{d}s > 0$ a.s., the sequence T_t^n converges in law (and even $\mathcal{F}^0$-stably in law) to $\mathcal{N}(0,1)$.*

All ingredients in the definition of T_t^n are known to the statistician, except of course the quantity $V(p)_t$ to be estimated. Therefore we can derive (asymptotic) confidence intervals for $V(p)_t$ in a straightforward way.

If one wants to estimate $V(p,q)_t$ one can also use Theorems 3.2 or 3.3, depending on the assumptions on the scheme. However, so far we have no estimators for the conditional variance $\int_0^t |\sigma_s|^{2p} a(2q+2)_s w(p,q)_s\,\mathrm{d}s$ in the second theorem, hence we have no feasible statistics for deriving confidence intervals for $V(p,q)_t$ when $q>0$, except under the strong assumption (D′$(q+1)$).

This theorem should be compared with Theorem 4.1 of [1], which states exactly the same result, when the process X has no drift and the volatility σ_t is possibly random but independent of W: in this paper the authors consider only deterministic sampling schemes but, if their assumptions are not formally comparable to ours, they are in some sense significantly weaker: namely, they suppose that $\min_{i:i\leq N_t^n} \Delta(n,i)^{2/3}/\pi_t^n \to \infty$ as $n\to\infty$. In [16], Theorem 3.1, is also the same result as above, for $p=2$ and again when there is no drift: in that paper the sampling scheme is somewhat similar to a mixed renewal scheme, and the precise assumptions are not directly comparable to ours, since basically the authors assume the existence of a CLT for the conditionally centered variables $\Delta(n,i)$. Note also that this theorem is proved when $p=2$ in [15] for deterministic schemes which satisfy an assumption slightly stronger than (D(2)) plus (D′(1)).

An interesting – and crucial – feature of our result, as well as in the results of [1], is that the properties of the observation scheme are not showing explicitly in the result itself, and in particular the knowledge of the process $a(2)$ and even of the rates r_n is not necessary to apply it. This is a good thing because those are generally unknown, whereas it is also dangerous because one might be tempted to use the property that T_t^n is (approximately) $\mathcal{N}(0,1)$ without checking that the assumptions on the sampling scheme are satisfied, here as well as in [1]. When they are satisfied, and although the rates r_n do not explicitly show up, these rates still govern the "true" speed of convergence.

Once more, r_n is unknown, but N_t^n is of course known. Then as soon as we have also (D(0)) (for example when (E(2)) holds), then $N_t^n/r_n \xrightarrow{\mathbb{P}} \int_0^t a(0)_s\,\mathrm{d}s$, and so the actual rate of convergence for the estimators is also $1/\sqrt{N_t^n}$, as it should be.

3.3. *Some remarks about optimality*

In this part we consider only the question of estimating the integrated volatility $V(2)_t$. The problem of asymptotic efficiency for estimators of $V(2)_1$ is unsolved yet, as far as we know, and even the definition of asymptotic efficiency is not quite clear, when the volatility is genuinely random, and even for regular sampling.

However, one can examine the asymptotic behavior of our estimators in the very special situation where $X_t = \sigma W_t$ and σ is a constant, so $V(2)_t = \sigma^2 t$. In this case we have a parametric model, and a very simple one indeed, and the MLE for estimating $\sigma^2 t$ is asymptotically efficient. For regular sampling the MLE is $\frac{t/\Delta_n}{[t/\Delta_n]} V^n(2)_t$, so obviously the realized quadratic variation $V^n(2)_t$ is also asymptotically efficient.

When the sampling is non-random but not regular, the MLE for estimating σ^2 is

$$V_n' = \frac{t}{N_t^n}\sum_{i=1}^{N_t^n} \frac{1}{\Delta(n,i)}\big|\Delta_i^n X\big|^2.$$

This is again asymptotically efficient, and for each n the variance of $\sqrt{N_t^n}(V_n' - \sigma^2 t)$ is $2\sigma^4 t^2$. Now, if further (D(2)) holds, the asymptotic variance of $\sqrt{r_n}(V^n(2,0)_t - \sigma^2 t)$ is $2\sigma^4\int_0^t a(2)_s\,\mathrm{d}s$, and if (D(0)) also holds we have $N_t^n/r_n \to$

$\int_0^t a(0)_s \, \mathrm{d}s$. Therefore, under both (D(0)) and (D(2)), the two (normalized) asymptotic variances Σ_t and Σ'_t of $V^n(2)_t$ and V'_n satisfy:

$$\Sigma_t = \alpha_t \Sigma'_t, \quad \text{where } \alpha_t = \frac{1}{t^2}\left(\int_0^t a(2)_s \, \mathrm{d}s\right)\left(\int_0^t a(0)_s \, \mathrm{d}s\right).$$

If we use (2.5) with $q = 0$ and $p = 1$ and $q' = 2$ and go the limit, we see that $\alpha_t \geq 1$, as it should be because V'_n is asymptotically efficient. It may happen that $\alpha_t = 1$, of course, but this is equivalent to saying that $A(0)^n_t A(2)^n_t - (A(1)^n_t)^2 \to 0$, and the equality $A(0)^n_t A(2)^n_t = (A(1)^n_t)^2$ implies that the $\Delta(n,i)$'s for $i \leq N^n_t$ are all equal (by Cauchy–Schwarz inequality): therefore *the realized volatility $V^n(2)_t$ is asymptotically efficient only when the sampling scheme is "asymptotically" a regular sampling*.

At this stage, one might wonder why we do not use V'_n instead of $V^n(2)_t$ in general. This is because, if we assume for example (D(0)), the sequence V'_n converges in probability to the variable

$$t \frac{\int_0^t \sigma_s^2 a(0)_s \, \mathrm{d}s}{\int_0^t a(0)_s \, \mathrm{d}s},$$

which is different from $\int_0^t \sigma_s^2 \, \mathrm{d}s$ unless either σ_s or $a(0)_s$ do not depend on time.

4. More about sampling schemes

The aim of this section is to prove Lemmas 2.1, 2.2 and 2.3, and it also contains a few complementary results, useful for the proofs of the main theorems. Below, and in all the paper, K denotes a constant varying from line to line and may depend on the characteristics of the process or of the sampling scheme (we write K_p if it depends on an additional parameter p).

Proof of Lemma 2.1. Since (D(1)) always holds under (C) and (1.5), the final claim follows from the estimate (2.5) and from classical results on the tightness and convergence of increasing processes.

As for (2.5), this follows from Hölder's inequality as such:

$$\begin{aligned} A(p)^n_t - A(p)^n_s &= r_n^{p-1} \sum_{i=N^n_s+1}^{N^n_t} \Delta(n,i)^p \\ &\leq \left(r_n^{q-1} \sum_{i=N^n_s+1}^{N^n_t} \Delta(n,i)^q\right)^{(q'-p)/(q'-q)} \left(r_n^{q'-1} \sum_{i=N^n_s+1}^{N^n_t} \Delta(n,i)^{q'}\right)^{(p-q)/(q'-q)}. \end{aligned}$$

□

Before proving Lemma 2.2, we introduce strengthened versions of Assumptions (D(q)) and (E(q)), which we will also use later on:

***Assumption* (SD(q)).** *We have (D(q)), and $\int_0^t a(q)_s \, \mathrm{d}s$ is bounded for each t, and the convergence in (2.4) takes place in $\mathbb{L}^1$.*

***Assumption* (SE(q)).** *We have (E(q)), and for some constant $L \geq 1$ we have $G(q)^n \leq L$ and $1/G(1)^n \leq L$ (hence $G(q) \leq L$ and $1/G(1) \leq L$ as well).*

Lemma 4.1. (a) *Let $1 \leq q_1 < \cdots < q_r$ and assume (D(q_j)) for $j = 1, \ldots, r$. For all $t > 0$, $k \geq 1$, one can find a scheme $(T_{k,t}(n,i))$ satisfying (SD(q_j)) for $j = 1, \ldots, r$ (with the processes $a_{k,t}(q_j)$), and subsets $\Omega_{k,t}$ and $\Omega_{k,n,t}$ of Ω, such that*

$$\left.\begin{aligned} &\mathbb{P}(\Omega_{k,t}) \geq 1 - 1/k, \qquad \liminf_n \mathbb{P}(\Omega_{k,n,t}) \geq 1 - 1/k, \\ &\omega \in \Omega_{k,n,t} \quad \Rightarrow \quad [0,t] \cap \{T(n,i)(\omega)\colon i \geq 1\} = [0,t] \cap \{T_{k,t}(n,i)(\omega)\colon i \geq 1\} \end{aligned}\right\} \tag{4.1}$$

and, for $j = 1, \dots, r$,

$$s \le t, \omega \in \Omega_{k,t} \quad \Rightarrow \quad a(q_j)(\omega)_s = a_{k,t}(q_j)(\omega)_s. \tag{4.2}$$

(b) *Assume* (E(q)) *for some* $q > 1$. *For all* $t > 0$, $k \ge 1$, *one can find a scheme* $(T_{k,t}(n,i))$ *satisfying* (SE(q)) (*with the processes* $G_{k,t}(p)$), *and subsets* $\Omega_{k,t}$ *and* $\Omega_{k,n,t}$ *of* Ω, *such that we have* (4.1) *and, for* $p \in [0, q]$,

$$s \le t, \omega \in \Omega_{k,t} \quad \Rightarrow \quad G(p)(\omega)_s = G_{k,t}(p)(\omega)_s. \tag{4.3}$$

Proof. (a) We may of course assume that $r_n \ge 1$ for all n, and also that $q_r > 1$ because (1.5) is enough to imply (SD(1)). Fix t and k. We can find $a > 1$ such that the set $\Omega_{k,t} = \bigcap_{1 \le j \le r} \{\int_0^{t+1} a(q_j)_s \, ds \le a - 1\}$ satisfies $\mathbb{P}(\Omega_{k,t}) \ge 1 - 1/k$, and thus by (D($q_j$)) the sets $\Omega_{k,n,t} = \bigcap_{1 \le j \le r} \{A(q_j)^n_{t+1} \le a\}$ satisfy $\liminf_n \mathbb{P}(\Omega_{k,n,t}) \ge 1 - 1/k$. Then we define the sampling scheme $(T_{k,t}(n,i))$ in the following way: we set $\alpha_n = (a r_n)^{1/q_r}/r_n$, which goes to 0; when $\alpha_n \ge 1$ we simply take $\Delta_{k,t}(n,i) = \Delta(n,i)$ (so the observation times are unchanged), and when $\alpha_n < 1$, that is for all n large enough, we put

$$\Delta_{k,t}(n,i) = \begin{cases} \Delta(n,i), & \text{if } \Delta(n,j) \le \alpha_n \text{ for all } j = 1, \dots, i, \\ 1/r_n, & \text{otherwise.} \end{cases}$$

We associate the processes $A_{k,t}(q)^n$ by (2.2). It is obvious that the scheme $(T_{k,t}(n,i))$ satisfies (C) and (1.5). On the set $\Omega_{k,n,t}$ we have $\Delta(n,i) \le \alpha_n$ for all i with $T(n,i) \le t+1$, hence $\Delta_{k,t}(n,i) = \Delta(n,i)$ for those i's, and we conclude the last part of (4.1). Moreover $A_{k,t}(q_j)^n_s = A(q_j)^n_s$ as long as $\Delta(n,i) \le \alpha_n$ for all i with $T(n,i) \le s$, and afterwards $A_{k,t}(q_j)^n_s$ increases in s like $[s r_n]/r_n$, so obviously the new scheme satisfies (D(q_j)) with a function $a_{k,t}(q_j)$ which coincides with $a(q_j)$ on $[0,t]$ on the set $\Omega_{k,t}$, giving (4.2). Finally, by construction $A_{k,t}(q_j)^n_t \le a + t$ as soon as $\alpha_n < 1$, so in fact this new scheme satisfies (SD(q_j)) for $j = 1, \dots, r$.

(b) Again, we fix $t > 0$ and $k \ge 1$. Since $G(q)$ is càdlàg $(\mathcal{F}^0_t)$-adapted, there is $a_k > 1$ such that $R_k = (t+1) \wedge \inf(s\colon G(q)_s + 1/G(1)_s \ge a_k)$ is an $(\mathcal{F}^0_t)$-stopping time and $\Omega_{k,t} = \{R_k > t\}$ satisfies $\mathbb{P}(\Omega_{k,t}) \ge 1 - 1/k$. Next,

$$R^n_k = R_k \wedge \inf\left(s\colon \left|G(q)^n_s - G(q)_s\right| + \left|G(1)^n_s - G(1)_s\right| \ge \frac{1}{2a_k}\right)$$

is an $(\mathcal{F}_t)$-stopping time and $\Omega_{k,n,t} = \{R^n_k > t\}$ satisfies the condition in (4.1) by (E(q)). Then the new scheme $(T_{k,t}(n,i))$ is defined by

$$\Delta_{k,t}(n,i) = \begin{cases} \Delta(n,i), & \text{if } T(n,i-1) < R^n_k, \\ 1/r_n, & \text{otherwise.} \end{cases}$$

(Note that it depends on t, because R^n_k implicitly depends on t.) It is obvious that the scheme $(T_{k,t}(n,i))$ satisfies (1.5), and (C) because all R^n_k are $(\mathcal{F}_t)$-stopping time, and the last part of (4.1) is also obvious.

Now, $r^p_n E(\Delta_{k,t}(n,i)^p \mid \mathcal{F}_{T_{k,t}(n,i-1)})$ is equal to $r^p_n E(\Delta(n,i)^p \mid \mathcal{F}_{T(n,i-1)})$ on the set $\{T(n,i-1) < R^n_k\}$, and to 1 otherwise. Therefore we can take, as the $(\mathcal{F}_t)$-optional process $G_{k,t}(p)^n$ satisfying (2.7) relative to the new scheme, the process which equals $G(p)^n$ on $[0, R^n_k)$ and 1 on $[R^n_k, \infty)$. Then clearly $G_{k,t}(q)^n \le a_k + 1$ and $1/G_{k,t}(1)^n \le 2a_k$. Moreover $\mathbb{P}(R^n_k = R_k) \to 1$, so the new scheme satisfies (2.9) for all $p \in [1, q]$, with the limit $G_{k,t}(p)$ equal to $G(p)$ on the interval $[0, R_k)$ and to 1 on $[R_k, \infty)$. Therefore the new scheme satisfies (SE(q)), and (b) is proved. □

Proof of Lemma 2.2. A standard localization argument, based upon the previous lemma, yields that it is enough to prove the result when (SE(q)) holds instead of (E(q)). So, in view of (2.8), we assume $G(p)^n \le L$ for all $p \in [0, q]$ and $1/G(1)^n \le L$ for some constant $L \ge 1$. We can also assume $r_n \ge 1$.

(a) In a first step, we show that

$$\mathbb{E}(N^n_t + 1) \le K_t r_n, \tag{4.4}$$

where K_t is a constant depending on t. We put $C_t = (t+1) \vee (2L^2)^{1/(q-1)}$. We also define a new sampling scheme by $\Delta'(n,i) = \Delta(n,i) \wedge C_t$, with the associated variables N'^n_t and $T'(n,i)$ and $G'(p)^n_{T(n,i-1)} =$

$\mathbb{E}(r_n^p \Delta'(n,i)^p | \mathcal{F}_{T'(n,i-1)})$. This new scheme does not necessarily satisfies (E(q)). However, it does satisfies $T(n,i) = T'(n,i)$ and $\Delta'(n,i) = \Delta(n,i)$ for all i such that $T(n,i) \le t$, and also $T'(n,i) > t$ if $T(n,i) > t$, because $C_t > t$. In particular, we have $N_t'^n = N_t^n$. Moreover, on the set $\{T(n,i-1) \le t\}$ we have

$$G(1)^n_{T(n,i-1)} - G'(1)^n_{T(n,i-1)} = r_n \mathbb{E}\big(\big(\Delta(n,i) - C_t\big)^+ | \mathcal{F}_{T(n,i-1)}\big) \le \frac{r_n}{C_t^{q-1}} \mathbb{E}\big(\Delta(n,i)^q | \mathcal{F}_{T(n,i-1)}\big) \le \frac{L}{C_t^{q-1}} \le \frac{1}{2L},$$

where the first inequality comes from Markov's inequality, the second one from $r_n \ge 1$, and the last one from $C_t^{q-1} \ge 2L^2$. Thus $G'^n(1)_{T(n,i-1)} \ge 1/2L$ if $i \le N_t^n + 1$, hence

$$A_t^n = \sum_{i=1}^{N_t^n+1} \frac{\Delta'(n,i)}{G'(1)^n_{T(n,i-1)}} \le 2L \sum_{i=1}^{N_t^n+1} \Delta'(n,i) \le 2L(t + C_t).$$

We have

$$\begin{aligned}\mathbb{E}\big(A_t^n\big) &= \sum_{i\ge 1} \mathbb{E}\left(\frac{\Delta'(n,i)}{G'(1)^n_{T(n,i-1)}} 1_{\{T(n,i-1)\le t\}}\right) \\ &= \sum_{i\ge 1} \mathbb{E}\left(1_{\{T(n,i-1)\le t\}} \mathbb{E}\left(\frac{\Delta'(n,i)}{G'(1)^n_{T(n,i-1)}} \Big| \mathcal{F}_{T(n,i-1)}\right)\right) \\ &= \frac{1}{r_n} \sum_{i\ge 1} \mathbb{E}(1_{\{T(n,i-1)\le t\}}) = \frac{1}{r_n} E\big(N_t^n + 1\big).\end{aligned}$$

Thus we deduce (4.4) with $K_t = 2L(t + C_t)$.

(b) Next we set, for $p \in [0,q)$ fixed:

$$B_t^n = \sum_{i=1}^{N_t^n+1} r_n^{p-1} \Delta(n,i)^p, \qquad C_t^n = \sum_{i=1}^{N_t^n+1} \Delta(n,i) \frac{G(p)^n_{T(n,i-1)}}{G(1)^n_{T(n,i-1)}},$$

$$\zeta_i^n = r_n^{p-1} \Delta(n,i)^p - \Delta(n,i) \frac{G(p)^n_{T(n,i-1)}}{G(1)^n_{T(n,i-1)}}.$$

We have $B^n_{(t-\pi_t^n)^+} \le A(p)_t^n \le B_t^n$ and $B_t^n - C_t^n = \sum_{i=1}^{N_t^n+1} \zeta_i^n$. Hence, in order to prove (2.4) with $a(p)$ given by (2.10), it is enough to prove the following two properties:

$$C_t^n \stackrel{\text{u.c.p.}}{\longrightarrow} \int_0^t \frac{G(p)_s}{G(1)_s} \, \mathrm{d}s, \qquad D_t^n = \sup_{s\le t} \left| \sum_{i=1}^{N_s^n+1} \zeta_i^n \right| \stackrel{\mathbb{P}}{\longrightarrow} 0. \tag{4.5}$$

(c) For the first property above we observe that (2.3) implies that for any càdlàg process H and any integer $k \ge 0$ and any sequence H^n of processes having $H^n \stackrel{\text{u.c.p.}}{\longrightarrow} H$, we have

$$\sum_{i=1}^{N_t^n+k} \Delta(n,i) H^n_{T(n,i-1)} \stackrel{\text{u.c.p.}}{\longrightarrow} \int_0^t H_s \, \mathrm{d}s. \tag{4.6}$$

This applied $H^n = G(p)^n / G(1)^n$ and $H = G(p)/G(1)$ and $k = 1$ yields the first part of (4.5).

(d) The variables ζ_i^n are martingale differences, relative to the discrete time filtration $(\mathcal{F}_{T(n,i)})_{i\geq 0}$, with respect to which the variable N_t^n+1 is a stopping time. Then, letting $r=2\wedge\frac{q}{p}\wedge q$, so $r\in(1,2]$, the Burkholder–Davis–Gundy inequality and the inequality $(\sum u_n)^{r/2}\leq\sum u_n^{r/2}$ when $u_n\geq 0$ yield

$$\mathbb{E}\big((D_t^n)^r\big)\leq K\mathbb{E}\left(\left(\sum_{i=1}^{N_t^n+1}(\zeta_i^n)^2\right)^{r/2}\right)\leq K\mathbb{E}\left(\sum_{i=1}^{N_t^n+1}|\zeta_i^n|^r\right).$$

Moreover, by (2.7) and (2.8), we have (with K below depending on r, p, L):

$$\begin{aligned}E\big(|\zeta_i^n|^r\,|\mathcal{F}_{T(n,i-1)}\big)&\leq 2^{r-1}\mathbb{E}\big(r_n^{r(p-1)}\Delta(n,i)^{pr}+L^{2r}\Delta(n,i)^r\,|\mathcal{F}_{T(n,i-1)}\big)\\&\leq K\mathbb{E}\left(\frac{1}{r_n^r}G(pr)^n_{T(n,i-1)}+\frac{1}{r_n^r}G(r)^n_{T(n,i-1)}\right)\leq\frac{K}{r_n^r},\end{aligned}$$

the last inequality coming from $pr\leq q$ and $r\leq q$. Then in view of (4.4),

$$E\big((D_t^n)^r\big)\leq\frac{K}{r_n^r}\mathbb{E}\big(N_t^n+1\big)\leq\frac{K}{r_n^{r-1}},$$

Then, recalling $r>1$, the second part of (4.5) holds, and the proof is complete. □

Proof of Lemma 2.3. (a) Property (C) comes from the fact that $T(n,i)$ depends only on $v^n_{T(n,i-1)}$ and $T(n,i-1)$, which are $\mathcal{F}_{T(n,i-1)}$-measurable, and on $\varepsilon(n,i)$, which is independent of $\mathcal{F}^0$. The second claim is obvious because $\sum_{i\geq 1}\varepsilon(n,i)=\infty$ a.s.

(b) By a standard localization argument and the subsequence principle, we may assume that in fact that $Y=\sup_{t,n}(v_t^n+1/v_t^n)$ is a finite positive variable, and $\sup_t|v_t^n-v_t|\to 0$ for each ω. Then we have $\varepsilon(n,i)/Yr_n\leq\Delta(n,i)\leq Y\varepsilon(n,i)/r_n$ and, since the $\varepsilon(n,i)$ are i.i.d. positive, it is obvious that (1.5) holds.

In the present situation (2.7) reads as $G(p)^n_{T(n,i-1)}=m'_p(v^n_{T(n,i-1)})^p$, so we can choose $G(p)^n=m'_p(v^n)^p$ and (2.9) holds for all $p\in[0,q]$ if $m'_q<\infty$, with $G(p)=m'_pv^p$, and (E(q)) holds. By Lemma 2.2 we have (D(p)) for any $p<q$, with $a(p)=G(p)/G(1)$.

It remains to prove (D(q)) for $q>1$ when $m'_q<\infty$, that is (2.4) with $a(q)$ as above. (1.5) implies $N_t^n\xrightarrow{\mathbb{P}}\infty$. Then the LLN yields

$$\frac{r_n^q}{N_t^n}\sum_{i=1}^{N_t^n}\big(\Delta(n,i)/v^n_{T(n,i-1)}\big)^q\xrightarrow{\mathbb{P}}m'_q.\tag{4.7}$$

In view of (2.11) and $G(1)=m'_1v$, we deduce that

$$r_n^{q-1}\sum_{i=1}^{N_t^n}\big(\Delta(n,i)/v^n_{T(n,i-1)}\big)^q\xrightarrow{\mathbb{P}}\frac{m'_q}{m'_1}\int_0^t\frac{1}{v_s}\,\mathrm{d}s.$$

This, being true for all t, implies that for an arbitrary càdlàg process H we have

$$r_n^{q-1}\sum_{i=1}^{N_t^n}\Delta(n,i)^q\frac{H_{T(n,i-1)}}{(v^n_{T(n,i-1)})^q}\xrightarrow{\mathbb{P}}\frac{m'_q}{m'_1}\int_0^t\frac{H_s}{v_s}\,\mathrm{d}s.\tag{4.8}$$

Then, plugging $H=v^q$ above and using the convergence $v_t/v_t^n\to 1$ (uniform in t for each ω) gives (2.4) with $a(q)=\frac{m'_q}{m'_1}v^{q-1}$.

(c) Since $G(p)^n_t=m'_p(v_t^n)^p$ and $G(p)=m'_p(v_t)^p$, the result is obvious. □

5. Two auxiliary stable convergence results

This section is devoted to proving two closely related key results on stable convergence in law. They are standard, except that here we have the $\mathcal{F}^0$-stable convergence in law only (the $\mathcal{F}$-stable convergence does not hold in general), and we need to be careful.

The setting is as before: we have the process X and the sampling scheme satisfying (1.5) and (C), relative to the filtration $(\mathcal{F}_t^0)$. We consider the σ-fields $\mathcal{G}_i^n = \mathcal{F}_{T(n,i)}$ and $\mathcal{G}_i'^n = \mathcal{F}_{T(n,i)} \vee \sigma(\Delta(n, i+1))$, and we denote by $\mathbb{E}_i^n$ and $\mathbb{E}_i'^n$, respectively, the conditional expectation with respect to $\mathcal{G}_i^n$ and $\mathcal{G}_i'^n$. Set

$$\beta_i^n = \frac{\sigma_{T(n,i-1)}\Delta_i^n W}{\sqrt{\Delta(n,i)}}, \tag{5.1}$$

$$U^n(p,q)_t = r_n^{q+1/2} \sum_{i=1}^{N_t^n} \Delta(n,i)^{q+1} \big(\big|\beta_i^n\big|^p - m_p|\sigma_{T(n,i-1)}|^p\big). \tag{5.2}$$

Proposition 5.1. *Suppose that σ is càdlàg and bounded, and that (SD($2q+2$)) holds for some $q \geq 0$. Then if $p \geq 0$ the processes $U^n(p,q)$ converge $\mathcal{F}^0$-stably in law to the process*

$$\overline{U}(p,q)_t = \sqrt{m_{2p} - m_p^2} \int_0^t |\sigma_s|^p \sqrt{a(2q+2)_s}\, \mathrm{d}W_s', \tag{5.3}$$

where W' is a Brownian motion defined on an extension of $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ and independent of $\mathcal{F}$.

Proof. (a) Set

$$\zeta_i^n = r_n^{q+1/2} \Delta(n,i)^{q+1} \big(\big|\beta_i^n\big|^p - m_p|\sigma_{T(n,i-1)}|^p\big). \tag{5.4}$$

Assumption (C) and the definition of $\mathcal{G}_i'^n$ yield, for any $r > 0$:

$$\mathbb{E}_{i-1}'^n\big(\big|\beta_i^n\big|^r\big) = m_r |\sigma_{T(n,i-1)}|^r, \tag{5.5}$$

so the $\mathcal{G}_i'^n$-measurable variable ζ_i^n satisfies $\mathbb{E}_{i-1}'^n(\zeta_i^n) = 0$. Moreover, $\{N_t^n = i\} = \{T(n,i) \leq t < T(n,i+1)\}$, hence N_t^n is an $(\mathcal{G}_i'^n)$-stopping time. Then, should we want to prove the usual stable (i.e. the $\mathcal{F}$-stable) convergence in law we would apply Theorem IX.7.13 of [13], in its version for triangular arrays and associated with the basic martingale $Z = W$.

Since we only want to prove the $\mathcal{F}^0$-stable convergence, and since (C) holds, a close look at the proof of that theorem shows that it is enough to prove the following properties:

$$\sum_{i=1}^{N_t^n} \mathbb{E}_{i-1}'^n\big((\zeta_i^n)^2\big) \xrightarrow{\mathbb{P}} \big(m_{2p} - m_p^2\big) \int_0^t |\sigma_s|^{2p} a(2q+2)_s \,\mathrm{d}s, \tag{5.6}$$

$$\sum_{i=1}^{N_t^n} \mathbb{E}_{i-1}'^n\big((\zeta_i^n)^2 1_{\{|\zeta_i^n|>\varepsilon\}}\big) \xrightarrow{\mathbb{P}} 0, \tag{5.7}$$

$$\sum_{i=1}^{N_t^n} \mathbb{E}_{i-1}'^n\big(\zeta_i^n \Delta_i^n M\big) \xrightarrow{\mathbb{P}} 0 \tag{5.8}$$

for all t and $\varepsilon > 0$ and $M \in \mathcal{M}$, where $\mathcal{M}$ is the union of the set $\mathcal{M}_1$ of all bounded $(\mathcal{F}_t^0)$-martingales which are orthogonal to W, and of the singleton $\{W\}$.

(b) We have $\mathbb{E}_{i-1}'^n(\zeta_i^n \Delta_i^n M) = 0$ when $M = W$, because $x \mapsto |x|^p$ is an even function. The same equality holds when $M \in \mathcal{M}_1$: see Lemma 6.8 of [10], the proof works for the conditional expectation $\mathbb{E}_{i-1}'^n$ here because of (C). So (5.8) is proved.

(c) Next, we prove (5.7). Under (SD(r)) the convergence in (2.4) holds in $\mathbb{L}^1$, locally uniformly in time. It follows that the supremum of the jump sizes of $A(r)^n$ over $[0,t]$, which is $r_n^{r-1}(\pi_t^n)^r$, goes to 0 in $\mathbb{L}^1$ because the limit is continuous. In other words,

$$\text{(SD}(r)), \qquad t \geq 0 \quad \Rightarrow \quad \mathbb{E}\bigl(r_n^{r-1}\bigl(\pi_t^n\bigr)^r\bigr) \to 0. \tag{5.9}$$

We have $m_p|\sigma_s|^p \leq C$ for some constant C. Take any $\theta \in (0, \varepsilon/2C]$. Then if $|\zeta_i^n| > \varepsilon$ and $r_n^{q+1/2}(\pi_t^n)^{q+1} \leq \theta$, we must have $|\beta_i^n|^p > \varepsilon/2\theta$ and also $|\zeta_i^n| \leq 2r_n^{q+1/2}\Delta(n,i)^{q+1}|\beta_i^n|^p$. Therefore, with the notation $a_n(\theta) = \mathbb{P}(r_n^{q+1/2}(\pi_t^n)^{q+1} > \theta)$, we have for any $\eta > 0$:

$$\left.\begin{aligned}
&\mathbb{P}\Biggl(\sum_{i=1}^{N_t^n}\mathbb{E}'^n_{i-1}\bigl(\bigl(\zeta_i^n\bigr)^2 1_{\{|\zeta_i^n|>\varepsilon\}}\bigr) > \eta\Biggr) \\
&\quad\leq a_n(\theta) + \frac{4r_n^{2q+1}}{\eta}\sum_{i\geq 1}\mathbb{E}\bigl(\Delta(n,i)^{2q+2}\bigl|\beta_i^n\bigr|^{2p} 1_{\{|\beta_i^n|^{2p}>(\varepsilon/2\theta)^2\}} 1_{\{T(n,i)\leq t\}}\bigr) \\
&\quad\leq a_n(\theta) + \frac{K\theta^2 r_n^{2q+1}}{\eta\varepsilon^2}\sum_{i\geq 1}\mathbb{E}\bigl(\Delta(n,i)^{2q+2}\bigl|\beta_i^n\bigr|^{4p} 1_{\{T(n,i)\leq t\}}\bigr) \\
&\quad\leq a_n(\theta) + \frac{K\theta^2}{\eta\varepsilon^2}\mathbb{E}\bigl(A(2q+2)_t^n\bigr)
\end{aligned}\right\} \tag{5.10}$$

(recall that K varies from line to line; for the last inequality we use $\mathbb{E}'^n_{i-1}(|\beta_i^n|^{4p}) \leq K$ and (C)). Observe that $a_n(\theta) = \mathbb{P}(r_n^{2q+1}(\pi_t^n)^{2q+2} > \theta^2)$ goes to 0 by (5.9) applied with $r = 2q+2$. Then using the boundedness of $\mathbb{E}(A(2q+2)_t^n)$, and by choosing first θ small, then letting $n \to \infty$, we deduce (5.7).

(d) Finally $\mathbb{E}'^n_{i-1}((\zeta_i^n)^2) = r_n^{2q+1}\Delta(n,i)^{2q+2}(m_{2p} - m_p^2)|\sigma_{T(n,i-1)}|^{2p}$, whereas by (2.4)

$$\text{(D}(r)) \quad \Rightarrow \quad r_n^{r-1}\sum_{i=1}^{N_t^n} H_{T(n,i)}\Delta(n,i)^r \stackrel{\text{u.c.p.}}{\longrightarrow} \int_0^t H_s a(r)_s\,\mathrm{d}s \tag{5.11}$$

for any càdlàg process H. Since σ_t is càdlàg, we deduce (5.6) by taking $r = 2q+2$. $\square$

Recalling (2.7), our second result will be about the processes

$$U'^n(p,q)_t = \sqrt{r_n}\sum_{i=1}^{N_t^n}\Biggl(r_n^q\Delta(n,i)^{q+1}\bigl|\beta_i^n\bigr|^p - m_p|\sigma_{T(n,i-1)}|^p\frac{G(q+1)^n_{T(n,i-1)}}{G(1)^n_{T(n,i-1)}}\Delta(n,i)\Biggr). \tag{5.12}$$

Proposition 5.2. *Let $p \geq 2$ and $q \geq 0$. Suppose that σ is càdlàg and bounded, and that (SE($2q+2$)) and (SD($2q+2$)) hold. Then the processes $U'^n(p,q)$ converge $\mathcal{F}^0$-stably in law to the process*

$$\overline{U}'(p,q)_t = \int_0^t |\sigma_s|^p\sqrt{a(2q+2)_s w(p,q)_s}\,\mathrm{d}W'_s, \tag{5.13}$$

where $w(p,q)$ is defined by (3.8) and W' is as in Proposition 5.1.

Proof. Instead of (5.4), we set

$$\zeta_i^n = r_n^{q+1/2}\Delta(n,i)^{q+1}\bigl|\beta_i^n\bigr|^p - r_n^{1/2}m_p|\sigma_{T(n,i-1)}|^p\frac{G(q+1)^n_{T(n,i-1)}}{G(1)^n_{T(n,i-1)}}\Delta(n,i). \tag{5.14}$$

Recall that $G(p)^n \le L$ for all $p \in [0, 2q+2]$ and $1/G(1)^n \le L$. Then by taking first the conditional expectation with respect to $\mathcal{G}'^n_{i-1}$, and second with respect to $\mathcal{G}^n_{i-1}$ and by using (C) and the fact that σ is adapted to $(\mathcal{F}^0_t)$, we get

$$\left.\begin{aligned}
&\mathbb{E}^n_{i-1}(\zeta^n_i) = 0,\\
&\mathbb{E}^n_{i-1}\big((\zeta^n_i)^2\big) = \frac{|\sigma_{T(n,i-1)}|^{2p}}{r_n}\bigg(m_{2p} G(2q+2)^n_{T(n,i-1)} + \frac{m_p^2 G(q+1)^n_{T(n,i-1)}}{(G(1)^n_{T(n,i-1)})^2}\\
&\qquad \times \big(G(q+1)^n_{T(n,i-1)} G(2)^n_{T(n,i-1)} - 2G(q+2)^n_{T(n,i-1)} G(1)^n_{T(n,i-1)}\big)\bigg).
\end{aligned}\right\} \tag{5.15}$$

At this stage we can reproduce the argument of the previous proof (with the σ-fields $\mathcal{G}^n_i$ instead of $\mathcal{G}'^n_i$), to the effect that it is enough to prove

$$\left.\begin{aligned}
&\sum_{i=1}^{N^n_t} \mathbb{E}^n_{i-1}\big((\zeta^n_i)^2\big) \xrightarrow{\mathbb{P}} \int_0^t |\sigma_s|^{2p} a(2q+2)_s w(p,q)_s \,\mathrm{d}s,\\
&\sum_{i=1}^{N^n_t} \mathbb{E}^n_{i-1}\big((\zeta^n_i)^2 1_{\{|\zeta^n_i|>\varepsilon\}}\big) \xrightarrow{\mathbb{P}} 0,\\
&\sum_{i=1}^{N^n_t} \mathbb{E}^n_{i-1}\big(\zeta^n_i \Delta^n_i M\big) \xrightarrow{\mathbb{P}} 0
\end{aligned}\right\} \tag{5.16}$$

for all t, and $\varepsilon > 0$, and all $(\mathcal{F}^0_t)$-martingales $M \in \mathcal{M}$ (notation of the previous proof).

The third property above is proved like in (b) of the previous proof, the left side being indeed identically 0. For the second property we also follow the previous proof: namely, $m_p|\sigma_s|^p G(q+1)^n_s/G(1)^n_s \le C$ for some constant C, and we take $\theta \in (0, (\varepsilon/2C)^{q+1}]$. We may assume that $r_n \ge 1$. Then if $r_n^{q+1/2}(\pi^n_t)^{q+1} \le \theta$, we have $r_n^{1/2}\pi^n_t \le r_n^{(q+1/2)/(q+1)}\pi^n_t \le \theta^{1/(q+1)} \le \varepsilon/2C$; hence if further $|\zeta^n_i| > \varepsilon$ we must have $|\beta^n_i|^p > \varepsilon/2\theta$, and also $|\zeta^n_i| \le 2r_n^{q+1/2}\Delta(n,i)^{q+1}|\beta^n_i|^p$. Therefore we have the same string of inequalities as in (5.10), and we conclude the second property (5.16) in the same way.

Finally, exactly as for (4.6), and under the same conditions on H^n and H, (2.11) yields

$$\frac{1}{r_n}\sum_{i=1}^{N^n_t} H^n_{T(n,i-1)} \xrightarrow{\text{u.c.p.}} \int_0^t \frac{H_s}{G(1)_s}\,\mathrm{d}s.$$

Recalling that $a(2q+2) = G(2q+2)/G(1)$ and that $G(p)^n \xrightarrow{\text{u.c.p.}} G(p)$ for $p \in [0, 2q+2]$, we then readily deduce the first property in (5.16) from the second one in (5.15). □

6. Proof of the main theorems

6.1. *Preliminaries*

Before starting the proofs, we strengthen the assumptions (A), and (B) as follows:

***Assumption* (SA).** (A) *holds, and for some constant C we have*:

$$\big|\sigma_t(\omega)\big| + \big|b_t(\omega)\big| \le C. \tag{6.1}$$

***Assumption* (SB).** (B) *holds, and for some constant C we have*

$$\big|\sigma_t(\omega)\big| + \big|b_t(\omega)\big| + \big|\widetilde{\sigma}_t(\omega)\big| + \big|\widetilde{b}_t(\omega)\big| + \big|a'_t(\omega)\big| + a_t(\omega) \le C. \tag{6.2}$$

Under (SB) we can replace (2.1) by

$$\sigma_t = \sigma_0 + \int_0^t \widetilde{b}_s \, \mathrm{d}s + \int_0^t \widetilde{\sigma}_s \, \mathrm{d}W_s + M_t, \tag{6.3}$$

the only difference with (2.1) being that M has still bounded jumps, but not necessarily bounded by 1 (of course, M is not the same in (2.1) and in (6.3)).

It is then standard, by a suitable localization procedure (see Section 6.3 in [10] for example), and upon using also Lemma 4.1, that it is enough to prove the theorems when (A), (B), (D(q)) and (E(q)), according to the case, are replaced by (SA), (SB), (SD(q)) and (SE(q)), and for Theorem 3.3 we may even assume that the Itô semimartingales $G(p)$ for $p \in [1, 2q+2]$ have coefficients satisfying the same boundedness assumptions than those of σ_t in (SB).

We end these preliminaries with some estimates. Recall (5.1) for β_i^n, and set

$$Z_i^n = 1 + \left|\beta_i^n\right| + \left|\Delta_i^n X\right| / \sqrt{\Delta(n,i)}, \qquad \xi_i^n = \Delta_i^n X - \sqrt{\Delta(n,i)} \beta_i^n. \tag{6.4}$$

For any càdlàg process H, possibly multi-dimensional, we set

$$\alpha(H)_i^n = \sup_{s \in I(n,i)} \|H_s - H_{T(n,i-1)}\|^2. \tag{6.5}$$

Then classical estimates, see e.g. [10], and in view of (C), give for all $r \geq 2$ and on the set $\{i \leq N_t^n\}$:

$$\left.\begin{array}{lll} (\mathrm{SA}) & \Rightarrow & \mathbb{E}_{i-1}^{\prime n}\big((Z_i^n)^r\big) \leq K_r, \\ (\mathrm{SA}) & \Rightarrow & \mathbb{E}_{i-1}^{\prime n}\big(\left|\xi_i^n\right|^r\big) \leq \Delta(n,i)^{r/2}\big(K_r \mathbb{E}_{i-1}^{\prime n}\big(\alpha(\sigma)_i^n\big) + \Delta(n,i)^{r/2}\big) \\ & & \qquad \leq K_{r,t} \Delta(n,i)^{r/2} \quad \text{on } \big\{i \leq N_t^n\big\}, \\ (\mathrm{SB}) & \Rightarrow & \mathbb{E}_{i-1}^{\prime n}\big(\left|\xi_i^n\right|^r\big) \leq K_r \Delta(n,i)^{1+r/2} \end{array}\right\} \tag{6.6}$$

(note that $\{i \leq N_t^n\} \in \mathcal{G}_{i-1}^{\prime n}$; the second inequality in the second display above comes from the boundedness of σ and $\Delta(n,i) \leq t$ if $i \leq N_t^n$).

Lemma 6.1. *For H a bounded càdlàg process and $r > 0$ and $p \geq 0$, we set*

$$v(H,p,r)_i^n = r_n^{p-1} \Delta(n,i)^p \Big(\Delta(n,i)^r + \sqrt{\mathbb{E}_{i-1}^{\prime n}\big(\alpha_i^n(H)\big)}\Big). \tag{6.7}$$

Then $\sum_{i=1}^{N_t^n} v(H,p,r)_i^n \xrightarrow{\mathbb{P}} 0$ as soon as (SD(p)) holds, or (SD(q)) holds for some $q \geq p$ when $p > 1$, or (SD(q)) holds for some $q \in [0,p)$ when $p < 1$.

Proof. We first observe that the sequence $A(p)_t^n$ is uniformly integrable: this is obvious under (SD(p)). If $1 < p < q$ and (SD(q)) holds, the sequence $A(q)_t^n$ is uniformly integrable and (2.5) implies $A(p)_t^n \leq t^{(q-p)/(q-1)} \times (A(q)_t^n)^{(p-1)/(q-1)}$, and $\frac{p-1}{q-1} < 1$, so we obtain the uniform integrability of the sequence $A(p)_t^n$. When $0 \leq q < p < 1$ the same argument, upon using $A(p)_t^n \leq t^{(p-q)/(1-q)} (A(q)_t^n)^{(1-p)/(1-q)}$, gives the uniform integrability of the sequence $A(p)_t^n$ again.

Letting $\overline{V}_t^n = \sum_{i=1}^{N_t^n} \overline{v}_i^n$ and $\widetilde{V}_t^n = \sum_{i=1}^{N_t^n} \widetilde{v}_i^n$, where

$$\overline{v}_i^n = r_n^{p-1} \Delta(n,i)^{p+r}, \qquad \widetilde{v}_i^n = r_n^{p-1} \Delta(n,i)^p \sqrt{\mathbb{E}_{i-1}^{\prime n}\big(\alpha_i^n(H)\big)},$$

it suffices to prove $\overline{V}_t^n \xrightarrow{\mathbb{P}} 0$ and $\widetilde{V}_t^n \xrightarrow{\mathbb{P}} 0$.

First, the inequality $\overline{V}_t^n \leq A(p)_t^n (\pi_t^n)^r$ is obvious. Since $\pi_t^n \leq t$ and $\pi_t^n \xrightarrow{\mathbb{P}} 0$ and the sequence $A(p)_t^n$ is uniformly integrable, we deduce that $\overline{V}_t^n \to 0$ in $\mathbb{L}^1$. Second, we set

$$w_i^n = r_n^{p-1} \Delta(n,i)^p \mathbb{E}_{i-1}^{\prime n}\big(\alpha_i^n(H)\big), \qquad w_i^{\prime n} = r_n^{p-1} \Delta(n,i)^p \alpha_i^n(H)$$

and $Y_n = \sum_{i=1}^{N_t^n} w_i^n$ and $Y'_n = \sum_{i=1}^{N_t^n} w_i'^n$. The Cauchy–Schwarz inequality yields $(\widetilde{V}_t^n)^2 \le A(p)_t^n Y_n$, thus for $\widetilde{V}_t^n \xrightarrow{\mathbb{P}} 0$ it suffices to prove that $\mathbb{E}(Y_n) \to 0$. Since $w_i^n = \mathbb{E}_{i-1}'^n(w_i'^n)$ we have

$$\mathbb{E}(Y_n) = \sum_{i\ge 1} \mathbb{E}\big(w_i^n 1_{\{i\le N_t^n\}}\big) = \sum_{i\ge 1} \mathbb{E}\big(w_i'^n 1_{\{i\le N_t^n\}}\big) = \mathbb{E}\big(Y'_n\big)$$

because, as seen before, $\{i \le N_t^n\} \in \mathcal{G}_{i-1}'^n$. Therefore we are left to prove $\mathbb{E}(Y'_n) \to 0$.

Since H is càdlàg bounded, $\alpha(H)_i^n \le K$ for all i and, if $\theta > 0$ and M_θ is the number of jumps of H of size bigger than $\theta/2$ on $[0,t]$, for all n large enough we have $\alpha(H)_i^n \le \theta^2$ for all $i \le N_t^n$ except at most M_θ values of i (corresponding to the intervals $I(n,i)$ containing such a "large" jump). In other words, $Y'_n \le \theta^2 A(p)_t^n + K M_\theta r_n^{p-1}(\pi_t^n)^p$ outside a set Ω_n tending to $\varnothing$, and thus

$$\mathbb{P}\big(Y'_n < \rho\big) \le \frac{\theta^2}{\rho}\mathbb{E}\big(A(p)_t^n\big) + \mathbb{P}\big(K M_\theta r_n^{p-1}\big(\pi_t^n\big)^p > \rho\big) + \mathbb{P}(\Omega_n)$$

for all ρ. In view of (5.9) and of the uniform integrability of the sequence $A(p)_t^n$, plus $\Omega_n \to \varnothing$, by choosing first θ small and then n large, we deduce that the above goes to 0 as $n \to \infty$. Finally, we have for all $\rho, L > 0$, and since $Y'_n \le K A(p)_t^n$:

$$\mathbb{E}\big(Y'_n\big) \le \rho + KL\mathbb{P}\big(Y'_n > \rho\big) + K\mathbb{E}\big(A(p)_t^n 1_{\{A(p)_t^n > L\}}\big) \le \rho + KL\mathbb{P}\big(Y'_n > \rho\big) + \Phi(L),$$

where $\Phi(L) \to 0$ as $L \to \infty$ (apply again the uniform integrability of the sequence $A(p)_t^n$ for the last inequality). We deduce $\mathbb{E}(Y'_n) \to 0$ by choosing first ρ small, then L large, then n large. $\square$

6.2. *Proof of Theorem* 3.1

Here we fix $p \ge 1$ and $q \ge 0$, and from what precedes we may assume (SA) and (SD($q+1$)). With the notation of (3.2), we have

$$\left.\begin{aligned}
&r_n^q V^n(p,q) - V(p,q) = F^n(1) + F^n(2) + F^n(3), \quad \text{where}\\
&\quad F^n(1)_t = \sum_{i=1}^{N_t^n} \zeta(1)_i^n, \quad \zeta(1)_i^n = r_n^q \Delta(n,i)^{q+1}\big(\Delta(n,i)^{-p/2}\big|\Delta_i^n X\big|^p - \big|\beta_i^n\big|^p\big),\\
&\quad F^n(2)_t = \sum_{i=1}^{N_t^n} \zeta(2)_i^n, \quad \zeta(2)_i^n = r_n^q \Delta(n,i)^{q+1}\big(\big|\beta_i^n\big|^p - m_p|\sigma_{T(n,i-1)}|^p\big),\\
&\quad F^n(3)_t = m_p\left(r_n^q \sum_{i=1}^{N_t^n} \Delta(n,i)^{q+1}|\sigma_{T(n,i-1)}|^p - \int_0^t a(q+1)_s|\sigma_s|^p\,\mathrm{d}s\right).
\end{aligned}\right\} \tag{6.8}$$

Observe that $\sqrt{r_n}F^n(2) = U^n(p,q)$, as given by (5.2). Then it is enough to prove that for $j = 1,2,3$:

$$F^n(j) \xrightarrow{\text{u.c.p.}} 0. \tag{6.9}$$

We have $|\zeta(1)_i^n| \le K r_n^q \Delta(n,i)^{q+1/2}(Z_i^n)^{p-1}|\xi_i^n|$ because $p \ge 1$, hence by (6.6) and Hölder's inequality (with the exponents $a = b = 2$ when $k = 1$, and when $k = 2$ under (SA), and with the exponents $a = 4/3$ and $b = 4$ when $k = 2$ under (SB)) we get for $k = 1,2$:

$$\mathbb{E}_{i-1}'^n\big(\big|\zeta(1)_i^n\big|^k\big) \le \begin{cases} K_k r_n^{kq}\Delta(n,i)^{k(q+1)}\big(\Delta(n,i)^{k/2} + \sqrt{\mathbb{E}_{i-1}'^n\big(\alpha(\sigma)_i^n\big)}\big), & \text{under (SA)},\\ K_k r_n^{kq}\Delta(n,i)^{kq+1+(k+1)/4}, & \text{under (SB)}.\end{cases} \tag{6.10}$$

Now, $\zeta(1)_i^n$ is $\mathcal{G}_i^n$-measurable, whereas N_t^n is a finite $(\mathcal{G}_i'^n)$-stopping time. Hence by Lenglart's inequality (see e.g. [13], Lemma I.3.30), in order to get (6.9) for $j=1$ it is enough to prove that $\sum_{i=1}^{N_t^n} \mathbb{E}_{i-1}'^n(|\zeta(1)_i^n|) \xrightarrow{\mathbb{P}} 0$. By virtue of (6.10), we have $\mathbb{E}_{i-1}'^n(|\zeta(1)_i^n|) \le K v(\sigma, q+1, 1/2)_i^n$, as given by (6.7), and we deduce (6.9) for $j=1$ from Lemma 6.1.

Next, the variables $\zeta(2)_i^n$ are martingale increments, relative to the discrete-time filtration $(\mathcal{G}_i'^n)$, and $\mathbb{E}_{i-1}'^n(|\zeta(2)_i^n|^2) \le K r_n^{2q} \Delta(n,i)^{2q+2}$. Then a well known result on martingale differences, using again Lenglart's inequality, yields that for proving (6.9) for $j=2$ it is enough to show that $w_n(t) := r_n^{2q} \sum_{i=1}^{N_t^n} \Delta(n,i)^{2q+2} \xrightarrow{\mathbb{P}} 0$ for all t. Observing that $w_n(t) \le r_n^q (\pi_t^n)^{q+1} A(q+1)_t^n$, we deduce $w_n(t) \xrightarrow{\mathbb{P}} 0$ from the fact that the sequence $A(q+1)_t^n$ is bounded in probability, and from $r_n^q (\pi_t^n)^{q+1} \xrightarrow{\mathbb{P}} 0$, which in turn follows from (5.9) with $r=q+1$.

Finally, (6.9) for $j=3$ follows from (5.11).

6.3. *Proof of Theorem* 3.2

Step 1. By localization again, it is enough to prove the theorem under the assumptions (SA) when $p=2$ and (SB) otherwise, and (SD(2)) when $q=0$, and (SD($q+1$)), (SD($2q+2$)) and (D$'$($q+1$)), when $q>0$. In view of (6.8) and of Proposition 5.1, it is enough to prove that $\sqrt{r_n} F^n(j) \xrightarrow{\text{u.c.p.}} 0$ for $j=1$ and $j=3$, and this amounts to the following three convergences (for all t):

$$\sqrt{r_n} \sum_{i=1}^{N_t^n} \mathbb{E}_{i-1}'^n\bigl(\zeta(1)_i^n\bigr) \xrightarrow{\text{u.c.p.}} 0, \tag{6.11}$$

$$\sqrt{r_n} \sum_{i=1}^{N_t^n} \bigl(\zeta(1)_i^n - \mathbb{E}_{i-1}'^n\bigl(\zeta(1)_i^n\bigr)\bigr) \xrightarrow{\text{u.c.p.}} 0, \tag{6.12}$$

$$\sqrt{r_n} \left(\sum_{i=1}^{N_t^n} r_n^q \Delta(n,i)^{q+1} |\sigma_{T(n,i-1)}|^p - \int_0^t |\sigma_s|^p a(q+1)_s \, ds \right) \xrightarrow{\text{u.c.p.}} 0. \tag{6.13}$$

Step 2. A martingale argument shows that for (6.12) it is enough to show that

$$r_n \sum_{i=1}^{N_t^n} \mathbb{E}_{i-1}'^n\bigl(\bigl|\zeta(1)_i^n\bigr|^2\bigr) \xrightarrow{\mathbb{P}} 0.$$

Since σ is bounded, the first part of (6.10) yields $r_n \mathbb{E}_{i-1}'^n(|\zeta(1)_i^n|^2) \le K v(0, 2q+1, 1)_i^n$, so the above convergence follows from Lemma 6.1.

Step 3. Next, we prove (6.11) under (SB), in essentially the same way as in [10]. Remembering (6.3) and (6.4), we write $\xi_i^n = \Delta_i^n X - \sqrt{\Delta(n,i)} \beta_i^n$ as $\xi_i^n = \widehat{\xi}_i^n + \widetilde{\xi}_i^n$, where

$$\begin{aligned}
\widehat{\xi}_i^n &= \int_{I(n,i)} (b_s - b_{T(n,i-1)}) \, ds + \int_{I(n,i)} \left(\int_{T(n,i-1)}^s \bigl(\widetilde{b}_u \, du + (\widetilde{\sigma}_u - \widetilde{\sigma}_{T(n,i-1)}) \, dW_u\bigr) \right) dW_s, \\
\widetilde{\xi}_i^n &= \widetilde{\xi}(1)_i^n + \widetilde{\xi}(2)_i^n, \quad \widetilde{\xi}(1)_i^n = \int_{I(n,i)} (M_s - M_{T(n,i-1)}) \, dW_s, \\
\widetilde{\xi}(2)_i^n &= b_{T(n,i-1)} \Delta(n,i) + \int_{I(n,i)} \widetilde{\sigma}_{T(n,i-1)} (W_s - W_{T(n,i-1)}) \, dW_s.
\end{aligned}$$

Then, using (SB), we obtain as in [10] that

$$\left.\begin{array}{l}
\bullet\ r \ge 2 \Rightarrow \mathbb{E}_{i-1}'^n(|\widehat{\xi}_i^n|^r) + \mathbb{E}_{i-1}'^n(|\widetilde{\xi}_i^n|^r) \le K_r \Delta(n,i)^{1+r/2}, \\
\bullet\ \mathbb{E}_{i-1}'^n(|\widehat{\xi}_i^n|^2) \le K \Delta(n,i)^2 (\Delta(n,i) + \mathbb{E}_{i-1}'^n(\alpha(b)_i^n + \alpha(\widetilde{\sigma})_i^n)), \\
\bullet\ h \text{ an odd function with polynomial growth } \Rightarrow \mathbb{E}_{i-1}'^n(\widetilde{\xi}_i^n h(\beta_i^n)) = 0
\end{array}\right\} \tag{6.14}$$

(the last property above is nearly obvious if we replace $\widetilde{\xi}_i^n$ by $\widetilde{\xi}(2)_i^n$. It also holds when we replace $\widetilde{\xi}_i^n$ by $\widetilde{\xi}(1)_i^n$ because $\widetilde{\xi}(2)_i^n = (M_{T(n,i)} - M_{T(n,i-1)})(W_{T(n,i)} - W_{T(n,i-1)}) - \int_{I(n,i)} (W_s - W_{T(n,i-1)})\,\mathrm{d}M_s$, hence $\widetilde{\xi}(1)_i^n h(\beta_i^n)$ can be written as the sum of products of a stochastic integral w.r.t. W, times a stochastic integral w.r.t. M, and we conclude by the orthogonality of M and W).

Next, a Taylor expansion and the differentiability of $f(x) = |x|^p$ (recall $p \geq 2$, we have $f'(\lambda x) = \lambda^{p-1} f'(x)$ and observe that $f''(\lambda x) = \lambda^{p-2} f''(x)$) give

$$\bigl|\zeta(1)_i^n - r_n^q \Delta(n,i)^{q+1/2} f'\bigl(\beta_i^n\bigr)\xi_i^n\bigr| \leq K r_n^q \Delta(n,i)^q \bigl(Z_i^n\bigr)^{p-2}\bigl|\xi_i^n\bigr|^2.$$

The last part of (6.14) and the fact that f is even, hence f' is odd, imply that $\mathbb{E}'^n_{i-1}(f'(\beta_i^n)\widetilde{\xi}_i^n) = 0$. Moreover we have $|f'(\beta_{i+1}^n)| \leq K(Z_{i+1}^n)^{p-1}$. Therefore

$$\bigl|\mathbb{E}'^n_{i-1}\bigl(\zeta(1)_i^n\bigr)\bigr| \leq K r_n^q \Delta(n,i)^q \mathbb{E}'^n_{i-1}\bigl(\bigl(Z_i^n\bigr)^{p-2}\bigl|\xi_i^n\bigr|^2 + \sqrt{\Delta(n,i)}\bigl(Z_{i+1}^n\bigr)^{p-1}\bigl|\widehat{\xi}_i^n\bigr|\bigr). \tag{6.15}$$

We deduce from the first part of (6.6), the third part of the same (with Hölder's inequality with exponents $a = 4/3$ and $b = 4$, recall that (SB) holds here) and from the second part of (6.14) with the Cauchy–Schwarz inequality, that

$$\bigl|\mathbb{E}'^n_{i-1}\bigl(\zeta(1)_i^n\bigr)\bigr| \leq K r_n^q \Delta(n,i)^{q+3/2}\bigl(\Delta(n,i)^{1/4} + \Delta(n,i)^{1/2} + \sqrt{\mathbb{E}'^n_{i-1}\bigl(\alpha(b)_i^n + \alpha(\widetilde{\sigma})_i^n\bigr)}\bigr).$$

Recall also that $\Delta(n,i) \leq t$ if $i \leq N_t^n$. Then with H being the two-dimensional process with components b and $\widetilde{\sigma}$, we see that $\sqrt{r_n}|\mathbb{E}^n_{i-1}(\zeta(1)_i^n)| \leq K_t v(H, q+3/2, 1/4)_i^n$ with the notation (6.7), and if $i \leq N_t^n$. Since $1 < q+3/2 < 2q+2$ and (SD$(2q+2)$) holds, we then deduce (6.11) from Lemma 6.1.

Step 4. When $q > 0$, the property (6.13) is a trivial consequence of Assumption (D$'(q+1)$).

Step 5. In this step, we prove (6.13) under (SB), when $q = 0$. This follows from the next two properties:

$$\sqrt{r_n}\int_{T(n,N_t^n)}^t |\sigma_s|^p\,\mathrm{d}s \stackrel{\text{u.c.p.}}{\longrightarrow} 0, \tag{6.16}$$

$$\sum_{i=1}^{N_t^n} \mu_i^n \stackrel{\text{u.c.p.}}{\longrightarrow} 0, \quad \text{where } \mu_i^n = \sqrt{r_n}\int_{I(n,i)} \bigl(|\sigma_s|^p - |\sigma_{T(n,i-1)}|^p\bigr)\,\mathrm{d}s. \tag{6.17}$$

Equation (6.16) readily follows from (5.9). The proof of (6.17) is similar to [10]: using the C^1 property of $f(x) = |x|^p$ and (6.3), and with $g(x,y) = f(x) - f(y) - f'(y)(x-y)$, we have $\mu_i^n = \mu(1)_i^n + \mu(2)_i^n + \mu(3)_i^n$, where

$$\mu(1)_i^n = \sqrt{r_n}\int_{I(n,i)} f'(\sigma_s)\,\mathrm{d}s \int_{T(n,i-1)}^s \widetilde{b}_u\,\mathrm{d}u,$$

$$\mu(2)_i^n = \sqrt{r_n}\int_{I(n,i)} f'(\sigma_s)\,\mathrm{d}s \int_{T(n,i-1)}^s (\widetilde{\sigma}_u\,\mathrm{d}W_u + \mathrm{d}M_u),$$

$$\mu(3)_i^n = \sqrt{r_n}\int_{I(n,i)} g(\sigma_{T(n,i)}, \sigma_s)\,\mathrm{d}s.$$

If $|x|, |y| \leq C$ we have $|f'(x)| \leq K_C$ and $|g(x,y)| \leq K_C|x-y|^2$. Therefore by (SB), which also implies that $\mathbb{E}'^n_{i-1}((\sigma_s - \sigma_{T(n,i-1)})^2) \leq K\Delta(n,i)$ on the set $\{s \in I(n,i)\}$, we get for $j = 1$ and $j = 3$:

$$\mathbb{E}\Biggl(\sum_{i=1}^{N_t^n}\bigl|\mu(j)_i^n\bigr|\Biggr) \leq K\sqrt{r_n}\mathbb{E}\Biggl(\sum_{i=1}^{N_t^n}\Delta(n,i)^2\Biggr) \leq \frac{K}{\sqrt{r_n}}\mathbb{E}\bigl(A(2)_t^n\bigr).$$

Finally $\mathbb{E}'^n_{i-1}(\mu(2)_i^n) = 0$ and $\mathbb{E}'^n_{i-1}(|\mu(2)_i^n|^2) \leq K r_n \Delta(n,i)^3$ by (SB) again, hence

$$\mathbb{E}\Biggl(\Biggl(\sum_{i=1}^{N_t^n}\mu(2)_i^n\Biggr)^2\Biggr) \leq K\mathbb{E}\bigl(A(2)_t^n \pi_t^n\bigr)$$

which goes to 0 by (SD(2)) and (1.5). This finishes the proof of (6.17).

Step 6. It remains to consider the case $p=2$ when we have (SA) instead of (SB): we need to prove (6.11), and also (6.13) when $q=0$. The proof is similar to the proof given in [12], and so we only give a sketch below.

We first assume that σ and b are piecewise constant, that is of the form

$$\sigma_t = \sum_{j\geq 1} \sigma_{t_{j-1}} 1_{[t_{j-1},t_j)}(t), \qquad b_t = \sum_{j\geq 1} b_{t_{j-1}} 1_{[t_{j-1},t_j)}(t), \tag{6.18}$$

where t_j is a strictly increasing sequence of reals, starting at $t_0=0$ and going to infinity. In this case, (6.12) has been proved in Step 2, and (6.13) holds from Step 4 when $q>0$, and when $q=0$ it is an obvious consequence of $\sqrt{r_n}\pi^n_t \xrightarrow{\mathbb{P}} 0$ (apply (5.9) with $r=2$). Moreover, by integration by parts the left side of (6.11) is

$$\sqrt{r_n}\sum_{j:t_j\leq t} \eta^n_j + \sqrt{r_n}\sum_{i=1}^{N^n_t} \eta'^n_i,$$

where $|\eta^n_j| \leq K r^q_n (\pi^n_t)^{q+1}$ and $|\eta'^n_i| \leq K r^q_n \Delta(n,i)^{q+2}$. Therefore if $s\leq t$ the left side of (6.11) is smaller than $K_t r_n^{q+1/2}(\pi^n_t)^{q+1} + K\sum_{i=1}^{N^n_t} v^n_i$, where v^n_i is given by (6.7) with $k=1$ and $H=0$. Then we deduce (6.11) from (5.9) for $r=2q+2$ and from Lemma 6.1. Therefore we have proved that under (6.18), Theorem 3.2 holds.

Next, we assume (SA) only, and we approximate σ and b with piecewise constants processes $\sigma(m)$ and $b(m)$, which are bounded uniformly in m, and in the sense that

$$\int_0^t \bigl(\bigl|b(m)_s - b_s\bigr| + \bigl|\sigma(m)_s - \sigma_s\bigr|\bigr)\,\mathrm{d}s \xrightarrow{\mathbb{P}} 0 \tag{6.19}$$

for all t, as $m\to\infty$. We introduce the semimartingales

$$X(m)_t = X_0 + \int_0^t b(m)_s\,\mathrm{d}s + \int_0^t \sigma(m)_s\,\mathrm{d}W_s, \tag{6.20}$$

and the associated processes $V^{n,m}(p,q)$. From what precedes, for each m the processes $\overline{V}^{n,m}_t = \sqrt{r_n}(r^q_n V^{n,m}(2,q)_t - \int_0^t \sigma(m)^2_s a(q+1)_s\,\mathrm{d}s)$ converge $\mathcal{F}^0$-stably in law, as $n\to\infty$, to the process $\sqrt{2}\int_0^t |\sigma(m)_s|^2\sqrt{a(2q+2)_s}\,\mathrm{d}W'_s$, and it remains to prove that

$$\int_0^t \bigl|\sigma(m)_s\bigr|^2\sqrt{a(2q+2)_s}\,\mathrm{d}W'_s \xrightarrow{\text{u.c.p.}} \int_0^t |\sigma_s|^2\sqrt{a(2q+2)_s}\,\mathrm{d}W'_s \quad \text{as } m\to\infty, \tag{6.21}$$

$$\lim_{m\to\infty}\limsup_{n\to\infty}\mathbb{E}\Bigl(\sup_{s\leq t}\bigl|\overline{V}^{n,m}_s - \overline{V}^n_s\bigr|\Bigr) = 0, \tag{6.22}$$

where $\overline{V}^n_t = \sqrt{r_n}(r^q_n V^n(2,q)_t - \int_0^t \sigma^2_s a(q+1)_s\,\mathrm{d}s)$. Equation (6.21) readily follows from (6.19), which also implies (6.22) by the same argument as in [12]: the (tedious) details are left to the reader here.

6.4. *Proof of Theorem* 3.3

The proof is basically the same as for Theorem 3.2. We can assume (SB) and (SE($2q+2$)) and (SD($2q+2$)) and (SD($q+1$)) because (D($q+1$)) holds by Lemma 2.2. Instead of the decomposition (6.8) we write, with the same $F^n(1)$ as in (6.8) and the notation (5.12):

$$\sqrt{r_n}\bigl(r^q_n V^n(p,q) - V(p,q)\bigr) = \sqrt{r_n}F^n(1) + U'^n(p,q) + m_p F'^n,$$

where

$$F'^n_t = \sqrt{r_n}\Biggl(\sum_{i=1}^{N^n_t} |\sigma_{T(n,i-1)}|^p \frac{G(q+1)^n_{T(n,i-1)}}{G(1)^n_{T(n,i-1)}}\Delta(n,i) - \int_0^t a(q+1)_s|\sigma_s|^p\,\mathrm{d}s\Biggr).$$

Note that the previous proof yields that (6.11) and (6.12) hold, that is $\sqrt{r_n}F^n(1) \stackrel{\text{u.c.p.}}{\longrightarrow} 0$. Then, in view of Proposition 5.2, it is enough to prove that $F'^n \stackrel{\text{u.c.p.}}{\longrightarrow} 0$.

Due to (E$'(q+1)$), to the boundedness of $G(q+1)^n$ and of $1/G(1)^n$ (uniform in n), and to the form of $a(q+1)$ recalled above, for getting $F'^n \stackrel{\text{u.c.p.}}{\longrightarrow} 0$ it is enough to prove the following:

$$\sqrt{r_n}\left(\sum_{i=1}^{N_t^n} \Delta(n,i) H_{T(n,i-1)} - \int_0^t H_s \,\mathrm{d}s\right) \stackrel{\text{u.c.p.}}{\longrightarrow} 0,$$

where H is the process $H = |\sigma|^p G(q+1)/G(1)$. Remember that under our assumptions σ, $G(q+1)$ and $G(1)$ are Itô semimartingale with bounded coefficients (in the sense of (SB)), and they are bounded and $1/G(1)$ is also bounded. Then by Itô's formula the process H is also an Itô semimartingale with bounded coefficients. Then exactly the same proof than in Step 5 of the previous proof shows that $F'^n \stackrel{\text{u.c.p.}}{\longrightarrow} 0$, and the proof is finished.

6.5. *Proof of Theorem* 3.10

Theorem 3.10 trivially follows from Theorems 3.1 and 3.2, upon recalling the following fact: if U_n and A_n are random variables defined on $(\Omega, \mathcal{F}, \mathbb{P})$, and if U_n converges $\mathcal{F}^0$-stably to a limit U (defined on an extension of the space) and $A_n \stackrel{\widehat{\mathbb{P}}}{\longrightarrow} A$, and if A is $\mathcal{F}^0$-measurable, then (U_n, A_n) converges in law to (U, A).

Acknowledgments

The authors are grateful for the careful reading of a referee who pointed out several mistakes in two successive drafts of this paper.

References

[1] O. E. Barndorff-Nielsen and N. Shephard. Power variation and time change. *Theory Probab. Appl.* **50** (2005) 1–15. MR2222740
[2] O. E. Barndorff-Nielsen, P. R. Hansen, A. Lunde and N. Shephard. Multivariate realised kernels: Consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading. *J. Econometrics* **162** (2011) 149–169.
[3] D. Duffie and P. Glynn. Estimation of continuous-time Markov processes sampled at random time intervals. *Econometrica* **72** (2004) 1773–1808. MR2095532
[4] M. Fukasawa. Central limit theorem for the realized volatility based on tick time sampling. *Finance Stoch.* **34** (2010) 209–233. MR2607763
[5] V. Genon-Catalot and J. Jacod. Estimation of the diffusion coefficient for diffusion processes: Random sampling. *Scand. J. Stat.* **21** (1994) 193–221. MR1292636
[6] J. E. Griffin and R. C. Oomen. Covariance measurement in the presence of non-synchronous trading and market microstructure noise. *J. Econometrics* **160** (2011) 58–68. MR2745867
[7] T. Hayashi and N. Yoshida. On covariance estimation of nonsynchronously observed diffusion processes. *Bernoulli* **11** (2005) 359–379. MR2132731
[8] T. Hayashi and N. Yoshida. Nonsynchronous covariance estimator and limit theorem. Preprint, 2006.
[9] T. Hayashi and N. Yoshida. Asymptotic normality of a covariance estimator for nonsynchronously observed diffusion processes. *Ann. Inst. Statist. Math.* **60** (2008) 357–396. MR2403524
[10] J. Jacod. Statistics and high frequency data. In *SEMSTAT Seminar*, 2007. To appear.
[11] J. Jacod. Asymptotic properties of realized power variations and related functionals of semimartingales. *Stochastic Process. Appl.* **118** (2008) 517–559. MR2394762
[12] J. Jacod and P. Protter. Asymptotic error distributions for the Euler method for stochastic differential equations. *Ann. Probab.* **26** (1998) 267–307. MR1617049
[13] J. Jacod and A. Shiryaev. *Limit Theorems for Stochastic Processes*, 2d edition. Springer, Berlin, 2003. MR1943877
[14] P. Malliavin and M. E. Mancino. Fourier series methods for measurement of multivariate volatilities. *Finance Stoch.* **6** (2002) 49–61. MR1885583
[15] P. A. Mykland and L. Zhang. ANOVA for diffusions and Itô processes. *Ann. Statist.* **34** (2006) 1931–1963. MR2283722
[16] P. C. B. Phillips and J. Yu. Information loss in volatility measurement with flat price trading. Preprint, 2008.